ST/ESA/SER.A/61

Methods of measuring the impact of family planning programmes on fertility: problems and issues

UNITED NATIONS

ST/ESA/SER.A/61

Department of Economic and Social Affairs

POPULATION STUDIES, No. 61

Methods of measuring the impact of family planning programmes on fertility: problems and issues

A project of the Population Division of the Department of Economic and Social Affairs of the United Nations Secretariat, in collaboration with the Committee on Demographic Aspects of Family Planning Programmes of the International Union for the Scientific Study of Population

UNITED NATIONS

New York, 1978

NOTE

ST/ESA/SER.A/61

UNITED NATIONS PUBLICATION

Sales No. E.78.XIII.2

Price: $U.S. 11.00
(or equivalent in other currencies)

FOREWORD

This publication is a project of the Population Division of the Department of Economic and Social Affairs of the United Nations Secretariat, undertaken in collaboration with the Committee on Demographic Aspects of Family Planning Programmes of the International Union for the Scientific Study of Population.

The entire publication, which is presented in three parts, is related to the work of the Expert Group Meeting on Methods of Measuring the Impact of Family Planning Programmes on Fertility, convened at Geneva, Switzerland, from 20 to 27 April 1976. The Meeting and substantive preparations for it were made possible by a grant from the United Nations Fund for Population Activities.

Part one comprises a background paper prepared by the United Nations Secretariat and three country studies developed by national experts for the Expert Group Meeting. The report of that meeting is contained in part two. Part three is composed of statements on methods of measuring the impact of family planning programmes on fertility, submitted by members of the Expert Group.

CONTENTS

Explanatory notes

The following symbols have been used in the tables throughout the publication:

Three dots (. . .) indicate that data are not available or are not separately reported

A dash (—) indicates that the amount is nil or negligible

A blank in a table indicates that the item is not applicable

A minus sign (−) indicates a deficit or decrease, except as indicated

A full stop (.) is used to indicate decimals

A slash (/) indicates a crop year or financial year, e.g., 1974/75

Use of a hyphen (-) between dates representing years, e.g., 1973-1975, signifies the full period involved, including the beginning and end years.

Reference to dollars ($) indicates United States dollars, unless otherwise stated.

Details and percentages in tables do not necessarily add to totals, because of rounding.

Part One

STUDIES OF THE MEASUREMENT OF THE IMPACT OF FAMILY PLANNING PROGRAMMES ON FERTILITY: BACKGROUND MATERIAL PREPARED FOR THE EXPERT GROUP MEETING

METHODS OF MEASURING THE IMPACT OF FAMILY PLANNING PROGRAMMES ON FERTILITY: PROBLEMS AND ISSUES*

*United Nations Secretariat***

Introduction

Purpose and scope of the study

At a time when a major portion of population policy efforts is committed to large-scale family planning programmes, there is naturally a growing concern about the results of those efforts and, hence, an increased interest in evaluative research. The Expert Group Meeting was devoted to an in-depth examination of one particular aspect of evaluation, namely, the measurement of the impact of family planning programmes on fertility. Specifically, the purpose of that meeting was to review the principal current methods of measuring programme impact on fertility and to determine which measurement method or group of methods are most appropriate in different circumstances.

Why was impact on fertility chosen for discussion? Programme impact can be assessed at various stages of implementation, by measuring changes occurring, for instance, in knowledge of birth control methods, use of birth control methods or desired family size. The interest in fertility derived at first mainly from the fact that a certain number of countries have population policies designed to reduce the rate of growth of their population and consider fertility reduction to be a major objective. In these countries, attempts to measure programme impact on fertility have been made on various occasions, but both the techniques employed and the interpretation of the results have raised a number of difficult issues.

The interest in measuring programme impact on fertility exists independently of policy objectives. Some policies favour fertility increases and others are not directed towards achieving any fertility changes. But fertility change is a plausible consequence of family planning programmes, irrespective of policy aims; and policy-makers, as well as programme administrators, may be interested, for various reasons, in the impact of such programmes on fertility. Indeed, measuring programme effects on fertility is not limited to finding out whether the programme is achieving its objectives. The population factor is often a major component of development planning; and demographic data about future population trends are generally required for the formulation of policies with respect to manpower, health, education etc. Population forecasts require the best available assumptions about future fertility trends. If family planning programmes are undertaken, planners will wish to know whether those programmes have any effects on fertility so as to be able to take such effects into account.

Measuring the effects of a family planning programme on fertility is recognized as a difficult task. If a change in fertility is believed to have occurred during a period of programme implementation, the evaluator is expected to determine what part of that change can be attributed to the programme. Even if no change in fertility is observed, the evaluator must investigate whether the unchanged fertility reflects the absence of any programme influence or whether the constant level of fertility results from the compensating effects or programme and non-programme variables. In other words, the evaluator is expected to establish whether and by how much the fertility of the population under study would have been different without the programme.

How does the evaluator undertake such measurements? The conditions under which he is working are characterized by the fact that many factors are at work in determining fertility, that their relationships are complex, that it is the complexity of these relationships rather than the number of variables involved which appear to raise the most difficult questions, and that wide gaps in knowledge still exist in understanding these interrelations. Ideally, a most rewarding method for evaluating programme impact would be one which, on the one hand, would account separately for programme and non-programme influences and, on the other hand, would weigh appropriately the influence of various variables within each category of influences. This method would, of course, require a model defining the various interrelations among variables, including a satisfactory set of assumptions whenever adequate knowledge was not available. In particular, the model would specify all the relevant variables and the form of their interrelationships, in order to elucidate direct and indirect effects, combined effects, overlapping effects, reciprocal effects etc.

With few exceptions, however, existing statistical and demographic techniques which have been adapted

---

* The original version of this paper appeared as document ESA/P/AC.7/1.

** Population Division of the Department of Economic and Social Affairs. The final draft benefited greatly from the revision by the consultant to the project, Albert I. Hermalin, Assistant Director of the Population Studies Center, University of Michigan, Ann Arbor, Michigan, United States of America.

and new methods devised for evaluation purposes do not attempt to quantify the contribution of specific programme and non-programme factors to fertility change. The main emphasis is placed on distinguishing between programme and non-programme influences by comparing the fertility observed under programme conditions with the fertility that would have been observed had the programme not been undertaken. The first subsection given below describes as simply and concisely as possible the principal current evaluation methods. Next, various difficulties and methodological issues associated with the application of those methods are outlined. The last subsection briefly states the justification and purpose of the three case studies in which the same evaluation methods are applied comparatively to individual countries.

Methods of measuring family planning programme impact on fertility

There are a number of possible ways of classifying the methods that have been proposed to examine the impact of family planning programmes on fertility. In this paper, the methods described are grouped as follows, recognizing that these categories are not mutually exclusive:

(*a*) Standardization approach;
(*b*) Trend analysis;
(*c*) Experimental designs;
(*d*) Couple-years of protection (CYP);
(*e*) Component projection approach;
(*f*) Analysis of the reproductive process;
(*g*) Regression analysis (including path analysis);
(*h*) Simulation models.

One purpose of this paper is to give a brief account of the procedures used for evaluating programme impact on fertility. In order to facilitate comparison, the description follows a standard format, referring systematically to distinguishing aspects of the approach used in each method. These various points of comparison are: (*a*) the estimating technique itself; (*b*) the programme and non-programme factors utilized; (*c*) the fertility indicators used for measuring fertility change; (*d*) the main assumptions involved; (*e*) the population covered; and (*f*) the time reference for which evaluation is made.

Standardization approach

The standardization approach as applied to measure programme impact on fertility requires two steps. The first step consists in measuring fertility at two points in time to determine whether any change has occurred during the period under study. The second step consists in trying to account for the observed change, if any, by standardizing for various non-programme components which, depending upon the fertility indicator used, may affect observed fertility without reflecting a genuine fertility change. Standardization will thus "explain" part of the observed change and the residual portion which cannot be accounted for by the standardized components will require an additional analysis. On the basis of reasonable assumptions and satisfactory evidence, all or part of the residual can thus be attributed to the family planning programme. Caution is, however, required in using this method. Factors other than those standardized for can evidently affect observed fertility, although not all of the non-programme factors can easily be taken into account. Standardization for socio-economic factors, apparently not attempted so far, might be a fruitful effort to approach this difficulty.

Trend analysis

Trend analysis, the fertility projection approach, is used to estimate, on the basis of reasonable assumptions, how the fertility of the population under study would have evolved had the family planning programme not been undertaken. This potential trend in fertility is then compared with the actual trend and an attempt to interpret the difference between the two trends can be made in order to assess the possible effects of the family planning programme. Caution has, of course, to be taken so as not to attribute to the programme a trend difference resulting from erroneous projection assumptions. The method can be applied on the aggregate level to estimate over-all country effects or, if data are available, to specific groups, such as acceptors only. In the latter case, however, additional problems arise.

Experimental designs

The experimental-design approach endeavours to compare two groups of population: one, the "experimental group", is assumed to have undergone a treatment which, in the present case, would be the family planning programme; the other, the "control group", is assumed to have the same characteristics as the experimental group, except that it was not exposed to the treatment. The fertility of each group is recorded at one or several points in time; and, assuming that the two groups are comparable except for the programme factor, the evaluator would consider any difference in fertility between the experimental and control groups as resulting from the programme. Comparability of groups or non-exposure to treatment is, however, rarely found in real social settings, so that evaluators resort to variations of the classical experimental design. In practice, researchers resort to various forms of *ex post facto* matching procedures.

Couple-years of protection

The couple-years of protection index is an estimate of the protection against pregnancy resulting from the differential use of various methods of birth control. It is used to produce a measure of programme achievement in a period, by assessing the joint impact of methods adopted, taking into account the length of time a couple is likely to be protected by each method. The CYP index can also be used to produce an estimate of prevalence of use during a period by taking into account protection resulting from past distribution as well as protection derived in the period from current distribution. From this prevalence measure, one can estimate the number of births averted on the basis of a simple translation equation of the form 1 CYP = *n* births averted, *n* varying with the fertility levels prevailing in each country. The sources of data, the quality of data and the assumptions involved in the determination of both terms of this equation define this method as very simple, but somewhat crude, whose interpretation is difficult and whose reliability is uncertain. Although this method has been recommended or used for various administrative purposes, its conclusion requires independent verification.

Component projection approach

The component projection approach is also based on data about birth control practice. The number of acceptors of a given method, the duration of use and the effectiveness of the contraception must be taken into account, either explicitly or implicitly. In addition, the fertility of these acceptors, had the family planning programme not been undertaken (potential fertility) must be estimated for the period of time under analysis. These data allow the evaluator to obtain an estimate of the number of births averted over a given time period. These estimates are usually worked out by five-year age groups on an annual basis. The summation of births averted for all ages of the women's reproductive span and for all 12-month periods studied provides the total number of births averted during the period under study. This approach is not without difficulties, both for estimating the number of continuing users and for estimating their potential fertility. The timing of adoption in relation to the woman's reproductive cycle, the switching of family planning methods and the use of abortion are factors which raise a number of additional problems. Since the pioneering work of Lee and Isbister[1] in component projection, there have been a number of new developments, the purpose of which is to examine the effects of family planning programmes on both marital age-specific fertility rates and crude birth rates. Examples of these new developments can be found in studies referred to in the selected bibliography at the end of this paper, the most recent work in this field being that published by the Economic and Social Commission for Asia and the Pacific (ESCAP).

Analysis of reproductive process

Analysis of the reproductive process has been used to estimate births averted by a segment of contraceptive use. As applied by Potter[2] and by Wolfers,[3] the mean duration of interruption of pregnancy due to the acceptors' effective use of programme contraception is estimated by the life-table technique and compared with the mean duration between births, used as an estimation of the acceptors' potential fertility. If, because of the adoption of the programme contraceptive, the acceptors remain in a non-pregnant state for an average period equivalent to the length of an average birth interval, then one birth has been averted. In addition to the average duration per birth and the length of use of the contraceptive from the life table, these methods often include adjustments for the proportion of acceptors who are sterile at acceptance and who become sterile subsequently, overlap of contraceptive use with post-partum amenorrhoea, pregnancy rates while using contraception and a time "penalty" for such pregnancies. These methods, devised for intra-uterine methods, have been adapted for other contraceptive methods. In the application of this method, caution is required in establishing assumptions dealing with special situations, such as when one contraceptive is substituted for another or when one method is supplemented by another.

Regression analysis

Multiple regression analysis, including path analysis, can also be performed if the required data are available. The method consists in determining an equation or system of equations where the dependent variable is a fertility indicator and the independent variables are programme and non-programme factors. Through such a functional model, the evaluator can attempt to calculate quantitative estimates of the weight of the various independent variables in explaining differences in the dependent variable. The relation of various family planning programme components to differences in level of fertility within a given country is interpreted by this approach as pin-pointing the effect of these components upon changes in fertility that have occurred in the past. This method is not without difficulties and involves a series of crucial steps—selection of the variables, procedure for estimating regression coefficients etc.—whose bearing on the results and their interpretation is of fundamental importance.

Simulation models

Simulation models are also used for the study of births prevented by the use of birth control methods.

---

[1] B. M. Lee and John Isbister, "The impact of birth control programs on fertility", in Bernard Berelson and others, eds., *Family Planning and Population Programs: A Review of World Development* (Chicago, Ill., University of Chicago Press, 1966), pp. 737–758.

[2] Robert G. Potter, "Application of the life-table techniques to measurement of contraceptive effectiveness", *Demography,* vol. 3, No. 2 (1966), pp. 297–304; and *idem,* "Estimating births averted in a family planning program", in S. J. Behrman, Leslie Corsa, Jr. and Ronald Freedman, eds., *Fertility and Family Planning: A World View* (Ann Arbor, Mich., University of Michigan Press, 1969), pp. 413–434.

[3] David Wolfers, "The demographic effect of a contraceptive programme", *Population Studies,* vol. XXIII, No. 1 (March 1969), pp. 111–141.

This approach is, however, largely theoretical and has as its objective the study of the effects of family planning practices under various hypothetical conditions regarding, for instance, the type and effectiveness of method used and the reproductive characteristics of the women. Research in this field has investigated births prevented by contraception, by abortion, by sterilization and by contraception supplemented by abortion.

Major methodological issues

The second section of this paper is devoted to an in-depth examination of the main difficulties which evaluators face in applying the evaluation methods described in the second paper. Its principal purpose is to generate a discussion focused on these specific issues. The presentation is, however, concerned chiefly with raising questions rather than with suggesting answers, the latter area being the responsibility of the experts participating in the meeting. Another purpose of this paper is to set the stage for an organized systematic discussion of the results of the three country studies prepared for this meeting.

The methodological issues selected for discussion have been classified, somewhat arbitrarily, as follows:

(*a*) Potential fertility;
(*b*) Data requirement problems;
(*c*) Interaction problems;
(*d*) Uncontrolled variables;
(*e*) Independence of methods;
(*f*) Cost-precision analysis.

Most of the main issues are common problems of the social sciences related to the interdependence among variables. On the other hand, potential fertility as such is an issue specific to family planning programme evaluation, although the measurement of non-events is a problem shared in many areas of evaluative research. Cost-precision problems are partially of an administrative nature. There is often a trade-off between cost and precision which may affect in one way or another the reliability of the evaluation results and conclusions. Each of these six issues is briefly commented upon below.

Potential fertility

Many methods assessing the effects of family planning programmes on fertility rely on a comparison between the actual fertility of the population under study and its potential fertility, i.e., the fertility that population might have experienced in the absence of the programme. Allied to the question of potential fertility is that of substitution of contraception from the programme for methods practised outside the programme, or taking account of the possibility that those who practise for the first time within the programme would have adopted through other means in the absence of a programme. The estimation of potential fertility, including substitution effects, raises a number of problems, notably with respect to the type of data used, the assumptions involved, the indicator selected and the computation technique applied. Some procedures are more straightforward than others; and, in all cases, an evaluation of the quality of the estimate is needed. Its validity and accuracy are important because over-estimating or underestimating potential fertility would result in an over-estimation or an underestimation of programme impact. It is true that one can never know with certainty what fertility would have been in circumstances other than those which occurred. Reasonable estimates of potential fertility, however, can usually be obtained if caution is exercised in formulating the underlying assumptions.

Data requirement problems

When data required by a particular method are not available, substitute data may be obtained by estimating the missing data, by using data from another population with a similar background, or by making other assumptions about the unavailable data etc. When available data are known to be unreliable, estimates or adjustments sometimes can be performed to bring those data in line. In other instances, data cannot be corrected, although the magnitude of the error can sometimes be assessed, as in the case of sampling error. Even when data are corrected, error may still remain. Thus, whether treated or not before being used for evaluation purposes, data may be the source of sometimes extensive uncertainties affecting the validity of the evaluation results. Furthermore, data quality is not the only ground on which results may be questioned. Certain indicators—of fertility, for instance—may be better than others for specific measurements. Failure to interpret a given indicator in the light of its limitations—a crude birth rate uncorrected for age structure—is an additional source of inaccurate evaluation. A fair appraisal of errors associated with data utilization is a prerequisite of reliable evaluation conclusions.

Interaction problems

Discriminating between programme and non-programme effects and calculating the differential impact of various factors on fertility is the essential feature of programme impact measurement procedures. A non-programme factor may affect fertility directly, or through one or more non-programme factors, or even through one or more programme factors. Likewise, programme factors can affect fertility directly or through other factors. In some cases, the effects of two variables may be additive; in other cases, their effects may be overlapping. Improper measurement of the contribution of any given variable to a fertility change may over-estimate or underestimate its influence and lead to erroneous conclusions about the impact of the programme. A satisfactory assessment of the various interaction effects of the variables utilized in the application of a given method is thus imperative for a satisfactory evaluation of programme impact.

Uncontrolled variables

The number and complexity of the interrelations between fertility and its determining factors do not allow the evaluator to utilize all the relevant variables. Many methods make a selective use of variables, giving emphasis to those most expected to provide a satisfactory measurement of programme impact. Some variables are neglected because their role is not immediately obvious, others because they are difficult to quantify, still others because the required data are not available. The output of a method application is, of course, directly dependent upon its inputs. If inappropriate variables are used, for instance, or if some important variables are not controlled for at all, the validity of the evaluation exercise may be affected to an extent which is, in many instances, not easy to determine.

Independence of methods

The measurement of programme impact on fertility is complicated not only by the recognized limitations of all current methods but by the differing conditions and circumstances involved in programme implementation in different countries. It would be of great help if some of the difficulties encountered in identifying the various influences on fertility could be overcome by using two or more methods simultaneously. This might strengthen the evidence, thus increasing the probability that the conclusions about the effects of the programme are correct. Which methods can be best used as complementary approaches to measure the impact of a family planning programme on fertility under specific conditions? It appears reasonable to assume that methods which would assess programme impact independently would afford a greater guarantee of objective evaluation. But what are independent methods? One may refer here to the items of the standard format used in describing the various methods and consider whether some of these items could be considered criteria of independence; and, if yes, which constitute the most valid basis for establishing independent verification.

Cost-precision analysis

The accuracy of the results that can be obtained from the different measurement methods is still a largely unexplored field. Evidently, some methods yield a more precise assessment of the programme effect than others and the question of method precision deserves more systematic study. How precise an estimate of programme impact should be is another question. In some cases, a bare approximation of the programme effects could be simpler, easier and less time-consuming to obtain and still be satisfactory for a specific purpose. In other cases, more accurate results would be necessary. Whenever the more precise measurement methods require more work for data gathering and corrections, more expertise, more complex techniques of evaluation and the use of electronic computers, the additional cost may impose a restriction on the degree of precision feasible. Taken together, these considerations lead to the conclusion that the precision required for a given evaluation will be, by and large, a function of both the purpose of the evaluation and the cost involved. In other words, a cost-precision analysis is required in the light of financial possibilities and evaluation objectives. So far, little attention has been devoted to this question. How does one compare the various methods in terms of accuracy? What are the alternative decision-making processes through which an acceptable cost-precision balance can be reached?

Country studies

The main objective of the three country studies was to complement the material presented in the main sections of this paper by providing concrete comparative models of method application. Practically, their conclusions are expected to point out the method or methods that present the least difficulties and yield the best results in specific circumstances. Each study is carried out with this purpose in mind and focuses primarily on: (*a*) problems that arise when an evaluation method is put to use in specified circumstances; and (*b*) a comparison of the results obtained by each method and an analysis of the probable reasons for whatever differences are found to exist. Emphasis is placed not on the description of method application, but rather on the problem areas of method application and on the comparative validity of the results. The difficulties encountered by the evaluators are discussed, whenever it is pertinent, in terms of the methodological issues described below.

The problems arising when specific family planning evaluation methods are applied in given circumstances may be considered from three viewpoints: (*a*) selection of the method; (*b*) application of the method; and (*c*) interpretation of the results. As concerns the selection of the method, there may be cases where the conditions of programme implementation prevent the use of one or more specific evaluation procedures. These limitations are mainly related to the unavailability of the required data, especially when estimates of the needed, missing or inadequate data cannot be worked out. For instance, a method using data on acceptors would have to be discarded, if data on acceptors were not recorded or were very incomplete. Other limitations may, however, hamper the sound application of a specific method. Certain types of experimental design studies may not be feasible if a satisfactory control group cannot be established. A real analysis may be prevented if the required number of observation units cannot be obtained or if regional fertility differentials would not yield meaningful results.

Once a method is selected, the general background conditions and trends in these conditions have to be carefully borne in mind so as to guarantee a sound application of the given evaluation procedure. The non-programme factors are the chief indicators of these general conditions and selection of the most relevant factors is fundamental for assessing influences on fertility. With varying background conditions, fertility determinants may have varying roles so that specific variables may be more important in one country than in another. Religion, for instance, may be a chief factor in Catholic countries, but a secondary factor in other cultures. In addition to their role in selecting the relevant fertility determinants, general conditions also influence the adoption of specific assumptions required by the evaluation method. Past and/or expected changes in the level of modernization, in natality and mortality, or nuptiality, in economic development etc. are incorporated in several methods as assumptions for the calculation of potential fertility estimates.

The general background conditions also constitute a frame of reference for the interpretation of the results. These conditions are not comprehensively taken into account by the available evaluation methods which generally focus unevenly on various background variables. Such methodological short-comings can be taken into consideration in interpreting results and drawing conclusions. Indeed, at the stage of interpretation and conclusions, the validity and reliability of the findings can be assessed in the light of the specific country circumstances. In other words, one can discuss the application of the method and try to establish whether a given impact on fertility is likely to have occurred in the given circumstances, whether no other explanation for the resulting impact can be assumed. whether underestimates or overestimates should be hypothesized. For instance, widespread dissemination of birth control practices prior to the family planning programme may mean large-scale substitution effects; rapid social and economic change during the period under evaluation may mean a substantive contribution of non-programme factors to the observed fertility decline; the existence of a well-established mass-communication network associated with well-organized information campaigns may mean that non-programme contraception has been stimulated by the programme and that the programme impact is eventually underestimated.

The Measurement Methods

This section attempts to review the methods utilized by researchers to measure the impact of family planning programmes on fertility. These methods have been classified as follows:

(*a*) Standardization approach;
(*b*) Trend analysis;
(*c*) Experimental designs (control and experimental groups);
(*d*) Couple-years of protection;
(*e*) Component projection; approach;
(*f*) Analysis of the reproductive process;
(*g*) Regression analysis (including path analysis);
(*h*) Simulation models.

It is not the purpose of this review of measurement methods to provide a manual on means of applying the various techniques for assessing programme impact on fertility. Detailed descriptions of the different procedures reviewed in this paper can be found in references cited in appropriate bibliographies. Since the ultimate objective of the meeting for which this text is intended to serve as documentation is to ascertain which method or methods should be applied in specific circumstances for the best results, the objective here is only to provide a simple and concise description of each type of method. In order to facilitate discussion and appraisal of the methods as applied in the country case studies, the methods are described under the following headings:

(*a*) Type of method;
(*b*) Programme and non-programme factors affecting fertility change;
(*c*) Measurement of fertility change;
(*d*) Main assumptions;
(*e*) Population covered;
(*f*) Time reference.

Suggested readings for each method are given in the selected bibliography at the end of this paper.

Standardization approach

Type of method

Approach

The standardization approach is a method of assessing family planning programme impact by decomposing an observed fertility change into component parts, or constructing standardized fertility measures that control factors considered extraneous to the movement of marital fertility. It is a logical first step in assessing impact since it seeks to establish whether a decline in fertility that could be related to the programme did indeed occur. This approach requires several steps. First, a measure of fertility of a specified population is observed or estimated at two points in time. This estimate is generally made on the basis of a crude birth rate[4] and the composite nature of this fertility indicator appears as the first reason for standardization. The next step consists of decomposing the observed change into marital fertility and other factors, or standardizing for the factors considered extraneous. If the crude birth rate or general fertility rate is used as the fertility indicator, these extraneous fac-

[4] The crude birth rate is, indeed, often used even when other fertility indicators are available. The reason is that programme impact is often perceived in terms of population growth rate and the crude birth rate is the fertility component of that growth rate. The need to standardize the crude birth rate arises, of course, from the fact that this rate is a composite indicator of fertility.

tors will more often be age structure and marital status distribution. The number and kind of factors taken into account are limited only by data availability and the ability to express the observed fertility measure as an algebraic function of those factors. Educational distribution may be important where it is changing rapidly and age/education—specific fertility rates differ sharply.[5] The result of the decomposition or standardization then is a measure of the proportion of change in the fertility measure due to "extraneous" factors (such as age structure and marital distribution) and a proportion due to marital fertility which presumably reflects changing birth control practices.

The next step is to gather information in order to determine what part of this increase in family planning practice, and hence of the fertility decline, can be credited to family planning activities.[6] An essential feature of the standardization approach is thus to reduce the observed change to a residue to be investigated for family planning programme effects. The method or methods used to determine whether increased family planning practices are due to the programme vary with specific circumstances. If, for instance, it can be established that there was no pre-programme use of birth control methods, the entire residue may eventually be attributed to the programme. If birth control was already practised, an attempt can be made to link decrease in fertility to increase in programme acceptors. The estimated decline in fertility due to the programme can also, eventually, be compared with the fertility decline due to the programme as assessed by another evaluation method.[7]

Procedure

Various standardization techniques can be used with this approach. A common procedure is the "direct method" of standardization although in some studies the "indirect method" also is used. Authors of evaluation studies sometimes do and sometimes do not refer to the type of technique they have used. Data availability and the evaluator's judgement are often the main criteria for selecting a particular procedure. Judgement is also required when decomposing a change into components since the change can be accounted for by alternate decompositions.

Factors affecting fertility change

Non-programme factors

Factors utilized. Non-programme factors taken into consideration by the standardization approach account for the changes in the observed fertility measure that are not due to programme activities. The kinds of factors selected will depend upon the type of indicators used to assess fertility change, upon specific background circumstances, upon data availability etc. Although the standardization approach, as applied in evaluation studies, appears to focus more on the demographic, non-programme factors, aspects of socio-economic factors and biological factors also are examined below:

(a) *Socio-economic factors.* Modernization and socio-economic development are phenomena generally considered to induce smaller family values and, hence, increased practice of family planning. Indicators of social change appear, however, to be neglected in the standardization approach even when the circumstances of the country under study warrant that attention be given to such factors;

(b) *Demographic factors.* Whenever applicable standardization of some chief demographic variables is always systematically performed. Fertility changes undertaken on the basis of the crude birth rate are analysed with standardized age structure and marital distribution. The total fertility rate and age-specific fertility can be more easily defined as a function of marital status and marital fertility, and the latter measure can be examined further if legitimate and illegitimate fertility data are available. Standardization for other factors, such as mean age at childbearing or duration of marriage, does not appear to have been attempted.

(c) *Biological factors.* The standardization approach does not easily lend itself to the treatment of this category of variable, due more to lack of data than to any theoretical restriction.

Data requirements and sources. A general prerequisite of the standardization of fertility rates for various factors is that appropriate data be available. Age structure and marital distribution information are relatively accessible in standard demographic sources, but standardization for socio-economic categories requires differential fertility data for each of the categories, and this type of information is often difficult to obtain.

Programme factors

Factors utilized. There is no specific set of programme factors needed when the standardization approach is used. Depending upon circumstances, programme factors might not be needed at all. This would be the case if the residual portion of an observed fertility decline could be entirely credited to the family planning programme on the evidence that all other factors remained constant during the period under evaluation. If, however, the residual decline need be further analysed, the programme factors required will be whatever factors are needed by the evaluator to undertake this additional analysis.

Data requirements and sources. As stated above,

[5] John E. Anderson, "The relationship between change in educational attainment and fertility rates", *Studies in Family Planning*, vol. 6, No. 3 (March 1975), pp. 72–81.

[6] See, for instance, Jack Reynolds, "Costa Rica: measuring the demographic impact of family planning programs", *Studies in Family Planning*, vol. 4, No. 11 (November 1973), pp. 310–316.

[7] Sui-ying Wat and R. W. Hodge, "Social and economic factors in Hong Kong's fertility decline", *Population Studies*, vol. XXVI, No. 3 (November 1972), pp. 455–464.

the data requirements will reflect the evaluator's needs. The type of data will, of course, define the type of sources.

The measurement of fertility change

Levels and changes in observed fertility

Fertility is measured at two points in time in order to assess the direction and magnitude of the change, if any. The fertility indicators used to make this assessment are standard: crude birth rate, general fertility rates, age-specific fertility and total fertility rates etc. In fact, the crude birth rate is the basic indicator used in this approach because of its analytical function as a component of population growth. Analysis of the change, however, benefits from the use of age-specific rates since changes in fertility are often not uniform for all ages or age groups.[8]

Levels and changes in fertility in the absence of the programme

The standardization approach does not compute explicitly estimates of fertility in the absence of the programme. However, comparing fertility levels at two points in time can be interpreted as a comparison between potential and actual fertility. From this viewpoint, the observed rate at the end of the period under evaluation stands for actual fertility, and the observed standardized rate stands for potential fertility. This interpretation implies, of course, appropriate assumptions regarding standardization factors. If, for instance, it were assumed that, in the absence of a family planning programme, all factors associated with the initial period would have remained constant during the period under study, the fertility level at the beginning of the period would be an estimate of potential fertility at the end of the period. Indicators of observed fertility can thus be used to obtain estimates of fertility in the absence of the programme.

Main assumptions

Except for the assumptions required for data adjustments and estimates, such as linearity assumptions for interpolation procedures, the standardization approach does not call explicitly for any prerequisites or assumptions. However, in the course of applying this technique for programme evaluation, various assumptions are usually made, explicitly or implicitly, in connexion with various procedural steps. For instance, when a factor is standardized for, it is implicitly assumed that it affects fertility independently of the programme. Standardization for proportion marrying, for instance, thus implies that nuptiality is not affected by programme factors. It is also assumed that family planning programmes affect fertility only through changes in birth control practices.[9] When standardizing for age structure and marital status, for example, it is usually assumed that the effects of these factors are independent and additive. When a factor generally known to be associated with fertility is not standardized for, it is implicitly assumed that the factor has no effect on the observed change. This would be the case, for instance, for the duration of marriage. Once the appropriate standardizations are made, the analysis used to determine what part of the residual decline can be credited to the programme may also be associated with explicit or implicit assumptions.

Population covered

The standardization approach does not define what population coverage is most appropriate; and, in principle, the method can be applied to any kind of population. Practically, its usefulness is fullest when crude birth rates are used; and hence, when the population of a whole country is evaluated. This does not preclude the use of the method for other population subgroups if the requisite data are available.

Time reference

The standardization approach provides measurements of fertility levels and trends, as well as of fertility changes, in terms of calendar years. All rates are period rates and all changes are period changes.

Trend analysis

Type of method

Approach

The various procedures referred to in this discussion of trend analysis attempt to measure the impact of a family planning programme on fertility by comparing indicators of actual (observed) fertility over a specific period of time to projected fertility data for the same time period. The projected fertility data are assumed to represent the potential fertility of the population studied, i.e., the fertility of that population had the programme not been undertaken. Actual, or observed, fertility is the fertility really experienced by the same population over the same time span. The difference between actual and projected fertility is thus assumed to yield an estimate of the impact of the family planning programme on fertility.

---

[8] Ronald Freedman and Arjun L. Adlakha, "Recent fertility declines in Hong Kong: the role of the changing age structure", *Population Studies*, vol. XXII, No. 2 (July 1968), pp. 181–198; and Ronald Freedman and others, "Hong Kong's fertility decline, 1961–68", *Population Index*, vol. XXXI, No. 1 (January-March 1970), pp. 3–18.

[9] The rationale for this assumption appears to be that family planning programmes encourage or promote the use of birth control methods and that the results of a successful programme would be a decline in fertility through decreasing family size rather than through, for instance, changes in age at marriage. Nuptiality changes resulting from a given population policy need to be taken into consideration if it is known that measures to change age at marriage are also taken.

This approach, although defined as a "projection" approach, differs from the standard meaning in that fertility is projected from the past up to the present rather than from the present into the future. An essential feature of this difference is that characteristics of the female population, such as number, age and marital status, are known and need not be estimated for the period under evaluation. Only the hypothetical fertility of this population has to be assessed. Another important feature is that the fertility assumptions need not be based on past fertility trends alone, but may also take account of social and economic changes during the period of evaluation. As with the standardization method, the impact of the family planning programme is identified as a residual. It is usually appropriate to analyse further this residual as is the case with the standardization approach.

Procedures

When studying the fertility of the entire population of a country or of a specific region, trend analysis often involves an extrapolation of past fertility trends from the date of initiation of the family planning programme. This set of projected fertility indicators is then compared with the actual fertility trend over the same time period to gauge the effect of the programme. Various procedures have been suggested for obtaining the hypothetical trend in the absence of the programme. One suggestion is to utilize the fertility assumptions incorporated into projections of the population made before the start of the programme and at a time when relatively little weight was given to such endeavours. Another is to carry forward the fertility trend prior to the programme by one or another method, such as projecting forward the annual rate of change in the fertility indicator; or fitting a trend line by inspection or by some technique such as least squares. Instead of using crude birth rates, resort could, of course, be made to the total fertility rate, age-specific or marital age-specific rates projected on the basis of pre-programme trends. Comparisons utilizing these indicators have the virtue of eliminating the effect of changing age structure and, in the case of marital rates, the effect of changing marital distributions.[10]

Mauldin states that this method is also used with subgroups of the population, particularly acceptors. Here, the observed post-acceptance fertility of acceptors is compared with that expected on the basis of their fertility levels prior to acceptance. In projecting the potential fertility of the acceptors, account is usually taken of their aging and, hence, their reduced fecundity, their reproductive status at the time of acceptance (the fact that typically acceptors are not pregnant and may be in a state of post-partum amenorrhoea), and sometimes the possibility that in the absence of the programme they would have resorted to other forms of contraception or abortion. Such projections of expected births or fertility levels are akin to simulation models based on parameters of the reproductive process; but, as the starting-point they take into account the known characteristics of acceptors.

Factors affecting fertility change

Non-programme factors

Factors utilized. The difference between a projected and an observed indicator of fertility does not necessarily constitute a straightforward estimate of the family planning programme impact. For a given situation, it should be ascertained that no other factors can account for the observed, or part of the observed, difference, if any. The three main categories of factors affecting fertility are examined below:

(1) *Socio-economic factors*. Whether this category of factors is taken into account depends upon the type of projection technique used. Fertility projections do not explicitly take into account the relevant socio-economic factors. Rather, when assumptions as to the hypothetical fertility trend are formulated, such assumptions take implicit account of these socio-economic factors. This type of projection permits explicit examination of ethnic differences or urban-rural differences, for example, through separate projections. New types of projection models are, however, being developed in which it should be possible to include certain socio-economic factors;

(2) *Demographic factors*. Simple extrapolation of past fertility trends usually appears to be unsatisfactory for programme evaluation; and such factors as age structure, nuptiality and non-programme family planning practice can be included in the projection procedure;

(3) *Biological factors*. Most techniques of extrapolating fertility trends are not designed to take biological factors into account. But when reproduction models are used to obtain the number of births expected to acceptors, biological factors are taken into consideration. Probabilities of conception or of foetal loss, length of sterile period etc. are some of the factors included in this category.

Data requirements and sources. The main categories of data needed concern those required to determine the fertility trend of the population prior to the programme, to ascertain the current size, age and sex distribution, and other characteristics of the population under study, and to assess the possible social, economic and cultural changes that occurred during the evaluation period. Sources do not raise difficulties specific to the evaluation purpose, at least in so far as standard indicators are concerned. In certain cases, however, only special record-keeping procedures can provide the appropriate data for a given evaluation. Age-specific marital fertility of acceptors or biological

[10] W. Parker Mauldin, "Births averted by family planning programmes", *Studies in Family Planning*, vol. 1, No. 33 (August 1968), pp. 2–3.

characteristics, for instance, require special surveys or the maintenance of special clinic record.

Programme factors

Factors utilized. The programme factors are not explicitly identified in the projection approach. In assessing programme impact by estimating the difference between projected and actual fertility, the aggregate effect of the programme is assessed as a residual rather than inferred from specific programme activities indicators. Thus, no programme factors are specifically taken into consideration. However, should there be a need to analyse further the difference between observed and projected fertility, programme indicators similar to those which might be needed in similar circumstances with the standardization method could be required for the projection method.

Data requirements and sources. The required programme indicators and their sources will reflect the type of analysis chosen by the evaluator.

Measurement of fertility change

Levels and changes in observed fertility

The estimation of levels and trends in observed fertility has two objectives: (*a*) determining past trends in order to make fertility assumptions; and (*b*) determining the current fertility trend in order to have a term of comparison for the projected trend. Current observed fertility is often approached with standard fertility indicators, such as the crude birth rate and age-specific or age-specific marital fertility rates. For estimating fertility prior to the period under evaluation, additional indicators can be used to identify the timing and magnitude of a given change, if any. Such factors may include average family size, open birth intervals, parity rates etc.

Levels and changes in fertility in the absence of the programme

All standard fertility indicators can be used as indicators of projected fertility provided they are the same as the observed indications so that comparison will be feasible. A common fertility measure used is the crude birth rate, even when other data are available, because as a main component of population growth this rate appears to provide a better perception of population change. As this indicator does not, however, provide a reliable estimate of changes in fertility due to contraceptive use, more detailed approaches to projected fertility are performed. The comparison of age-specific fertility rates for five-year age groups, for instance, easily improves the analysis of potential changes. Reproduction models have also made use of pregnancy rates as a more sensitive indicator of short-term programme impact.

Main assumptions

The assumptions regarding trend analysis are associated with the estimation of fertility in the absence of the programme, specifically with the amount, the direction, the pattern and the date of onset of the assumed fertility changes. These explicit assumptions, along with the non-programme factors taken into consideration, tend to include in this approach a varying number of implicit assumptions. If, for instance, fertility is assumed to remain constant, it is implicitly assumed that non-programme background conditions have not changed. If the nuptiality factor is not taken into account explicitly, it is implicitly assumed that the population has remained homogeneous with respect to marital status. A linear extrapolation of a pre-programme fertility trend implicitly assumes that fertility is linearly associated with its determinants. Thus, in numerous instances, explicit assumptions are associated with implicit assumptions; likewise, failure to take some factors into consideration also may reflect implied assumptions. In the latter case, one may say that all non-programme factors that are not explicitly taken into account are implicitly accounted for in the very formulation of the fertility assumptions: if a fertility decline is assumed, it implicitly reflects the aggregate effects of all non-programme fertility determinants.

Population covered

Two main categories of population can be evaluated by trend analysis, the fertility projection approach. The first category includes the population living in a certain area: this means the entire country, administrative or geographical units, or areas where a family planning programme is being implemented. The second category includes populations that can be identified by a certain characteristic: the acceptors of programme services is the most common group in this category. Evaluation objectives and data availability are the major criteria for selecting a given coverage.

Time reference

Trend analysis yields results in terms of calendar years. It thus allows the evaluation of the effects of the programme beginning at a given date, say, from programme initiation or any time during programme implementation.

Experimental designs: control and experimental groups

Type of method

Approach

The experimental-design approach attempts to employ features of the classical experimental design to estimate the effects of a family planning programme on fertility. The classical design requires that two groups of population, as similar as possible before a treatment is given, be established. Often a "before" measure is taken in each group to determine the point from which changes are expected to occur and to provide a check on the equivalence of the two groups regarding the variable which is supposed to be affected. One of the two groups, the experimental group, is then exposed to a given treatment while the other group, the control group, is excluded from that treatment. At the end of the experiment, an "after" measure of the treated variable is taken for both the experimental and the control group and is compared. The effect of the treatment is taken as the difference between the two groups in the amount of change that occurred. Translated in terms of programme evaluation, the treatment consists of family planning programme activities and the "before" and "after" measures are one or more selected indicators of fertility. If fertility is found to change more rapidly in the experimental than in the control group, the difference is assumed to be due to the family planning programme. Here again, the effect of the programme is obtained as a residual.

In the classical design, the experimental and the control groups are constituted prior to the treatment and the persons included in each group are assigned randomly. When family planning evaluation is undertaken, the persons included in the experimental group are not assigned at random, but are self-selected on the basis of their decision to participate in the programme or are arbitrarily selected on the basis of an area chosen for programme implementation. As a result, comparability is not ensured between the experimental and the control group. Similarity and comparability of the two groups can, however, be greatly improved upon by a procedure of selective matching.[11]

Procedures

Designs without matching can utilize the general population of an area or programme acceptors as the experimental group. When a family planning programme is undertaken in only part of a country, that population constitutes the experimental group; and another part or the entire population outside the programme area forms the control group.[12] If programmes with different intensities are undertaken in different areas, the population of the area with the more intensive programme becomes the experimental group while the population of the area with the less intensive programme (or without programme) becomes the control group.[13] In one case, the purpose of the comparison is to evaluate the effects of the additional inputs provided to the more intensive programme area; in the other case, the over-all effect of the programme is assessed. When a family planning programme is undertaken on a nation-wide basis, so that no control area is available, the two comparative groups are formed on the basis of programme acceptance. Acceptors constitute the experimental group, and the control group is drawn either from the general population or from the non-acceptors only.

If, for instance, it is decided to use programme acceptors as the basis for the experimental group, comparability with this self-selected group can, as stated earlier, be improved by a matching procedure. Basically, this procedure consists of establishing a list of acceptors according to a number of demographic and/or socio-economic characteristics. This is the experimental group. The control group is then constituted by selecting from another source, household registration for instance, a number of non-acceptors with equivalent characteristics. This matching procedure can be carried out either on the basis of individuals or of the whole group. In the first case, each single characteristic of an acceptor has to be found in a corresponding non-acceptor. In the second case, individuals are ignored and acceptors are matched as a group where only the total number of individuals and the total number of each characteristic are considered relevant. Variations in socio-cultural and programme conditions lead, of course, to variations in the criteria for defining acceptance and non-acceptance as well as in matching criteria. When a programme is established in only one area of a country, another area with similar aggregate demographic and socio-economic characteristics is sometimes matched with it to serve as the control group.

Factors affecting fertility change

Non-programme factors

Factors utilized. Non-programme factors that may have affected fertility are taken into consideration implicitly and, when matching is performed, explicitly. Without matching, it is assumed that socio-economic and other factors which may affect fertility during the period under evaluation operate equally on the experimental and the control group, so that it is not necessary to identify those factors explicitly. However, since randomization has not been performed, a difference in fertility change might result from the fact that the experimental group has non-programme characteristics that are different from those of the control group. This possible differential effect of non-programme factors is mitigated when the two groups are matched for non-programme characteristics, thereby avoiding a confusion of programme and non-programme effects. Non-programme factors for matching therefore, must be explicit, and they usually can be subsumed under the three following categories:

(1) *Socio-economic factors*. This category contains

[11] L. M. Okada, "Use of matched pairs in evaluation of a birth control program", *Public Health Reports*, vol. 84, No. 5 (May 1969), pp. 445–450.

[12] Population Council, "India: the Singur study", *Studies in Family Planning*, vol. 1, No. 1 (July 1963), pp. 1–4.

[13] Jae Mo Yang, "Fertility and family planning in rural Korea", *Proceedings of the World Population Conference, Belgrade, 30 August–10 September 1965*, vol. II, Selected Papers and Summaries: *Fertility, Family Planning, Mortality* (United Nations publication, Sales No. 66.XIII.6), pp. 309–312.

a wide range of indicators, such as income, education, occupation, religion, ethnic group and type of residence. One or several of the most relevant indicators are chosen on the basis both of data availability and of their theoretical relationship with fertility in the group under study.

(2) *Demographic factors.* This category includes standard demographic indicators, such as sex, age, marital status, number of births and family size. When appropriate, other characteristics may be added, such as open birth interval on number of living sons, as well as the relevant Knowledge-Attitude-Practice (KAP) indicators.

(3) *Biological factors.* Although these factors might be of importance for certain matching studies, they are seldom used. This category would include length of anovulatory period, sterility etc.

Data requirements and sources. Non-programme data are required when matching is undertaken; they consist of indicators relevant to the selected matching factors. Since matching studies usually pertain to acceptors versus non-acceptors, KAP studies, follow-up surveys, clinic records and service statistics in general constitute the most appropriate source for the acceptor group; for the non-acceptors, data are obtained from standard demographic sources, such as fertility surveys, censuses and household registration. Certain types of non-programme data are also relevant when matching is not undertaken. This concerns age structure and nuptiality data, for instance, when crude birth rates are used as fertility indicators.

Programme factors

Factors utilized. As in the two methods previously described, the effect of the programme is calculated as a residual, as there are no provisions for taking programme factors explicitly into account. Programme factors would only be needed if any type of additional analysis were undertaken regarding the population under study. Such analysis would go beyond the experimental design approach.

Data requirements and sources. The type of data needed would depend upon the type of analysis undertaken. Service statistics would certainly be the most appropriate source.

Measurement of fertility change

Levels and change in observed fertility

Measures of both level and change are used in the experimental-design approach. As stated above, the "before" (or bench-mark) fertility measures of each group is needed in order to equate the two groups as concerns their fertility. The measure of fertility change is needed for the actual assessment of the impact of the programme. The "before" and "after" fertility measures in each group are obtained from observed data, and the change measure is obtained as the difference between the "before" and "after" measures. The data used to make these estimates usually consist of standard fertility indicators, such as the crude birth rate and age-specific fertility rates.

Levels and changes in fertility in the absence of the programme. The estimates of fertility trends in the absence of the programme are computed from the "before" and "after" fertility measures of the population constituting the control group. This group is, indeed, assumed to represent the fertility that would have been experienced by the population of the experimental group had the family planning programme not been undertaken. The fertility indicators selected to make this estimate are, of course, the same as those used to assess fertility changes among the population of the experimental group.

Main assumptions

A major condition for obtaining valid results with the experimental-design approach is that the experimental and the control groups be comparable. In the absence of randomization, a number of assumptions are implicitly made regarding their comparability. When no matching procedure is undertaken, it is usually assumed that the two comparative groups are, nevertheless, of similar characteristics and that the population of the control group is not influenced by the programme activities to which the experimental group is subjected. These assumptions are supposed to apply to populations of particular areas, or to non-matched acceptors and non-acceptors. When the experimental and the control groups have undergone a matching procedure, this procedure is assumed to have ensured comparability, but it is still implied that the programme activities do not affect the control group. More generally, it is also implied that the procedure used to record the "before" and "after" measures does not have a differential effect.

A crucial assumption related to the specific objective of evaluating a family planning programme is that the fertility of the population in the control group is assumed to stand for the fertility the population in the experimental group would have experienced had the family planning programme not been undertaken. It is thus further implicitly assumed that the social, economic, political and cultural events occurring during the period of time elapsing from the "before" to the "after" measure affect likewise the fertility of the population in each group; that the population in each group remains homogenous during the period of evaluation; and that the period for identifying the "before" and "after" measures refers to the same time span.

Population covered

Like the other methods discussed above, the experimental-design approach does not inherently specify the most appropriate population coverage. The method is flexible enough to cover the population of an area or of the whole country provided assumptions are met, data are available etc.

Time reference

The experimental-design approach, like the preceding approaches, measures fertility changes and programme effects on fertility in terms of calendar years and period rates.

Couple-years of protection

Type of method

Approach

Unlike the preceding methods, which assess the impact of the programme by comparing fertility experienced under programme conditions to estimates of fertility in the absence of the programme, the couple-years of protection method determines the impact of the programme directly from data on birth control methods and programme acceptors. The impact is thus assessed directly from information on programme activities rather than indirectly as a residual. The estimating approach consists of two stages. The first stage is directed towards determining the number of couples protected against the risk of pregnancy during one year, and it yields an index of couple-years of protection. This index is calculated by estimating the length of time a couple is likely to be protected by each "application" of a family planning method and then using this factor in conjunction with the number of units of each method distributed to obtain CYP by method and over all. The second stage consists of inferring, for a given amount of CYP prevalent in a year, the number of births which have been averted.

Procedure

Ideally, factors for estimating CYP from each method should be worked out separately (for each country) and changed with time (within countries) inasmuch as the conditions that determine the length of protection afforded by a unit of contraception are not constant. The couple-years of protection estimate can be computed from Wishik's prevalence index.[14] The prevalence of CYP in year T is the sum of protection conferred to cohorts of acceptors from previous years who are still wearing the device in year T, including insertions of intra-uterine devices (IUD) performed in year T:

CYP prevalence index in year T

$$= \frac{a}{r} \sum_{i=0}^{j} N_{T-i} \{e^{-ri} - e^{-r(i+1)}\} \qquad \text{where } j \leq 6$$

where r = rate of IUD attrition during one year:

N_{T-i} = number of insertions in year $(T-i)$;

j = the maximum value of i, which indicates the number of years the IUD program has been in effect.

An example of another formula which has been used to estimate CYP prevalence is:

$$CYP_n = 0.01\,C_n + (0.75\,I_n + 0.50\,I_{n-1} + 0.35\,I_{n-2}) + (S_n + 0.95\,S_{n-1} + 0.90\,S_{n-2})$$

where C = the number of conventional contraceptives;

I = the number of intra-uterine devices inserted;

S = sterilizations (vasectomies and tubectomies).

The coefficient 0.01 reflects the assumption that 100 units of conventional contraceptives must be distributed to provide one couple with protection for one year. The coefficients for I reflect the continuation rates assumed to be in force and serve to bring forward to the current year those women still protected from insertions in earlier years.[15] Where possible, these rates should be based on actual experience and changed with time. The effect of sterilizations performed in earlier years is assumed to decrease over time to reflect the diminished protection associated with the probability of marital dissolution, the onset of menopause etc. The formula can, of course, be modified to express separately the number of vasectomies and tubectomies, or by adding a pill component if this method happens also to be used. The number of births averted is then obtained by applying to the number of couple-years of protection the ratio of how many couple-years of protection are needed to prevent one birth in one year.

Factors affecting fertility

Non-programme factors

Factors utilized. Non-programme factors affecting fertility are not explicitly taken into consideration by the CYP method. The coefficients that appear in the formula of the CYP index do, however, account for some such factors indirectly. In theory, these coefficients define the extent of protection given by a particular method. Notably, in the case of sterilization, for example, the coefficient would account for secondary sterility or disruption of marital union, or both. With respect to IUDs, the coefficient as used focuses on device retention, but might include more than that particular factor. Defining the non-programme factors

[14] Samuel M. Wishik and K. H. Chen, *The Couple-Year of Protection Index: A Measure of Family Planning Program Output,* Manuals for Evaluation of Family Planning and Population Programs, No. 7 (New York, Columbia University, International Institute for the Study of Human Reproduction, 1973).

[15] Enver Adil, "Measurement of family planning progress in Pakistan", *Demography,* vol. 5, No. 2 (1968), pp. 659–665; and Lee L. Bean and William Seltzer, "Couple years of protection and births prevented: a methodological examination", *Demography,* vol. 5, No. 2 (1968), pp. 947–959.

utilized in the CYP approach consists mainly of describing the various factors that have been applied in the calculation of the coefficients.

Data requirements and sources. The needed data and their sources are determined by the non-programme factors that are taken into account in the calculation of the various coefficients. These requirements may involve life tables for calculating the probability of widowhood and clinic or survey data for the age distribution of secondary sterility.

Programme factors

Factors utilized. The impact of the programme is determined directly from indicators of programme services, namely, the number of contraceptives distributed, sterilizations performed and, eventually, abortions undertaken. These categories of programme services are thus the only summary indicators of programme factors utilized in the CYP method.

Data requirements and sources. The data needed are statistics of the number of conventional and oral contraceptives distributed, intra-uterine devices inserted, and sterilizations and abortions performed, as the case may be. The definition of "distributed" often depends upon the type of data-gathering procedure; as emphasized in the literature on the subject, distribution may mean "distribution to suppliers" or "acceptance by couples" or "acceptance and actual use" etc.[16] The whole range of service statistics, as well as private sources and follow-up surveys, can be considered useful sources.

Measurement of fertility change

The CYP method does not require the measurement of fertility changes during the period under evaluation. The impact of the programme is estimated directly from programme activities.

Main assumptions

A number of assumptions are explicit in the coefficients found in the CYP formula shown. The coefficient of C, for instance, means that it is assumed that 100 units of conventional contraceptives per couple prevent one birth per annum, provided that sexual intercourse takes place twice a week, on the average, during that year. The coefficients both for the intra-uterine devices and for sterilizations are the assumed protection provided by those methods during the initial and successive 12-month periods of use. A final explicit assumption is that regarding the number of couple-years of protection needed to prevent one birth. In Pakistan, it is estimated that three or four couple-years of protection are needed to prevent one birth. This figure will, of course, vary with the type of population being evaluated.

[16] L. L. Bean and W. Seltzer, *loc. cit.*

In addition to these explicit assumptions, however, there are a number of important implicit assumptions whose validity will determine the reliability of the method. Most notably, it is assumed that all family planners included in the computation are programme acceptors; and that all conventional methods distributed or accepted are used, that they are used during periods of pregnancy risk and that they are used efficiently. It is also assumed that the retention rates for IUDs are constant and evenly distributed between and within age groups; that couples do not substitute programme methods for non-programme methods already being used, that couples do not switch from one method to another during the evaluation period, that couples do not supplement one method by another; that couples are fecund at the time of acceptance, that they remain at risk of conceiving during the evaluation period and that their fecundability remains constant.

Population covered

The population covered by the CYP method includes, in principle, all and only programme acceptors, i.e., couples who receive and actually use family planning methods provided by the programme or under programme auspices. A definition of the concept of "acceptor" is, however, required for each particular utilization of the method if a meaningful interpretation of the results is expected.

Time reference

The couple-years of protection approach is a calendar-year approach. Results are provided in terms of 12-month periods.

Component projection approach

Type of method

Approach

Like the CYP technique, the component projection approach is directed towards calculating births prevented and makes use of data on programme acceptors as a basis for computation. It is thus also a method which infers programme impact directly from programme acceptance. The rationale of the method consists of estimating the acceptors' potential fertility and assuming that the number of births they would have had in the absence of the programme are all births averted by the programme. This method can be used to estimate future changes in age-specific fertility rates or, as in the present interpretation, to provide estimates of births averted from the past up to the current time. In the latter use, the recorded number of acceptors and continuing users are used instead of projected estimates of acceptors. In addition to number of acceptors still in the programme at a given period of time, estimates of their potential fertility are also needed.

Procedure

The number of births averted by the family planning programme in age group i in year t is given by:

$$Q_{i,t} \cdot g_i$$

where Q_i is the number of acceptors belonging to age group i who were practising totally effective contraception in year $t-1$; and g_i is the potential fertility of the acceptors in age group i.[17] Various approaches to estimating Q_i and g_i have been developed, some being more elaborate than others. Venkatacharya[18] introduces, for instance, a factor for the family planning method use-effectiveness. The Q_i can be calculated from continuation rates applied to the number of couples who entered the programme in year $t-x$, $x=1, 2, \ldots n$, or can be obtained directly from a follow-up study.

The procedure to estimate g_i also varies. In applying their method to the Republic of Korea, Lee and Isbister estimated potential age-specific marital fertility rates of acceptors by increasing the marital fertility rates of the general population by 20 per cent on the assumption that acceptors have higher fertility than the general population. In studying the probable impact of IUDs and sterilizations in India, Venkatacharya estimated the potential number of births on the basis of a matrix of annual probabilities of live births taking into account, among other factors, the acceptors' initial susceptibility to conception. Any appropriately designed model could, of course, be used to estimate the probable number of births in the absence of a programme. Estimates of g_i can also be based on the acceptors' own pre-programme fertility rates, or a range of values can be used to study the effect of alternate estimates on the number of births averted.

Factors affecting fertility change

Non-programme factors

Factors utilized. Which particular non-programme factors are utilized in the component projection approach depends upon the procedures used to estimate Q_i and g_i. Where non-programme factors are not explicitly taken into consideration, they are, in fact, often implicitly accounted for. For instance, if Q_i is calculated on the basis of the original number of acceptors, mortality can be explicitly included in computing the number of surviving continuing users; if Q_i is based on a follow-up study, the mortality factor is implicitly taken into account. Such factors include:

(a) *Socio-economic factors.* This category of factors is not explicitly taken into consideration. However, in establishing the levels of potential fertility of the acceptors, assumptions can be made which permit imputations as to expected social and economic influences, especially in countries that are undergoing rapid social change;

(b) *Demographic factors.* The effect or lack of effect of some factors, such as divorce, mortality or widowhood, can be determined explicitly or implicitly. This is also true for interruption of contraceptive use or for the use of non-programme family planning methods;

(c) *Biological factors.* Consideration for this category of factors depends mainly upon the procedure used to estimate the acceptors' potential fertility. If potential fertility is calculated in terms of age-specific rates, biological factors are not explicitly included; but the use of reproduction models would include biological parameters, such as post-partum amenorrhoea and fecundity.

Data requirements and sources. The requisite non-programme data consist mainly of those acceptor characteristics which affect their number and whose effects on fertility should not be credited to the family planning programme. For example, if an IUD acceptor gets divorced, or becomes a widow, she is no longer exposed to the risk of pregnancy. Data of this type can be obtained from service statistics, follow-up and evaluation surveys or, eventually, from life tables as regards mortality and widowhood. The data requirements for estimating potential fertility are determined by the procedure utilized. They may involve fertility rates or marital fertility rates, by age prior to or during the period of fertility evaluation of the acceptors' own pre-programme fertility. If a reproduction model is used, the data required will be, as indicated above, of a more biological nature.

Programme factors

Factors utilized. The impact of the family planning programme depends upon the number of programme acceptors[19] in each age group, by the method of fertility control applied. In so far as acceptors of contraceptives are concerned, the rate of continuation is a factor of vital importance.

Data requirements and sources. The required data are the number of acceptors by age, type of conventional contraceptives, oral contraceptives, intra-uterine device insertions, and/or the number of abortions and sterilizations performed, as indicated. The annual number of current family planners can be obtained from appropriate continuation rates applied to the number of initial acceptors. Depending upon the data-gathering procedures, the statistics may be obtained from follow-up recordings or follow-up surveys; or from private clinics or physicians, if they are included in the programme plan etc.

[17] B. M. Lee and J. Isbister, *loc. cit.*

[18] K. Venkatacharya, "A model to estimate births averted due to IUCDs and sterilizations", *Demography,* vol. 8, No. 4 (November 1971), pp. 491–505.

[19] The term "programme acceptors" refers to acceptors registered in the records of an official programme.

Measurement of fertility change

As with the CYP method, the component projection approach yields estimates of family planning programme impact in terms of births averted and thus does not require measuring fertility changes that may have occurred during the evaluation period.

Main assumptions

A number of assumptions are implied in this approach, and their nature and number depend upon the particular procedures used to obtain $Q_{i,\ t}$ and g_i. Lee and Isbister assumed that births averted in year t result from acceptors who were current users in year $t-1$ and that there is no mortality between years $t-1$ and t. In the computation of the current number of acceptors, assumptions are made regarding the rate of "consumption" of renewable contraceptives (e.g., condoms) or the rate of retention of non-renewable contraceptives (e.g., IUDs), as well as regarding the rate of use effectiveness of each family planning method. An assumption can also be included regarding the use of family planning for postponing a birth rather than for limiting the size of the family. Assumptions for divorce, mortality, widowhood etc. may also be necessary. Of course, some of these assumptions may not be necessary, if acceptors can be followed individually on an annual basis.

Assumptions regarding the acceptors' potential fertility can also vary in nature and number, depending upon how this factor is obtained. It can be assumed that the potential fertility is equal to, higher or lower than the acceptors' own pre-programme fertility, or than the fertility of the general population or the married population. The use of reproduction models to estimate potential fertility requires assumptions regarding such biological factors as post-partum amenorrhoea, foetal losses and sterility.

Population covered

The population covered by the component projection approach should consist of programme acceptors only, i.e., couples who have accepted a family planning method through any of the official or private programme services. As mentioned previously, the concept of "acceptor" requires an operational definition.

Time reference

The various evaluation procedures associated with this method yield estimates of births averted in calendar years.

Analysis of the reproductive process

Type of method

Approach

The two methods of analysing the reproductive process, described below, also attempt to estimate the number of births averted by a family planning programme using data on contraceptive acceptance and use. They have in common the utilization of the life-table technique to calculate the proportion of acceptors who, after a specified period, are still using a given family planning method. Specifically, the technique is intended to estimate births averted per segment of IUD[20] use, by comparing the average duration that childbearing is interrupted because of the programme contraceptive used, with the average duration of interruption per birth had the IUD not been adopted. Thus, if a woman uses programme contraception effectively and prolongs her non-pregnancy state for a period equal to this average, it is assumed that a birth has been prevented.

Procedures

Potter[21] summarizes his procedure as follows:

$$B = \frac{I}{D}$$

$$I = F(R - A - PW)$$

where B = births averted per first segment of contraception;

I = average duration that childbearing is interrupted;

D = average duration per birth required in the absence of programme contraception;

F = proportion of couples not sterile at time of acceptance;

R = mean time programme contraception is used among couples not sterile at time of acceptance;

A = allowance for post-partum amenorrhoea;

P = proportion becoming accidentally pregnant;

W = penalty for accidental pregnancy.

In computing R, allowance is made for mortality, secondary sterility, and discontinuation of contraceptive use. Ideally, R is estimated by a life-table analysis in which the components are entered as competing risks. As applied to the intra-uterine device, discontinuation is due either to accidental pregnancy, or to expulsion or removal of the device.

Other factors that can affect the R-value, such as divorce, can also be entered if applicable. In so far as family planning is being adopted for birth postponement rather than limitation, the results should be considered to be applicable for the short run only.

---

[20] Reference is made here only to the IUD because the method was originally devised for this contraceptive.

[21] For a complete description of the method, see R. G. Potter, "A technical appendix on procedures used in the manuscript 'Estimating births averted in a family planning program'," paper prepared for Major Ceremony V, University of Michigan Sesquicentennial Celebration, 1 June 1967. See also R. G. Potter, "Application of life-table techniques to measurement of contraceptive effectiveness"; and *idem*, "Estimating births averted in a family planning program".

Wolfers'[22] procedure, like Potter's, was devised to compare periods of effective contraceptive use due to the programme with mean birth intervals that would have been experienced by the acceptors in the absence of the programme. The number of births averted is obtained in several stages. The first stage consists of calculating (as applied to intra-uterine devices) monthly continuation rates with allowance made for accidental pregnancies, expulsions and removals. The procedure is then extended, at a second stage, to take into consideration such factors as post-partum amenorrhoea, secondary sterility and probability that an accidental pregnancy will result in a live birth. Simultaneously, the expected birth intervals in the absence of the programme are included in the computation, and estimates of births averted per 100,000 acceptors are obtained for each year after acceptance.

Factors affecting fertility changes

Non-programme factors

Factors utilized. Analysis of the reproductive process makes allowance for various factors which may influence the calculation of the number of births averted by the programme, in order not to credit the programme with effects originating outside the programme:

(a) *Socio-economic factors.* These factors are seen as a major determinant of family planning acceptance in the absence of a programme. Hence, it is assumed that couples who became acceptors might, in the absence of the programme, have become contraceptive users, had social change favoured such a development. The procedures of both Potter and Wolfers leave room to take this possibility into account.

(b) *Demographic factors.* A number of standard demographic factors which may affect the measurement of programme impact are taken into consideration. These factors include mortality and widowhood, as well as divorce occurring among acceptors. Another major factor affecting the number of births averted estimate is the extent of family planning practice among programme acceptors which existed before acceptance. Further, the number of acceptors who discontinue the programme method or have an accidental pregnancy are taken into consideration by one or the other procedure.

(c) *Biological factors.* The biological factors taken into account include post-partum amenorrhoea and sterility.

Data requirements and sources. The data required concern acceptors' characteristics: family planning practice before acceptance; practice after acceptance; and, of course, reliable data on continuing users. With Potter's procedure, the estimation of potential fertility requires that the acceptors' own pre-programme fertility be known; Wolfers' procedure requires data on birth intervals for the general population. In addition, data are also needed to estimate the biological parameters included in the analysis. All standard and family planning programme sources are used to obtain these data, including service statistics, KAP studies, follow-up studies and appropriate life tables for mortality estimates.

Programme factors

Factors utilized. The impact of the family planning programme is obtained from the number of programme acceptors.

Data requirements and sources. The needed data consist of the number of acceptors of each category of family planning method provided by the programme. In the case of non-renewable contraceptives, such as the intra-uterine device, data on reinsertions are also needed if accurate estimates of programme impact are expected. Service statistics provide the main source of required data regarding the number of acceptors enrolled in the programme. Statistics from private sources that distribute contraceptives are also needed if they are considered as being part of the official programme effort.

Measurement of fertility change

As with previous methods measuring programme impact directly from contraceptive inputs, no measure of changes in fertility is needed.

Main assumptions

Although the two procedures have a similar structure, some differences both in approach, related in part to the programmes, and in estimating techniques can be mentioned. Potter, for instance, assumes that acceptors have higher fecundity than the average population, whereas Wolfers does not share this view. Potter also assumes homogeneity of risk within five-year age groups and treats age-specific risks as constants unvarying with time elapsed from insertion. He also assumes that a proportion of acceptors are sterile at the time of acceptance, while a proportion of those fertile discontinue "immediately" to account for the high rates of discontinuation during the initial two or three months. While Wolfers also assumes homogeneity within age groups, he does not make the other assumptions and takes into account time elapsed since acceptance. Both procedures make assumptions regarding secondary sterility and post-partum sterility. Potter, in addition, makes explicit assumptions regarding mortality and, if applicable, divorce. Both procedures assume that family planning is undertaken for the purpose of limiting rather than postponing births, and both also allow inclusion of assumptions regarding acceptors who would have used contraception in the absence of the programme.[23]

[22] D. Wolfers, *loc. cit.*

[23] References to these main assumptions are found in R. G. Potter, "Estimating births averted in a family planning program", pp. 418–422 and 430; and D. Wolfers, *loc. cit.*, pp. 115–118 and 140.

Population covered

The population covered includes all acceptors included in the programme record-keeping system, whether the services are obtained from official or private sources.

Time reference

Estimates of births averted are obtained in total for first segment of use or for successive years of contraceptive use since time of acceptance.

Regression analysis

Type of method

Approach

Determination of programme impact by regression analysis is rather different from the methods previously described, in that this method is directed specifically towards inferring the role of all the factors included in the analysis. This method allows the inclusion of a large range of explanatory variables; and when applied to family planning programme evaluation, it permits the handling of both programme and non-programme factors. In dealing with programme impact on fertility, a fertility indicator becomes the main dependent variable; and socio-economic, demographic and possibly biological factors, as well as programme indicators, constitute the independent variables. Programme impact is thus assessed directly through the regression parameters, rather than indirectly as a residue. As discussed here, the unit of analysis is an areal subdivision of a country rather than an individual, and aggregate data for these units are used in the analysis.

A special type of regression analysis, path analysis, has also been used to assess programme impact on fertility. A path analysis is always based on an explicit model where all variables are ordered in time, the direction of the relationship is explicitly stated and the presence of direct and indirect effects upon the dependent variables is specified.[24] This approach, which is based on standard multivariate regression techniques, is specially designed to assess the magnitude of the indirect effects, in addition to the direct effects obtained by a standard regression analysis.

Procedure

Estimation of the regression parameters can be carried out through various procedures which vary generally with the type of regression model. A simple and much-used estimating procedure is the ordinary least-square technique. This technique can be applied to a single equation model or to a set of several equations, provided that the assumptions associated with a particular model satisfy the least-square criteria. If some of the conditions are not met, other techniques, usually more complex, are available. The estimating procedure must therefore be selected in accordance with the type of regression model and the relevant assumptions.

Factors affecting fertility changes

Non-programme factors

Factors utilized. The regression approach does not identify the particular variables that can or should be included in the regression equations. The selection of these variables depends upon the model used to account for changes in fertility behaviour. The types of socio-economic, demographic or biological factors taken into account are thus chosen on the basis of the model developed to "explain" fertility levels or changes in fertility, as discussed below:

(a) *Socio-economic factors.* A number of socio-economic factors that are associated with industrialization and economic development are usually considered to be associated with fertility and can be included in a regression analysis to account for these non-programme effects. Examples of such factors are income, literacy, level of education, industrialized or agricultural composition of the labour force and urbanization indexes;

(b) *Demographic factors.* Which of these factors are included depends in part upon the measure of fertility used as a dependent variable. Thus, factors representing age structure or nuptiality may be introduced when the dependent variable is the crude birth rate or age-specific fertility rates. In addition, a measure of infant mortality is often utilized because of the presumed effect of this factor on subsequent fertility. Occasionally, a measure of the initial fertility level of an area or a lagged value of the dependent variable is included in the model;

(c) *Biological factors.* Aggregate values of biological factors are not usually available and hence cannot be included in a regression analysis. By now, it will have become clear that the biological variables commonly taken into consideration in family planning programmes evaluation are fecundity, sterility, post-partum amenorrhoea; no attempt has been made, so far, to include such variables in a regression model.

Data requirements and sources. The data utilized will depend both upon the factors selected for inclusion in the regression and upon the availability of data. In addition, the choice of specific data is also guided by the particular interrelations between variables that the evaluator has assumed. Thus, it may be subsumed, for example, that the relationship between family planning acceptance and industrialization is best indicated by measures of the proportion of labour force engaged in agriculture or by the proportion of males engaged in agriculture. Similarly, the education factor can be identified by data on literacy, the educational

[24] An effect is defined as "indirect" when it is exerted on the dependent variable through another variable.

level attained by males or females; or by the proportion of children of a given age attending school etc. The demographic data used can range from crude birth rates to age-specific marital fertility rates and from crude death rates to infant mortality or expectation of life. Many kinds of data can be utilized in this method, and a wide range of sources is needed.

Programme factors

Factors utilized. A variety of programme factors can be employed to allow for the programme effects on fertility. Services provided by the programme, abortions, sterilizations, contraceptives distributed; and administrative, medical and paramedical personnel involved in the programme are only a few among such possible factors. Regression analysis is the only method that allows for the inclusion of programme indicators beyond acceptors for assessing programme impact on fertility.

Data requirements and sources. Man-month of use of various personnel components, number of acceptors of various family planning methods etc. constitute significant indicators of the programme activities. Service statistics and follow-up surveys represent the best sources of data for construction of these indicators.

Measurement of fertility change

Levels and changes in observed fertility

The main dependent variable that appears in regression analysis is, of course, fertility. The determination of fertility levels for the various units of observation can be measured according to various indicators: crude birth rates and standardized crude birth rates; general fertility rates; age-specific and age-specific marital fertility rates; etc. Data on observed fertility are gathered for the year of observation, but depending upon the model used, fertility of previous years can also be included in the analysis as lagged dependent variables.

Levels and changes in fertility in the absence of the programme

The regression approach does not require estimates of potential fertility, nor is there any estimate of potential fertility associated with this method. The use of observed fertility measures and potential fertility estimates implies a comparison between these two indicators; regression analysis does not rely on such a comparison.

Main assumptions

The ordinary least-square technique for estimating regression parameters is a straightforward and relatively simple procedure; hence, it is often preferred to more complex approaches. But this procedure requires that a number of conditions associated with the regression model be met, thus restricting somewhat its use. However, only assumptions related to ordinary least squares are examined here, without any attempt to be exhaustive. These assumptions vary with the type of regression model used—a single equation versus a system of equations, for instance—and should, of course, reflect the process described by that particular model.

The ordinary least-square procedure assumes that the association between the dependent and the explanatory variables is linear. It also assumes that the error term has an independent distribution, a mean equal to zero and a constant variance. It further assumes that the explanatory variables are independent of the error term. As stated above, assumptions may vary with the type of regression model: in the case of a system of equations, for instance, ordinary least squares can be applied only if it can be safely assumed that the system is recursive. It is also assumed that the variables are measured without error.

Population covered

Regression analysis can be applied to any category of population provided the required data are available. This method is easier to apply on a nation-wide basis, because aggregate data for a sufficient number of units are more likely to be at hand.

Time reference

No particular specifications are set regarding regression analysis: a regression model can be a cross-section model, a time-series model or a mixed model, and thus may encompass both cross-section and time series. In each case, results are in calendar years.

Simulation models

Type of method

Approach

A number of simulation models have been designed to study the occurrence of births, and some have been used to investigate the effects of family planning practice as well as other fertility determinants on the number of children born to a particular population. When used for evaluation purposes, the effect on family planning practice is the only programme factor taken into account. The models developed may be distinguished along several different dimensions. On the one hand, one may distinguish demographic models which focus on the vital events—i.e., birth, death, marriage—which change a population over time from biological models which decompose the birth probabilities into their biological components—fecundity, sterility, foetal loss, gestation and anovulatory periods etc. On the other hand, models may be distinguished as analytical or numerical, stochastic or deterministic (expected value models); and as macro-simulation (in which probabilities or proportions are applied to subgroups of the population of interest) or micro-simulation (in which probabilities are applied to individuals).

In studying the effects of a family planning programme on fertility, the usual practice is to simulate the natality process of two or more populations in which one or more would be subject to the programme and one or more would operate in the absence of the programme. By comparing the resultant fertility of say, two cohorts of women, similar in all respects except for the practice of family planning, an indication of the births averted by the programme is obtained.

Procedures

The procedures used to generate sequences of births over time vary with the type of model. In a deterministic model, the proportion of births expected is applied to women (or couples) with specified characteristics and the fertility outcomes are uniquely determined by the proportions used. In a stochastic model, a Monte Carlo simulation is used whereby women (or couples) are exposed to the appropriate probabilities in conjunction with sequences of random numbers. This procedure generates states or events on a stochastic basis.[25]

An analytical procedure, used mainly with biological models, resorts to the stochastic theory of renewal processes to simulate the reproductive history of women. It formulates the family-building process in terms of recurrent events and waiting times, with intervals between consecutive births taken as independent random variables having the same distribution. A particular formulation of this theory allows the initial birth interval (from marriage to first birth) to be distributed differently from the subsequent intervals. This procedure, however, requires a number of simplifying assumptions about homogeneity among women and over time.[26] Some of these assumptions are avoided by other models; for example, by a macro-simulation deterministic model called ACCOFERT.[27]

Factors affecting fertility change

Non-programme factors

Factors utilized. Since the natality process is approached through a variety of models, the factors taken into consideration can vary widely. A major difference, already noted, concerns models that do or do not take into account the biological factors leading to a birth. Another difference arises from the point in time when the simulation is initiated. If a group of women is taken through its reproductive life from age 15 to age 45 or 50, it is theoretically required that the probability of primary sterility, of dying before marriage and of marrying be included in the model, whereas a simulation of a group of fecund married women does not take these factors into account. Likewise, a model dealing with a hypothetical cohort of homogeneous women will resort to different characteristics than a model attempting to represent a more diverse hypothetical population. Non-programme factors include:

(a) *Socio-economic factors.* Only in cases where the models attempt to reproduce a realistic population are these factors taken into account. In such cases, the initial population can be simulated with due consideration to such factors as ethnic group, residence (urban or rural) and income;

(b) *Demographic factors.* First marriage, death, widowhood, divorce, remarriage, desired family size etc., are the most common factors in this category to be taken into consideration. Some models also consider infant or child mortality. Few models, however, take all these factors into account;

(c) *Biological factors.* Only models designed to simulate the reproductive process include biological factors. The most common factors taken into account by these models are fecundity, sterility, miscarriages, spontaneous abortions, stillbirths, length of gestation and anovulatory periods.

Data requirements and sources. The factors included in the models as parameters must be quantified in a sufficiently realistic way to make the results of the simulation meaningful; they must, therefore, be based on relevant empirical data. This is true of all the factors and particularly so for factors that may strongly affect estimates of programme impact. Thus, models that give strong weight to the non-susceptible period may require highly accurate data on breast-feeding and the anovulatory period. This requirement raises the question of sources, and although demographic data may be relatively accessible for the population under study, biological data are often much more difficult to obtain. However, these data may sometimes be supplied by imputation using data from available studies of similar populations.

Programme factors

Factors utilized. Simulation models are usually designed to assess the effects of one or several family planning methods on fertility; therefore, family planning practice and its use-effectiveness are the only programme factors used.

Data requirements and sources. The type of data needed depends mainly upon the parameters inherent

[25] See, for example, Albert Jacquard, "La reproduction humaine en régime malthusien", *Population* (Paris), vol. 22, No. 5 (September–October 1967), pp. 897–920.

[26] Edward B. Perrin and Mindel C. Sheps, "Human reproduction: a stochastic process", *Biometrics*, vol. 20, No. 1 (March 1964), pp. 28–45; and Robert G. Potter, "Renewal theory and births averted", in *International Population Conference, London, 1969* (Liège, International Union for the Scientific Study of Population, 1971), vol. I, pp. 145–150.

[27] Robert G. Potter, "Description of ACCOFERT II", Providence, Rhode Island, Brown University; and Ann Arbor, Michigan, University of Michigan (mimeographed).

in a particular model. Some models account for just one method of contraception; others include several methods, and still others also include sterilization or abortion as family planning methods. Available data on the age distribution of acceptors, continuation rates, use-effectiveness etc., drawn from such sources as service statistics and follow-up studies, may be needed as input by some models.

Measurement of fertility change

Levels of and change in observed fertility

In the usual approach, the observed fertility is the fertility obtained by simulating the natality history of a group whose family planning practice is assumed to result from programme activities. Unless the initial population is assumed to have already reached a given fertility level, changes in levels are not relevant since most simulations begin when both observed and potential fertility are still zero. Indeed, populations are often simulated from the beginning of their reproductive life.

Levels of and change in fertility in the absence of the programme

Levels of fertility in the absence of the programme are obtained from the simulation of women or couples who are not practising family planning or who are practising non-programme family planning methods.

Main assumptions

Each simulation model involves a number of assumptions, and the variety of models gives rise to a long list of assumptions that cannot be detailed here. Mention can be made, however, of some of the more general categories. One source of assumptions arises from the mathematical structure employed. For example, models based on renewal theory assume, in generating the reproductive histories, that the same probabilities apply to all women, that all parameters are fixed in time and are independent of age, and that the reproductive period is sufficiently long.[28] Monte Carlo methods, as well as others, do not necessarily have such constraints but produce results subject to sampling fluctuation. Differences between models also arise in treating factors as fixed or subject to certain distributions and, in the latter case, in the distributions employed. Models of course also differ in the number and type of factors incorporated, in the interrelationships assumed to exist among factors, and in the values assigned to the parameters of the model.

Population covered

The population covered is defined and characterized by the simulation model utilized. Certain models deal with single women at age 15; other models simulate cohorts of married women, beginning at age 20. In some cases, the hypothetical population is homogeneous; in other cases, an initial population is generated with heterogeneous characteristics. The population is thus model-specific.

Time reference

Most models are longitudinal studies of cohorts of women. Some model applications can, however, yield results in terms of calendar periods.

Major methodological issues

This section deals with some of the more important methodological issues encountered in the application of the evaluation methods described above. These issues, which have been selected somewhat arbitrarily, are as follows:

(*a*) Estimation of potential fertility;
(*b*) Data requirement problems;
(*c*) Correlated variables and interaction;
(*d*) Uncontrolled variables;
(*e*) Independence of methods;
(*f*) Cost-precision analysis.

In examining these issues, emphasis is placed on the definition of problems encountered in evaluation, rather than on possible solutions to the problems. The relation of each evaluation technique to each methodological issue is merely illustrated; the country case studies that were prepared for the meeting constitute the basis for a more systematic, analytical discussion. The purpose here is to provide a framework for discussions of the extent to which evaluation techniques currently in use are capable of solving methodological issues and of the advantages and disadvantages of each method with respect to those issues.

Potential fertility

As was made clear in the preceding section, a number of methods used to measure the impact of family planning programmes on fertility do so by comparing, under specific sets of assumptions, the actual fertility of a particular population with the fertility that population would have experienced had the programme not existed. "Potential fertility" is the concept identifying this hypothetical fertility, which can be defined as the fertility a population subjected to a programme would have experienced in the absence of that programme. Except with regard to evaluation methods that do not utilize this concept (i.e., correlation analysis), potential fertility is a major component of evaluation procedures and, as such, requires estimates of the highest possible accuracy.

Indeed, the difference (or the ratio) between actual and potential fertility indicators is assumed to yield the magnitude of the programme impact on fertility. Con-

[28] R. G. Potter, "Estimating births averted in a family planning program", *loc. cit.*, pp. 413–434.

sequently, any over-estimation or underestimation of one of the components of comparison will result in a corresponding over-estimation or underestimation of the magnitude of the programme impact. The computation of satisfactory estimates of the potential fertility of a population presents a number of problems.

Problems in assessing potential fertility

In order to deal with the problems of assessing potential fertility, it is necessary at the outset to bear in mind two factors. The first is that estimating potential fertility consists of determining a particular fertility level that did not materialize. It is not possible to know with certainty what such a fertility level would have been and this problem cannot be solved. The purpose of the procedure is thus to compute reasonable estimates. The second consideration is that there are two types of population which can be observed for evaluation purposes: (*a*) the population living within defined boundaries, namely, that living in a country, an administrative area or a family planning programme area;[29] (*b*) the population known to participate in the family planning programme and which constitutes the group referred to as "acceptors". Problems are somewhat different for these two groups and are treated separately.

Potential fertility of the general population

Conceptual problems. Substantial efforts have been made to define the concept of "acceptor" and its significance, but no such attempts appear to have been made in respect of the "general population". The concept of "general population" is ambiguous, because there are no criteria for defining it and researchers rarely attempt to identify the potential biases that may occur when a given population is used as the general population. Thus, the general population is usually that which the evaluator arbitrarily selects on the basis of unspecified criteria. Theoretically, the general population can include all women in the area under study, or non-acceptors only, or married women only, or non-sterile, non-pregnant women only, or women who did or did not give birth during a defined period of time preceding the interview etc. The estimate of potential fertility will vary, depending upon which of these groups is taken as the general population. The problem is thus to develop a concept of the general population that would provide, under specified conditions, a satisfactory estimate of "potential fertility".

Methodological problems. Generally speaking, there are basically two distinct approaches to estimating the potential fertility of a general population. In the first, some set of observed fertility rates are equated with potential fertility; in the second, a set of observed fertility rates is combined with various assumptions to produce the estimates of potential fertility. When a family planning programme is introduced into a country or area, the fertility levels observed prior to and at the time of introduction in that country or area may be taken as the potential fertility that would have occurred in the absence of a programme. In effect, the prior experience of the country or area serves as a pseudo-control group. In a more strict experimental design, if the programme is introduced into certain areas only, non-programme areas deemed otherwise comparable may be selected as the control group; and the fertility for those areas over the same time period for which the programme was operational in the experimental areas is taken as a measure of potential fertility. In the latter case, the adequacy of the approach depends, among other things, upon the comparability of the two types of areas and upon the absence of contamination of the non-programme areas by the programme activities ongoing in the experimental areas.

In the second approach to estimating potential fertility, one projects the past fertility of the population under observation into the future, on the assumption that no family planning programme had been initiated. This approach involves, as a first step, proper identification of the past trend, which will be influenced by the length of the reference period used and the method of sorting out short-term fluctuations. Then, assumptions are made about the nature and magnitude of social and economic factors that might have affected fertility even in the absence of a family planning programme. Though it is sometimes helpful to establish a range of values for potential fertility under different assumptions, the possibility of inferring anything useful about the effect of the programme diminishes as the range for potential fertility increases.

The evaluation methods that explicitly or implicitly produce a measure of potential fertility for the general population use one or another of the two approaches outlined above and, in addition, differ in the way they take into account the social and economic changes that may affect fertility. The experimental design method, for example, utilizes the first approach and assumes that any social changes that occur over the time period of observation affect the control and experimental groups equally. In the decomposition or standardization procedure, it is implicitly assumed that the factors used for standardization capture the changes bearing on fertility. A projection approach, which simply extrapolates pre-programme trends into the period of observation, is implicitly assuming that the rate of change of non-programme factors is the same in both periods, an assumption which may or may not be realistic.

Potential fertility of acceptors

The evaluation methods based on data pertaining to acceptors emphasize the usefulness of information on

[29] The population defined by geographical boundaries is, for the purpose of this discussion, called the "general population"; this term is meant also to include the population of a community, a neighbourhood, a maternity ward etc.

persons who have been reached by the programme and the consequent changes in their fertility. It is assumed that if acceptors can be identified, changes in their fertility can be credited more easily to the programme than is possible when dealing with the whole population. Thus, for practical reasons[30] and to increase the probability that a correct inference is made, a number of methods tend to limit the evaluation of programme impact to those who are defined as acceptors.

Conceptual problems. As concerns the concept of "acceptors" it would appear at the outset that a fertility decline among acceptors might be readily linked to programme activities. There are reservations regarding this relationship, but the problem of covariance between these two variables is not the only question. A preliminary question is how the concept of "acceptor" is to be understood. Definitions of this concept vary to the extent that they affect markedly the composition and size of a group of acceptors and, consequently, also the impact on fertility. An acceptor may be defined as a person who accepts and uses a contraceptive method (or a sterilization or an abortion) offered by either private or public family planning programmes. Whether the acceptor participates in the programme for the first time or whether he or she accepts a particular programme method for the first time is also of particular relevance. A distinction may also be made between "acceptors" having used contraceptives prior to entering a programme and persons who accept (and use) a family planning method for the first time.[31]

A number of studies of acceptors' fertility prior to acceptance have indicated that this group may be a self-selected group characterized by higher than average fertility. If the past fertility of this group is above average, it may be hypothesized *a priori* that in the absence of the programme, the acceptors' fertility level would have remained above average and, hence, that their potential fertility should be estimated on the basis of their own past fertility. This assumption is critical to the whole problem of estimating the acceptors' potential fertility and raises, among others, the following questions: (*a*) whether acceptors have genuinely higher than average fertility; (*b*) if so, whether their fertility would have remained at a higher than average level had the programme not been undertaken.

With respect to the question of the acceptors' genuinely higher fertility, two levels of measurement are implied in the concept of genuine fertility. One is the amount of "excess" fertility that is observed in empirical studies comparing acceptors with non-acceptors or with the general population. The other concept of excess fertility, defined here as "genuine" higher fertility, refers to the assumption that acceptors have a higher than average probability of producing a live birth. Genuinely higher fertility cannot be directly observed but only inferred, and this assumption is critical in deciding whether higher than average fertility would have continued to be experienced by acceptors had the programme not existed.

References about genuinely higher fertility and its effect on potential fertility will be aided by decomposing the differentials in observed fertility between acceptors and non-acceptors. To the extent that one can determine whether the observed differences are due to the factors related to exposure to intercourse (length and stability of sexual union); exposure to risk of conception (e.g., coital frequency, use of contraception, breast-feeding patterns); probability of conceiving (fecundability, primary and secondary sterility) and risks of foetal wastage, one is better able to judge whether the observed fertility differentials are likely to persist in the future.

Certain features of family planning programmes help distinguish acceptors and non-acceptors on these various components. For example, in so far as women become acceptors shortly after a birth, they are less likely to be sterile than a woman of the same age selected at random from the population, who may be primarily sterile or who may have developed secondary sterility since her last birth, which occurred, on average, earlier than that of the acceptor. The differential in the likelihood of sterility will be relatively large at the older ages. By the same token, if the observed difference in fertility between acceptors and non-acceptors is based on a short period of observation, the difference may be largely a reflection of the fact that acceptors have had a recent birth rather than any difference in fecundity.

Though differences in observed fertility between acceptors and non-acceptors are often taken to imply differences in fecundity, this assumption must be examined carefully even when based on moderately long periods of observation. Simulation studies[32] have shown that chance factors in fecundity operating on a group with identical probabilities will produce a subgroup with above-average fertility, which is thereby disposed towards acceptance.

The chance factor in fecundity also complicates, in another way, the assumption that acceptors would continue to have higher fertility than non-acceptors. Brass[33] has shown that even if acceptors are correctly assumed to have higher fecundity than non-acceptors, due to the chance factor, the future differential in fertility may be less than that observed in the past. That is, over short durations in particular, the chance

[30] As is shown in the discussion of "interaction", there are serious difficulties in identifying who has been influenced by the programme among those who are not recorded for having done so.

[31] This procedure requires accurate and detailed record-keeping; one should be able to identify not only "acceptors" who are simply switching methods but those who increase their use-effectiveness as a result of the programme.

[32] Jeanne C. Ridley and others, "On the apparent subfecundity of non-family planners", *Social Biology,* vol. 16, No. 1 (March 1969), pp. 24–28.

[33] William Brass, "Assessing the demographic effect of a family planning programme", *Proceedings of the Royal Society of Medicine,* vol. 63, No. 11 (November 1970), pp. 29–31.

element may be more important than systematic differences in fecundability.[34] Moreover, empirical studies that compare the post-acceptance fertility of acceptors with non-acceptors may be misleading when based on short periods of observation, in that acceptors are rarely pregnant at the time of acceptance, but the comparison group of non-acceptors will have a certain proportion of pregnant women.[35] Thus, the period of observation is important when comparing acceptors and non-acceptors, with regard both to pre-acceptance fertility and to post-acceptance fertility.

Another conceptual difficulty has to do with the so-called "substitution problem". Acceptors are often distinguished from non-acceptors not only by higher fertility but by a greater interest in family planning. In determining the potential fertility of acceptors one must take into account the possibility that some acceptors are substituting programme methods of family planning for methods privately obtained. In this case, one would want to credit the programme only with any net gain in effectiveness that resulted from this substitution. Another facet of this problem is the possibility that acceptors who are not current or prior users of family planning might have adopted some method privately even in the absence of a programme. These problems cause severe complications in the estimate of potential fertility and permit no ready solution even at a conceptual level.

As Wolfers'[36] review of several studies indicates, the proportion of ever-users who are current users varies greatly. The fact that a woman was using a family planning method at the time of acceptance or at some time previously does not demonstrate that she would have used a method privately if she had not become an acceptor.

It should also be mentioned in passing that in so far as attention is restricted to the potential fertility of acceptors, the programme cannot be credited with any indirect effects of encouraging couples to adopt family planning privately. Indeed, if this practice takes place to any appreciable extent, it may detract from the apparent impact of the programme by diminishing the fertility differentials between acceptors and non-acceptors.

Methodological problems. With respect to measurement of the acceptors' past fertility the various conceptual problems discussed above would be meaningless if the acceptors' fertility were measured as higher than average only because of the unreliability of the data or method, or if the magnitude of the excess fertility were over-estimated or underestimated for similar reasons. The measurement of the acceptors' excess fertility is obtained by comparing fertility data of acceptors, on the one hand, and those of non-acceptors or the general population, on the other hand. The difference between these observed rates is considered to be the excess fertility. Various factors can affect the magnitude of this difference. If comparisons are based on samples, both the absence of biases and the significance of the fertility differences have to be ascertained. Excess fertility of acceptors can vary, depending upon whether one uses non-acceptors or the general population as the unit of comparison. Excess fertility will also vary with respect to the accuracy of the data being compared, especially with respect to the difference in accuracy between the two sets of fertility indicators.[37]

The length of the span of time over which fertility is estimated may also affect the estimated level of fertility and, hence, the magnitude of the acceptor's excess fertility. In addition to points discussed previously, it should be noted, for example, that if acceptors enter a programme shortly after a birth, an upward bias may be introduced if too short a reference period is used. Estimates of the acceptors' own fertility prior to the programme is often recorded for three- or five-year periods. Assuming that the length of the reference period is the same for both groups, the question of an optimum length—and minimum bias—to estimate fertility for evaluation purposes is still unanswered.

As concerns measurement of the acceptors' potential fertility a number of procedures utilized to estimate potential fertility have been outlined in the preceding section. The discussion here deals only briefly with some issues that may affect the validity of a particular procedure. One problem is whether the fertility of the general population or that of the acceptors themselves should be used as a basis for calculating potential fertility. Most methods use the acceptor's own fertility on the assumption that their fertility is and will remain higher than average. Only one author[38] recommends the use of data obtained from selected groups of the general population.

Another question is what kind of fertility measure is best for the purpose of estimating potential fertility. This problem has been raised mainly in connexion with Potter's and Wolfers' procedures,[39] the former using age-specific fertility rates and the latter birth intervals. Both procedures are based on a comparison of "units of time", namely, the period of useful retention or use of a contraceptive during which no conception occurs, as compared with the average interval between two births

---

[34] David Wolfers, "Births averted", in C. Chandrasekaran and Albert I. Hermalin, eds., *Measuring the Effect of Family Planning Programs on Fertility* (Liège, International Union for the Scientific Study of Population for the Development Centre of the Organisation for Economic Co-operation and Development, 1975), pp. 163–214.

[35] W. Brass, *loc. cit.*

[36] D. Wolfers, "Births averted".

[37] John A. Ross, "Cost of family planning programs", Bernard Berelson and others, eds., *Family Planning and Population Programs: A Review of World Development* (Chicago, University of Chicago Press, 1966), pp. 759–778.

[38] D. Wolfers, "The demographic effect of a contraceptive programme".

[39] R. G. Potter, "Estimating births averted in a family planning program"; and D. Wolfers, "The demographic effect of a contraceptive programme".

had the contraceptive not been used. To compute the potential fertility estimate, Potter records the acceptors' own fertility rates for a given pre-programme span of time; he then translates these rates into a unit of time defined as the average duration per birth that might have been required had the IUD not been adopted. The translation of birth rates into time periods is obtained by using a method based on the reciprocity of birth intervals and fertility rates.[40]

Wolfers[41] questions this procedure for obtaining mean birth intervals, preferring to obtain the birth intervals required for estimating potential fertility directly from interviewing randomly selected women at the time of delivery or from women all of whom have given birth within a certain specified period.

Conclusion

A number of approaches to solving some of the problems of estimating potential fertility rely on model simulation. According to one suggestion, for instance, use should be made of a theoretical model to determine the appropriate length of the pre-programme reference period so as to minimize the biases in estimating past fertility. Use of stochastic models of human reproduction has also been recommended as a means of estimating the expected number of births during specific periods of time in the absence of a family planning programme.

As the estimate of potential fertility is crucial to a number of evaluation procedures and is beset with many difficulties, close examination of the advantages and disadvantages of the various approaches is warranted.

Data requirement problems

The problems referred to in this section concern a selected number of difficulties associated with the selection, measurement, utilization and interpretation of data needed for applying methods of measuring the impact of family planning programmes on fertility. The quality of the data and the use of appropriate measures are, among other data requirements, major preconditions for drawing meaningful conclusions. Even the more sophisticated evaluation procedures cannot be expected to provide reliable results if the data inputs are unsatisfactory. It is to be anticipated that some problems cannot be completely solved. Their presence should, however, be explicitly acknowledged at the interpretation phase so as to give a proper perspective to the conclusions drawn by the evaluator.

The data problems reviewed and the items discussed are illustrative rather than exhaustive. Special emphasis is given to the use of fertility data, since measuring whether an observed difference in fertility over a given period of time reflects actual changes is a major component of many methods. Common statistical problems, such as the necessity of testing for significance of differences, are therefore mentioned. Specific data problems peculiar to certain methods can be examined in more detail when the country case studies are reviewed. Most of the problems, which are illustrated by fertility data, also arise in connexion with other demographic data, as well as with standard socio-economic data. The items discussed are as follows: (*a*) measurement problems are examined as they appear to be a major source of obtaining invalid results; (*b*) a number of questions are raised regarding the interpretation of results, even when data can be assumed to have been measured without errors.

Measurement problems

The measurement problems reviewed here are illustrative of a number of difficulties that are expected to be encountered in the measurement of the impact of family planning programme on fertility. These problems result from a variety of sources: unfavourable data-collection circumstances as they are found in the developing countries; faulty or unreliable data-collection procedures; reporting errors; classifications errors; chance errors, etc. An extensive body of literature is available on the major types of errors that are encountered in both sample surveys and complete enumeration, and on solutions that may be applied. The primary concern here is with a selected number of questions that might arise in connexion with evaluation studies and that may be illustrated by the case studies that complement the present review.

The problems at issue are specifically errors in measurement. They fall, for the purpose of this discussion, into two categories. The first category includes non-random errors and biases that are either constants and affect all units alike or that follow a particular distribution and affect the true values differentially. Although errors in this category are often suspected, it is nevertheless difficult to ascertain their magnitude and direction. Information originating outside the scope of the particular study is often required, or specific new information is generally needed to establish the nature and magnitude of these non-random errors. Once these errors are established, the observed data are often subject to adjustment as in the case of the age distribution of a population. Assessing the presence and magnitude of the errors, on the one hand, and determining the quality of the adjustments undertaken, on the other hand, constitute the major issues of measurement as far as non-random errors are concerned.

The second category of errors includes random errors. They are classified separately because of their

---

[40] The actual formula is $D(w) = 12{,}000/b(w)$, where $b(w)$ denotes the birth rate for women of average age w and $D(w)$ is the corresponding duration, in months, per birth. See R. G. Potter, a technical appendix on procedures used in manuscript "Estimating births averted in a family planning program".

[41] D. Wolfers, "The demographic effect of a contraceptive programme"; and *idem*, "The estimation of potential fertility for family planning evaluation", *Proceedings of the Royal Society of Medicine*, vol. 63, No. 11 (November 1970), pp. 41–44.

stochastic nature. Specifically, the approach to random errors is quite different from that of non-random errors, mainly because they can be dealt with by means of statistical theory. A large body of research has been devoted to this type of error; it is thus expected that some of the solutions and insights can be of use to problems encountered in evaluation studies. Although standard statistical theory considers the total error in measurement as resulting from sampling and non-sampling errors, the terminology used here, namely, random and non-random errors, is utilized because the category of random errors includes both sampling errors (i.e., random errors due to the sampling procedure) and random errors of non-sampling origin (i.e., random errors resulting from chance factors that are assumed to occur even in complete enumerations).

From a conceptual standpoint, the measurement of a variable can be shown in the formula:

$$X_i = X'_i + B_i + e_i \text{ where the error } = B_i + e_i$$

where X = the observed value of the individual characteristic being measured;
X'_i = the true value of that characteristic;
B_i = the bias;
e_i = the random error associated with the measurement.

The survey bias can be further defined both for total enumerations and for sample surveys:

$$B = E(X) - E(X') \text{ or } B = \bar{X} - E(X')$$

where B = the survey bias;
$E(X)$ = the expected value of the observed individual characteristics of the universe;
$\bar{X}$ = $E(\bar{X})$ = the expected value of the sample mean; and
$E(X')$ = the true mean of the universe.

Non-random errors

Non-random errors constitute a major source of biases in the measurement of observed variables, and the field of family planning programme evaluation is not exempt from this problem. Generally speaking, the problem of errors can be approached in two steps. The first consists of assessing the existence and the magnitude of suspected errors; the second step consists of adjusting the erroneous data. In the social sciences, in general, and in the field of demography in particular, such sources of errors as under-enumeration, non-responses, digit preferences, recall lapses, constitute major sources of biases, whether data are obtained from samples or from attempted complete enumerations. The experience gained by demographers in the gathering and adjustment of data often makes them readily aware of inaccurate results. But estimating the magnitude of such errors and correcting them presents great difficulties, as do their effects, which constitute yet another problem.

Although the assessment of random sampling errors can be undertaken directly on the basis of statistical theory and techniques of inference, non-random errors require some complementary information for assessing both the existence and the magnitude of the suspected biases. Response errors, for instance, as well as under-reporting, can be estimated by matching the data with statistics from other sources. The effect of non-random errors may not only affect the measurement of a particular indicator but, under given conditions, may affect the measurement of the accuracy of sample results. Assuming, for instance, that a simple random sample n is taken from a universe N where N_1 would respond and N_2 would not, the proportion $N_1/N = W_1$ is then the proportion of response in the population, and $N_2/N = W_2$ is the proportion of non-response in the population. The amount of bias in the sample mean x resulting from non-responses would be

$$\begin{aligned} E(\bar{x}_1) - \bar{X}' &= \bar{X}_1 - \bar{X}' \\ &= \bar{X}_1 - (W_1\bar{X}_1 + W_2\bar{X}_2) \\ &= W_2(\bar{X}_1 - \bar{X}_2) \end{aligned}$$

where the bias is equal to the product of the proportion of non-response and the difference between the two means in the two population strata.[42] The question is how such theoretical findings can be of benefit in evaluation studies: in other words, how the amount of bias can actually be estimated and what magnitude can be considered negligible, say, with fertility data.

Generally speaking, when sampling errors are estimated on the basis of the standard deviation, the estimate of the error is not satisfactory when the bias B is large compared with the standard error.[43] In the presence of large biases, the square root of the mean square error, $\sqrt{\text{MSE}}$, appears to be the best measure of accuracy of an estimate. The mean square error, which is equal to the variance of the estimate plus the square of the bias, is obtained, for a sample mean, as follows:

$$\text{MSE}_{\bar{x}} = \sigma^2_{\bar{x}} + (\bar{X} - \bar{X}')^2$$

where $\bar{X} = E(\bar{x})$ and $\bar{X}'$ is the true value. The effects of biases have been studied on the basis of specific models[44] and, according to Cochran,[45] have led to the conclusion that constant biases pass undetected in the sample data and that the computation of confidence

[42] W. G. Cochran, *Sampling Techniques* (New York, John Wiley and Sons, Inc., 1961), p. 294.

[43] According to Cochran, the effect of a bias on the accuracy of an estimate is negligible if $B/\sigma < 0.10$.

[44] Models used are of the type $X_i = X'_i + B_i + e_i$, where X' is the correct value; B_i a constant or a variable bias; and e_i the random component.

[45] W. G. Cochran, *op. cit.*, p. 307.

limits become, under specified conditions, misleading. The problem here is to determine under what conditions the effects of biases can be estimated and corrected as far as evaluation studies are concerned.

Selection of an appropriate technique for adjusting non-random errors is a difficult matter. A number of methods are available and the question of relevance is what method is appropriate in particular circumstances. Som[46] indicated three methods as being appropriate to correct the effects of non-response: the use of selected random substitutes from the responding units; the Hansen-Hurwitz method, i.e., a random subsample of the non-responding units; and the Hartley-Politz-Simmons method which proposes a means of adjusting for biases without resubmitting questionnaires. Then, a number of methods have been devised specifically for handling biases common in demographic inquiries, for example the Coale-Demeny stable population approach, the Brass method and the Som method for recall lapses.[47] Of course, this raises the basic question whether these methods can be used properly for evaluation purposes. The first aspect of this question concerns the requirements set forth for using the techniques. The Brass method, for instance, assumes relative constancy of age-specific fertility as does the stable population approach.[48] These techniques assume, in fact, precisely the contrary of what the evaluation methods are expected to measure, namely a fertility decline due to family planning practice. The second aspect is whether these methods provide sufficiently precise corrections for estimating small fertility changes that may occur over short periods of time.[49]

There are alternative techniques that can be used to adjust for suspected or known errors. The use of interrelations, such as linearity, between proportions tabulated as childless and as "parity not given" can be used to adjust such errors.[50] Corrections have also been made on the basis of data of populations assumed to have "similar" characteristics. In various cases, where phenomena are not directly observable, such as fecundability or post-partum anovulation, missing information is based on indirect observations and assumptions. Fecundability estimates have been obtained by Henry from observed fertility data and are still used, with some adjustments, for modelling and estimating programme impact in developing countries.[51] The duration of post-partum anovulation has been equated, in some cases, with the length of the breast-feeding period. Since a number of evaluation studies take no account of post-partum amenorrhoea, the question raised here is twofold: what is the magnitude of the bias introduced by ignoring such a factor; and how valid are the assumptions used to perform any particular adjustment. It will be interesting to see how the case studies account for such problems.

Non-random errors may affect all categories of variables; and in some cases, depending upon the factor affected, they may have greater or less influence on the final results of the evaluation. The accuracy of the fertility data, for instance, is crucial and an observed fertility decline should, of course, be verified on the basis of strong evidence. Because non-random errors are difficult to pin-point, they present particularly dangerous features when affecting, for example, the socio-economic variables utilized in the matching procedure or in the regression analysis approach. In the latter case, there can be as many errors as there are variables in an equation; and if several equations are involved, errors attached to one equation may reappear in another. Since under unsatisfactory data-gathering conditions the amount of inaccuracy due to non-random errors may be large, compared with random errors, particular attention should be paid to this category of errors.

Random errors

Two types of non-exclusive random errors can be thought of conceptually. One type includes random errors resulting from the use of a sampling procedure. In this case, an estimate of the sampling errors can be obtained on the basis of statistical theory and inference techniques. The existence of such errors arises from the sampling, and only their magnitude and direction need be estimated. The sampling error is commonly estimated by the standard deviation of the estimates. Naturally, the magnitude of the standard error has important implications for any conclusions; sampling techniques and sample size can be set up to control for this type of error. However, as mentioned earlier, biases affect the measure of accuracy of sample estimates. Control for sampling errors and level of significance through sample size needs to be done if biases are suspected, on the basis of the root mean square error as exemplified by Seltzer.[52]

The other type of random error is not necessarily associated with sampling procedures. These errors are assumed to derive from the net cumulative effect of a large number of small influences originating in a variety of factors difficult to identify. They are associated

[46] R. K. Som, *A Manual of Sampling Techniques* (London, Heinemann, 1973), pp. 296–297.

[47] *Manual IV. Methods of Estimating Basic Demographic Measures from Incomplete Data* (United Nations publication, Sales No.: 67.XIII.2), pp. 7–30 and 31–40 and R. K. Som, *Recall Lapse in Demographic Inquiries* (Bombay, Asian Publishing House, 1973).

[48] *Ibid.*, p. 33 and p. 46.

[49] P. M. Hauser, "Family Planning and Population Programs. A Book Review Article", *Demography* (Washington, D.C.), vol. 4, No. 1 (1967), p. 406.

[50] M. A. El-Badry, "Failure of enumerators to make entries of zero errors in recording childless cases in population censuses", *Journal of the American Statistical Association* (Washington), vol. 56 (December 1961), pp. 909–924.

[51] See, for instance, D. Wolfers, "The demographic effect of a contraceptive programme", p. 118.

[52] William Seltzer, "Measurement of accomplishment: the evaluation of family planning efforts", *Studies in Family Planning*, vol. 1, No. 53 (May 1970), pp. 9–16.

with phenomena assumed to be of a stochastic nature, such as fecundability, and sometimes with the unpredictable element of randomness found in human responses.[53] In some cases, random errors are also assumed to result from the fact that indicators generally do not adequately represent theoretical concepts; in other cases, they are merely errors of observation that happen to be distributed at random.

Often, the presence of random measurement errors may be suspected if the observed variable fluctuates when measured on an annual basis. In the light of this observation, it may be underlined that measuring fertility changes, especially over short periods of time, becomes difficult even if no systematic bias is assumed. Annual measures of fertility may be subject to variations of considerable magnitudes. When fertility is measured at two points with the view of determining changes in its level, it is of crucial importance to determine what part of the change reflects a trend and what part may reflect random factors. Assuming that in a time series the observed crude birth rate B_y is the result of the sum of two unobservable components, a polynomial of degree n (or less) and random disturbance component:

$$B_y = f(y) + e_y$$

Seltzer and Fand[54] attempted to estimate the level of disturbance associated with annual crude birth-rate series and to examine how stable these estimates of residual variability appear to be.

The amount of misinformation regarding the change is, of course, greater when the period of observation is shorter, and fertility is sometimes averaged over several calendar years in order to attenuate chance effects. In this case, the fertility level and, hence, the amount of change becomes a function of the length of the reference period utilized. The question is, therefore, what optimum length of the reference period can be used to obtain the best estimate of fertility, and how such an optimum length can be determined. One proposed solution was the use of simulation models. While the focus here is on fertility data, it should be borne in mind that all variables can be subject to measurement errors of a stochastic nature. This point is important notably for applying regression analysis procedures to assess programme impact, where both dependent and independent variables can be affected.

Determining the effects of random measurement errors and, eventually, the amount of misinformation resulting from such errors appears to be a complex undertaking. When random errors are studied, it is generally assumed that there are no other systematic biases. In some cases, conclusions can be reached easily. If, for instance, a factor X_i is measured with:

$$X_i = X'_i + e_i \text{ and } E(e_i) = 0$$

where X' is the true value and e the random error, it can be shown that under specified assumptions the expected value of X_i is not affected by the error. Its variance, however, is affected and is shown to become:

$$\sigma_x^2 = \sigma_{x'}^2 + \sigma_e^2$$

The question, then, is whether such a bias is important or negligible and under what conditions it would affect conclusions.

The assessment of the effects of measurement errors appears somewhat more complex when correlation and regression analyses are under scrutiny. The misinformation yielded by correlation or regression coefficients, for instance, varies with the type of regression model used, the assumptions made, the estimating procedure utilized etc. Research on regression parameters, for example, have concluded that in the case of a two-variable linear relation of the type:

$$Y = a + bX + w$$
$$\text{with } X = X + u,\ Y = y + v, \text{ and } y = a + bx$$

where X and Y indicate the observed measures, x and y the true values; and u, v and w the measurement errors, the slope coefficients estimated through the least-square method are, under specified assumptions, underestimated, compared with the true values.[55] A similar conclusion has been reached regarding the squared multiple correlation coefficient whose value becomes reduced in multivariate normal models with independent errors of measurement.[56] Models for tracing the effect of more complex error patterns, as when there is correlation among the error terms or when the error is correlated with the variable being measured, have also been developed.[57]

---

[53] J. C. Barrett and W. Brass, "Systematic and chance components in fertility measurement", *Population Studies*, vol. 28, No. 3 (November 1974), p. 473, underline the fact that even if couples had the same fecundity characteristics, there would still be differences in family size because of chance effects.

[54] William Seltzer and R. S. Fand, "A note on the annual variability of the crude birth-rate", *Proceedings of the Social Statistics Section, 1973* (Washington, D.C., American Statistical Association, 1974), pp. 326–391.

[55] See J. Johnston, *Econometric Methods* (New York, McGraw-Hill, 1963), p. 150. Johnston shows that in the case of the two-variable model and on the assumptions, notably, that the errors are independent of one another and of the true values, B_n, the least-square estimator of b on the basis of n sample observations is both biased and inconsistent, with:

$$\text{plim } B_n = \frac{b}{1 + \sigma_u^2/\sigma_x^2}$$

[56] W. G. Cochran, "Some effects of errors of measurement on multiple correlation", *Journal of the American Statistical Association*, vol. 65, No. 329 (March 1970), p. 22.

[57] Paul M. Siegel and Robert W. Hodge, "A causal approach to the study of measurement error", in *Methodology in Social Research*, H. M. Blalock and Ann B. Blalock, eds., (New York, McGraw-Hill, 1968), pp. 28–59; and John J. Chai, "Correlated measurement errors and the least-squares estimator of the regression coefficient", *Journal of the American Statistical Association*, vol. 66, No. 335 (1971), pp. 478–483.

Though the models described above are helpful in understanding the effects of measurement error, the results obtained hold only for the assumptions specific for a model and wider generalizations might prove misleading and unwarranted. In practice, it is often difficult to determine which error model is most appropriate for the data obtained; and, in many cases, the actual pattern of measurement error will be more complex than that specified by any model. This situation has led one observer to conclude that efforts must be directed at eliminating measurement error at the source, as the models that account for it are only simplifications of the reality.[58]

Questions of interpretation

The correct interpretation of analytical results is a crucial phase of all scientific investigations. Attention is drawn here only to aspects of particular relevance to the evaluation of family planning programmes and to the problems of data discussed in this section.

An important aspect of any interpretation is acknowledging the presence of any errors in the observed data and assessing their effects on the results. As mentioned above, this effort can often be accomplished to a limited degree, particularly with non-random errors. As a result, there is some tendency to focus on sampling errors, for which statistical techniques are often available. Even here, however, care must be exercised to utilize techniques appropriate to the data. For example, estimates of sampling error based on simple random sampling may be misleading when a more complex sampling design has been employed, as is often the case. Where possible, techniques appropriate to the design should be employed.[59] At the same time, non-sampling errors usually merit more attention than they are given because they often constitute a large part of the total error. It should also be recognized, however, that in certain circumstances, biased results may be acceptable if they have a small variance and lie closer to the expected value than unbiased results with large variances.

Non-random errors are difficult not only to observe but to correct. At times, however, methods for correcting unreliable, missing or incomplete data are undertaken. In such cases, attention should be given to the robustness of the procedure employed when the data deviate from the assumptions of the technique. An assessment should also be made of the precision of the correcting technique in relation to the magnitude of the event being measured. In some cases, for example, the range of precision associated with a technique may be large in relation to the amount of change observed. This is a troublesome problem in the evaluation of family planning programmes, where one is often dealing with rather small and short-range changes in fertility and other key variables.

As so many of the variables used in evaluation are rates of one type or another, it is worth noting that assessment of error should be carried out both for the numerator and for the denominator. Often it is the numerator, such as the number of acceptors or the number of births, that is more problematical, but the denominator cannot be assumed to be error-free. This procedure is particularly important when changes in rates are a major component of evaluation, as a relatively small change in the degree of error in the denominator can account for a large proportion of the total amount of change observed. Caution is also required with respect to rates or ratios when sampling error is being assessed. When the numerator and denominator are both random variables, the sample ratio is not an unbiased estimate of the population ratio, though in most applications the amount of bias is very small.[60]

In interpreting the results of an analysis, it is also desirable to assess the effect that alternate definitions or estimates of variables might have on the results. In the evaluation of family planning programmes, this need can arise in many ways. For example, note has been taken previously that the concept of acceptors may be made operational in a number of different ways; and where a specific definition has been used, consideration should be given to the possible effect on the results of alternative definitions. The same consideration applies to other key concepts, such as potential fertility and to fertility itself. A related concern is the sensitivity of results to estimates utilized in the evaluation procedure. Would slight changes in these estimates affect the results to an appreciable extent?

On a somewhat more general level, it is well at the interpretation stage to consider the possible effects of alternate specifications of the models utilized: whether any variables have been omitted from the model which might affect the results appreciably; or whether alternative assumptions about the interrelationships among the variables would produce different answers. Although a final resolution to such issues is not to be expected, an awareness of their importance will contribute to comparative analytical investigations which will lead to more reliable and secure results.

In addition to the foregoing aspects, appropriate inference must also be attentive to the level of analysis. When the unit of analysis is an area, the results obtained should not be taken as holding among individuals because, except in very particular circumstances, analyses at different levels of aggregation will not produce the same results. If areas are employed, the evaluation problem should be formulated as appropri-

---

[58] R. Schoenberg, "Strategies for meaningful comparison", in Herbert L. Costner, ed., *Sociological Methodology, 1972*, (San Francisco, Jossey-Bass, 1972), pp. 1–35.

[59] L. Kish and M. R. Frankel, "Inferences from complex samples", *Journal of the Royal Statistical Society*, vol. 26, No. 1, series B (1975), pp. 1–37.

[60] L. Kish, N. K. Namboodiri and K. Pillai, "The ratio bias in surveys", *Journal of the American Statistical Association*, vol. 57, No. 300 (December 1962), pp. 863–876.

ate to that level and the interpretation should be consistent with the aggregate nature of the data.

Correlated variables and interaction: interaction problems

On the general level, measuring the impact of a family planning programme on fertility is an attempt to estimate the magnitude and direction of the effects of specified programme and non-programme factors on fertility. From this view, one identifies a set of cultural, socio-economic, demographic and programme factors as determinants of fertility. The variables identified, plus their posited interrelationships, constitute in effect a system or theoretical model within which it is usual to distinguish the main dependent variable, fertility; the independent or exogenous variables, which are not determined within the model; and a set of intervening or intermediate variables, which are determined by the independent factors and which also have an effect on the dependent variable.

From this standpoint, it may appear that a fruitful system or model will account for or "explain" a high proportion of the variation in the dependent variable, in this case changes or levels of fertility, and allow an allocation of this proportion into programme factors and non-programme factors. However, the nature of the interrelations among the variables in the model may prevent the successful decomposition of the explained variance in this manner. The difficulty arises in many models because of correlations between the explanatory variables and/or interaction effects. In what follows, these difficulties are first illustrated with reference to multiple regression, and then their relevance for other evaluation techniques is shown.

Correlated variables in multiple regression

Consider a three-variable linear multiple-regression equation of the type:

$$Y_i = a + b_{y_{1\cdot 2}} X_{1i} + b_{y_{2\cdot 1}} X_{2i} + e_i \qquad (1)$$

where Y = a measure of fertility;
X_1 = a non-programme variable;
X_2 = a programme variable.

From the data, it is possible to estimate values of the partial regression coefficients, $b_{y_{1\cdot 2}}$ and $b_{y_{2\cdot 1}}$, which measure, respectively, the effect of the non-programme variable on fertility after taking into account or "holding constant" the programme factor; and the effect of the programme variable on fertility, after taking into account or "holding constant" the non-programme variable. The regression coefficients show how much change in fertility is expected from a change of one unit in the independent variables. They thus indicate the relative importance of each variable and are useful in cost-benefit analysis when the relative cost of achieving a change in each independent variable can be determined.

It is also possible to determine from the data the proportion of the total variation in fertility accounted for by both independent variables, sometimes referred to as the coefficient of determination or R^2. Only in special circumstances, however, can this proportion be decomposed into the amount due to X_1 and the amount due to X_2.

If the two independent variables are not correlated with each other, then the regression coefficients obtained above will be the same as those obtained from two simple regression equations in which fertility is regressed against each of the variables in turn. Similarly, the coefficient of determination, R^2, obtained above will be equal to the sum of the proportion of variation obtained from each of the two simple equations. However, these relations do not hold if the two independent variables are correlated. In this case, the regression coefficients from the multiple regression equation may be greater or less than that obtained from the simple bivariate case, while the multiple R^2 likewise may be greater or less than the sum of the R^2 from each equation.

Since in most systems the factors do not exercise their influence independently of one another, attention is directed to the pattern of interrelationships or structure of the model, assumed to hold among the variables. Indeed, as one observer states,[61] insufficient attention is accorded to appropriate representation of the phenomenon in question at the expense of concern with the proportion of variance explained and its partitioning among variables. Description of the structure is needed as a guide to the most appropriate measures and proper inferences from the results.

As illustration of the utility of giving explicit attention to the structural model underlying a given problem, figures I and II present two possible configurations of the variables in equation (1).

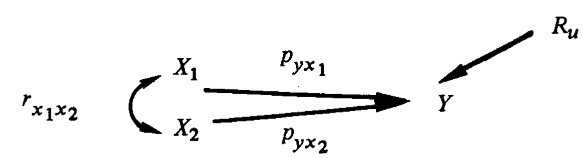


Figure I. Structural model without intermediate variable

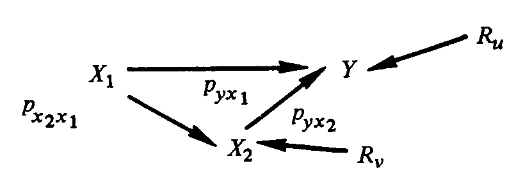


Figure II. Structural model with one intermediate variable

The representation uses the conventions of path analysis, a multivariate technique useful in explicating linear causal models.[62]

---

[61] Otis Dudley Duncan, "Partials, partitions, and paths", in Edgar F. Borgatta and G. Bohrnstedt, eds., *Sociological Methodology, 1970* (San Francisco, Jossey-Bass, 1970), pp. 38–47.

[62] For detailed expositions of these techniques, see S. Wright,

The variables in the figures are arranged in the presumed temporal and causal sequence. Each variable occurs earlier in time than those appearing to the right of it and later than those to the left, and can be affected by all the variables that precede it. The direct influences from one variable to another are represented by the one-way straight arrows and the sign and magnitude of these are the path coefficients, P_{ij} (read as representing the direct effect on variable i from variable j). The curved bidirectional arrow represents the correlation between the exogenous variables not analysed within the model, and the value shown is the zero order correlation coefficient, r.

In the algebraic representation of the model, each variable is in standard form and each dependent variable is treated as completely determined by some set of other variables in the model, including as necessary the residual effects of unmeasured variables, represented by the R_i in the diagrams. Thus, for figure I there would be only one equation:

$$Y = p_{yx_1} X_1 + p_{yx_2} X_2 + p_{yu} R_u \qquad (2)$$

For figure II, there would be two equations, one for Y exactly as presented in equation (2) and one for X_2:

$$X_2 = p_{x_2x_1} X_1 + p_{x_2v} Rv \qquad (3)$$

To understand the nature of the interrelationships among the variables and the difficulty of partitioning R^2, it is instructive to solve for the path coefficients in terms of the observed correlations among the variables.[63] Since the variables are in standard form:

$$r_{x2y} = \frac{\Sigma X_2 Y}{N}$$

and substituting for Y in equation (2) yields:

$$r_{x2y} = \frac{1}{N} \Sigma X_2 (p_{yx_1} X_1 + p_{yx_2} X_2 + p_{yu} R_u)$$

$$r_{x2y} = p_{yx_1} r_{x_1x_2} + p_{yx_2} \qquad (4)$$

since $\frac{\Sigma x_1 x_1}{N} = 1$, and $r_{x_2} R_u = 0$, on the assumption that the residual effects are uncorrelated with the exogenous variables.

The same process applied to r_{x_1y} yields:

$$r_{x_1y} = p_{yx_1} + p_{yx_2} r_{x_1x_2} \qquad (5)$$

For figure I, since only equation (2) is involved, one is able to see from equations (4) and (5) how the model generates the observed correlations. Equation (4) shows that the correlation between Y and X_2 is made up of the direct effect of X_2 on Y given by p_{yx_2} and by a joint effect it shares with X_1, that is, an effect arising from the correlation of X_2 with another cause of Y, namely, X_1. This effect is represented by the first term on the right side of equation (4). Similarly, the observed correlation between X_1 and Y can be seen to be the sum of a direct effect of X_1 on Y plus a joint effect of the type just described.

The proportion of variance in Y accounted for by X_1 and X_2 can be expressed in terms of the same notation as:

$$R^2 = p^2_{yx_1} + p^2_{yx_2} + 2 p_{yx_1} p_{yx_2} r_{12}$$

From this, it can be seen that there are three components of the explained variance: those due to X_1, to X_2, and to X_1 and X_2 jointly. Thus, a decomposition of the explained variance into a portion attributable to X_1 and a portion attributable to X_2 is not possible.[64] (It is also true that:

$$R^2 = p_{yx_1} r_{yx_1} + p_{yx_2} r_{yx_2}$$

but this does not provide the desired decomposition as either of the terms can be negative and it has been shown in equations (4) and (5) that the two correlations include joint effects.)

For figure II, there is the additional relation represented by equation (3). The same algebra as used in the preceding paragraph provides for this equation the result that the path coefficient is simply the zero-order correlation:

$$r_{x_1x_2} = p_{x_2x_1} \qquad (6)$$

Substituting for $r_{x_1x_2}$ in equations (4) and (5) gives:

$$r_{x_2y} = p_{yx_1} p_{x_2x_1} + p_{yx_2} \qquad (7)$$

$$r_{x_1y} = p_{yx_1} + p_{yx_2} p_{x_2x_1} \qquad (8)$$

Though no values have changed, the interpretation of the situation is different. Looking first at equation (8) along with figure II, the correlation between X_1 and Y is seen to be due to the direct effect p_{yx_1}, and the indirect effect by means of X_2 represented by the second term on the right. Equation (7) and figure II reveal that the correlation between X_2 and Y consists of the direct effect p_{yx_2} and a component of correlation due

"The method of path coefficients", *Annals of Mathematical Statistics*, vol. 5, No. 3 (September 1934), pp. 165–215; Otis Dudley Duncan, "Path analysis: sociological examples", *American Journal of Sociology*, vol. 72, No. 1 (July 1966), pp. 1–16; D. Heise, "Problems in path analysis and causal inference", in Edgar F. Borgatta, ed., *Sociological Methodology, 1969* (San Francisco, Jossey-Bass, 1969), pp. 38–73. An application in terms of an areal multivariate analysis of family planning programme effects is given in Albert I. Hermalin, "Regression analysis of areal data", in C. Chandrasekaran and Albert I. Hermalin, eds., *Measuring the Effect of Family Planning Programs on Fertility* (Liège, International Union for the Scientific Study of Population for the Development Centre of the Organisation for Economic Co-operation and Development, 1975), pp. 245–300.

[63] Under the assumption used here that the exogenous variables in each equation are uncorrelated with the residual term, the path coefficients are equal to the standardized partial regression coefficients, the beta coefficients, obtained from the ordinary least-squares solution of each equation.

[64] O. D. Duncan, "Partials, partitions, and paths".

to the fact that X_2 and Y are both affected by common cause, X_1.

Thus, a given set of observed correlations which yields a particular solution for the partial regression coefficients of equation (1) can be interpreted as arising, in one case, from direct and joint effects; or, in the other, from direct, indirect and effects due to a common cause, depending upon the system of interrelationships assumed to hold among the variables.[65]

Figure II also allows some additional perspective on partitioning the explained variance, though it does not permit of a unique decomposition. It can be shown[66] that:

$$R^2 = r^2_{yx_2} + p^2_{yx_1}(1 - r^2_{x_1x_2}) \quad \text{or} \quad (9)$$

$$R^2 = r^2_{yx_1} + p^2_{yx_2}(1 - r^2_{x_1x_2}) \quad (10)$$

Equation (9) expresses R^2 as the total effect due to the programme variable (X_2) including that transmitted from the non-programme variable (X_1), plus an increment due to including X_1. Equation (10) expresses R^2 as the total effect arising from the most remote cause X_1, and an increment gained by including the intervening variable X_2.

In the foregoing sections, the problems of correlation between independent variables have been treated within the context of a linear additive model represented by equation (1). In such a situation, it is possible to obtain estimates of the direct effect of each variable, represented essentially by the partial regression coefficients, even though it is not possible to partition the explained variance and the interpretation of the interrelationships will depend upon the underlying model posited, as illustrated by figures I and II.

In some situations, however, there may be so-called "interaction effects", whereby the effect of one variable depends on the level of a second variable. This might be the case, for example, where the effect of a family planning programme is different across countries, or in different areas within a country, depending upon the level of socio-economic development. This situation may be interpreted as an example of a particular effect arising from the presence of both explanatory variables that is not simply the sum of their independent effects. Thus, an additive model no longer is applicable. Often a regression analysis of such a model will employ an equation of the following type:

$$Y = a + b_1 X_1 + b_2 X_2 + cX_1 X_2 + e \quad (11)$$

This equation assumes that there are independent effects plus an interaction effect captured by the simple multiplicative term. In some cases, the interaction effect will be better represented by other than a simple cross-product. Whether a model of the type represented by equation (11) is to be preferred to that of equation (1) can be tested by noting whether the additional variance explained is significant. Where interaction is present, the ability to partition effects in the sense of equation (1) is no longer possible.

The problems treated in this section are, of course, directly relevant to the application of regression analysis to the study of family planning effects. They also enter into other evaluation techniques. For example, in the decomposition or standardization approach, if one attempts to decompose a change in crude birth rate into components on the assumption that the factors remained constant as of the initial date, one or more components representing correlation of change across factors may be needed to account fully for the observed change in crude birth rate.[67] More generally, in the standardization approach, the question of the independence of factors, often assumed, should be explicitly considered.

The logic of this section should also prove helpful in analysing key concepts in evaluation, such as the way socio-economic variables and the family planning programme influence motivation to accept, or the possible influence of the programme on adoption of private means of contraception. This section may also be relevant in analyses that seek to determine the nature of and interrelationships among factors associated with such important parameters as the probability of secondary sterility or the length of post-partum amenorhoea.

Uncontrolled variables

The problem of uncontrolled variables arises when one or several factors that are important in the understanding of variations in a variable under study have not been taken into consideration. The reasons for such short-comings vary. A major reason is, of course, that social phenomena are complex and often not well understood, so that one cannot identify all relevant determinants or the relevancy of certain determinants is imperfectly known. In other cases, it may not be possible to apply a particular evaluation method to as many variables as would be desired or to certain types of factors (for example, qualitative variables); or the method may not deal satisfactorily with certain categories of variables (such as biological factors). Lastly, variables may be excluded simply because the required data are not available. The omission of one or more variables from a model is referred to as "specification error". This term also refers to the situation, illustrated in the previous sections, of an incorrect

[65] Other possible arrangements of the three-variable case and analysis of four-variable systems are given in O. D. Duncan, "Partials, partitions, and paths"; and *idem, Introduction to Structural Equation Models* (New York, Academic Press, 1975).

[66] O. D. Duncan, "Partials, partitions, and paths", pp. 41–42.

[67] Evelyn M. Kitagawa, "Components of a difference between two rates", *Journal of the American Statistical Association*, vol. 50, No. 272 (December 1955), pp. 1178–1179.

temporal or causal arrangement of the variables included in the model or of an incorrect functional form.

Conceptually akin to specification error is the problem of errors in measurement and of unobserved variables. If a variable is incorrectly measured, then in a sense that variable is not properly controlled in the model and the desired variable is unobserved. The term "unobserved variables" is also used for the situation where the underlying concept is more highly abstract and measurements relevant to the concept are regarded as indicators rather than operational definitions.[68]

A simple example of the effect of omitting a variable can be illustrated with reference to the equations in the preceding section. Assume that in equation (1) the non-programme variable, X_1, is omitted and one is interested in the possible effects of this omission on the estimated effect of the programme variable X_2. If X_2 is uncorrelated with X_1, then the regression coefficient for X_2 is unaffected by the omission of X_1. On the other hand, if the two variables are correlated then the estimate of the effect of X_2 is biased and its regression coefficient may be larger or smaller when X_1 is included in the equation. The strategy suggested for studying the effect of omitted variables is to counterpose against the model being used a competing "true" model and to trace out the differences between the models and their consequences.[69]

Various aspects of measurement error have been discussed above and therefore are not further pursued here. The point worth noting in the light of the discussion in this and the previous section is that no study of the effects of measurement error is possible without assumptions about their nature. In sum, models of error should be incorporated into the substantive models in order to improve the inferential process.

One strategy for dealing with abstract concepts which do not lead to a single operational definition is to regard various observations as indicators or manifestations of the unobservable variable. Under certain conditions, developments in causal modelling permit inferences about the unobservable variable to be made on the basis of the behaviour of their indicators.[70] These developments appear relevant to problems in family planning evaluation in that certain key concepts, such as modernization or programme inputs, might be regarded as unobservables for which there are multiple indicators.

Independence of method

The multiplicity of evaluation techniques arises in part from the complexity of the relationships involved in analysing fertility and in part from the type, amount and quality of data available for the purpose of evaluation. Some techniques focus on certain dimensions of the problem and certain categories of data while other techniques have different foci. Since the country studies will reveal the results of applying different techniques to the same setting, it is important to consider the possible outcomes of multiple application.

There are a large number of combinations of possible outcomes from the application of two or more techniques, and no attempt is made here to enumerate them all. For example, if two techniques differ in their implications of programme effect, both may be tenable because they differ in coverage or the dimension of the problem being analysed. Or it may turn out that the data associated with one method are much more unreliable than that of another. A third possibility, of course, is that the results differ because of differences in basic assumptions, and it then remains to determine whether there is any evidence to support the plausibility of one set of assumptions as against the other.

If two techniques agree in their implication of programme effect, by how much are the conclusions strengthened as against having just one of the results? The question may be conceptualized as one of independence of methods. If the methods overlap considerably so that the results obtained by one technique are largely constrained to parallel those from the other, then relatively little is gained. On the other hand, if the methods were viewed as largely independent, one would interpret agreement of results from each method as strengthening the particular conclusion. The question thus resolves itself into determining the degree of independence among the several evaluation techniques.

Intuitively, two methods may be viewed as independent if they utilize different frames of reference in assessing programme impact. What is needed, however, is a set of rather specific criteria which can be applied uniformly by different investigators to assess the relative distinctness of each method. This is clearly a complex question which is likely to engender considerable discussion in the review of the country studies. The criteria listed below are designed to facilitate such discussion without attempting to reach any conclusion on this important issue:

(*a*) Number and type of assumptions utilized by each technique: how independence of assumptions can be analysed;

(*b*) Type of factors utilized: demographic *versus* biological etc.;

(*c*) Type of estimating technique employed: standardization *versus* regression *versus* projection etc.;

(*d*) Direct *versus* indirect measurement of programme effect: some techniques assess programme

---

[68] H. M. Blalock, Jr. "Making causal inferences for unmeasured variables from correlations among indicators", *American Journal of Sociology,* vol. LXIX, No. 1 (July 1963), pp. 53–62.

[69] The application of this strategy to more complex models is given in O. D. Duncan, *Introduction to Structural Equation Models,* pp. 101 ff.

[70] Philip M. Hauser and A. S. Goldberger, "The treatment of unobservable variables in path analysis", in Herbert L. Costner, ed., *Sociological Methodology, 1971* (San Francisco, Jossey-Bass, 1971), pp. 81–117; and O. D. Duncan, *Introduction to Structural Equation Models,* pp. 129 ff.

effect as a residual, while others obtain an estimate of effect directly from programme factors;

(*e*) Coverage: whether a method that focuses on over-all fertility change is independent of a method that utilizes only data on acceptors.

The mere enumeration of several possible criteria illustrates the difficulty of the problem. Within any one criterion there must be further specification of how differences across techniques is to be analysed. Following this, consideration must be given to the degree of overlap across criteria. Are all the criteria listed, and possibly others, required to assess independence of methods or would a small subset suffice?

Cost-precision analysis

The problem of cost-precision analysis is not specifically a methodological issue. The selection of one or several programme evaluation methods constitutes the first step in family planning programme evaluation. In light of the cost and objectives of programme evaluation and the limited resources available for the exercise, evaluators should determine which approach would yield the highest return in terms of quality of results. As quality of results is highly dependent upon the precision obtained, the problem is one of balance between method precision and cost.

In dealing with this problem, the first assumption is that the precision of the method is known, or can be determined or assessed. The second assumption is that some criteria can be worked out to assess the precision required. Depending upon the purpose of the evaluation, greater or less precision may be needed. The third assumption considers that the cost of an evaluation undertaking can be assessed fairly well. Another assumption is that a methodology exists or can be worked out to undertake the cost-precision trade-off that would give the evaluator a means of selecting the evaluation method or methods that would yield the greatest precision for a given amount of expenditures. These four assumptions are examined below.[71]

Precision of method

The meaning of precision is to be understood here as the magnitude of the errors that may accompany measurements of programme impact on fertility. Method precision depends chiefly upon the data-gathering process and the estimating techniques involved. In some cases, the concept of precision is well established. In probability sampling, for example, precision means the difference beteen the sample results and the results that would have been obtained from a complete enumeration under similar over-all conditions. This means that biases present in both sample and total enumerations would not be revealed by standard statistical techniques for measuring error.

The field of sampling error has been widely investigated, and the determination and assessment of sampling error can be undertaken in varying circumstances. Errors may be reduced by manipulating sampling design, although it is not worth while to engage in complex sampling techniques when non-sampling errors are suspected to be large. The contribution of standard confidence intervals to appraisal of precision needs no comment, and both the standard deviation and the mean square error are also indicators that can be utilized in assessing method precision. In applying evaluation methods, loss of precision can occur at various stages. A major source of error is, of course, that which occurs at the data-gathering stage. But errors and imprecise results can also occur as a consequence of erroneous application of the method (e.g., assumptions not met), of model simplifications (e.g., omission of important variables), use of unsatisfactory indicators (e.g., unstandardized crude birth rates) etc.

The problem of method precision is, therefore, not only inherent in the method itself but inherent in its proper application. In order to assess method precision, one should assume that the method is properly applied. A discussion on method precision might also treat the problems of random errors and non-random errors separately, except when their combined effect can be examined in terms of mean square error. In some cases, direct measurement of precision cannot be accomplished, but experience can provide information on the existence and the direction of certain types of errors.

Due to the novelty of the field of family planning programme evaluation, there have not been thorough studies of over-all precision of the evaluation methods.

Precision required

The precision required from a particular evaluation depends chiefly upon the objectives of the exercise. If only a rough approximation is needed for, say, a specific administrative purpose, a method that would give only a trend direction or an order of magnitude of the change might suffice. If, on the other hand, the evaluation is directed to obtaining a precise measurement of fertility change, to assess the role of a new programme component (a new contraceptive, for instance), a different method with a more specific and precise outcome would be needed. The precision required is thus a function of both the known precision of the various methods and the objectives of the evaluation. It is the judgement of the administrators of the programme and the evaluators that will decide this question, which implies, of course, that there is some criterion for determining an "acceptable" margin of error with respect to a given evaluation objective. For instance, the standard deviation of a sample estimate could be selected so that the sample estimate would permit detection, with relative certainty, of a 3 per cent change or more.

[71] It is also often implicitly assumed that the more precise the results required, the more expensive the evaluation undertaking. This assumption, not discussed here, may often be misleading.

Cost of evaluation

Determining the cost of evaluation is mainly an accounting problem. Evaluation expenditures will, of course, depend upon the type of data-gathering procedure, the data available and data to be collected, geographical circumstances of the area under study, qualification of evaluation personnel, equipment etc. For methods using sample data, once the amount of precision is chosen and the associated probability decided, a sample design can be worked out and its cost assessed. To this cost must be added the additional expenditures for data processing, analyses etc. In brief, information on all costs associated with the application of a particular method must be available or hypothesized.

Cost-precision trade-off

If the cost of applying a given method is acceptable, the evaluation can, of course, be undertaken as planned. If the cost is not acceptable, some balancing of precision against cost must be considered. In this respect, it appears that the basic questions to answer are:

(*a*) What are the advantages of obtaining the precision required?

(*b*) What are the disadvantages of settling for less precision?

(*c*) What is the differential cost of settling for the higher or the lesser degree of precision?

The most important item in considering these questions appears to be the purpose of the evaluation. If the results of the evaluation cannot be utilized as planned because of lack of precision, then the evaluation should not be undertaken unless more funds can be made available. For instance, if programme administrators wish to identify, say, a 5 per cent change in fertility over a given period of time, they might need a very large sample, which might be a very expensive undertaking. Depending upon the purpose of the evaluation study, administrators may consider that some relaxation in precision would not invalidate the results, or that a reduced sample which would identify only an 8 per cent fertility change would still provide useful information. Such decisions can be made in light of the basic questions listed above and can be based on intuitive judgement of the known evaluation conditions.

Selected bibliography

Standardization approach

Anderson, John E. The relationship between change in educational attainment and fertility rates. *Studies in family planning* (New York) 6:72–81, March 1975.

Cho, Lee-Jay *and* Robert D. Retherford. Comparative analysis of recent fertility trends in East Asia. *In* International Population Conference, Liège, 1973. Liège, International Union for the Scientific Study of Population, 1974. v. 2. p. 163–181.

Freedman, Ronald. A comment on "Social and economic factors in Hong Kong's fertility decline" by Sui-ying Wat and R. W. Hodge. *Population studies* (London) 27:589–595, November 1973.

Freedman, Ronald *and* Arjun L. Adlakha. Recent fertility declines in Hong Kong; the role of the changing age structure. *Population studies* (London) 22:181–198, July 1968.

Hong Kong's fertility decline 1961–68, *By* R. Freedman *and others*. *Population index* (Princeton, New Jersey) 36:3–18, January-March 1970.

Kitagawa, Evelyn M. Components of a difference between two rates. *Journal of the American Statistical Association* (Washington) 50:1168–1194, December 1955, no. 272.

Lapham, Robert J. Family planning and fertility in Tunisia. *Demography* (Washington) 7:241–253, May 1970.

Reynolds, Jack. Costa Rica; measuring the demographic impact of family planning programs. *Studies in family planning* (New York) 4:310–316, November 1973.

Vallin, Jacques. Limitation des naissances en Tunisie; efforts et résultats. *Population* (Paris) 26:181–204, special issue, March 1971.

Wat, Sui-ying and R. W. Hodge. Social and economic factors in Hong Kong's fertility decline. *Population studies* (London) 26:455–464, November 1972.

Trend analysis

Bogue, Donald J. Family planning improvement through evaluation; a manual of basic principles. Chicago, University of Chicago, 1970. 82 p. (Community and Family Study Center. Family Planning Research and Evaluation Manual, 1)

Mauldin, W. Parker. Births averted by family planning programs. *Studies in family planning* (New York) 1:2–3, August 1968, no. 33.

Sivin, Irving. Fertility decline and contraceptive use in the international post-partum family planning program. *Studies in family planning* (New York) 2:248–256, December 1971.

Wolfers, David. An evaluation criterion for a national family planning program. *American journal of public health* (Washington) 58:1447–1451, August 1968.

Analysis of reproductive process

Potter, Robert G. Application of life-table techniques to measurement of contraceptive effectiveness. *Demography* (Washington) 3:297–304, 1966, no. 2.

———. A technical appendix on procedures used in manuscript "Estimating births averted in a family planning program". Paper prepared for Major Ceremony V, University of Michigan Sesquicentennial Celebration, 1 June 1967.

———. Estimating births averted in a family planning program. *In* Fertility and family planning; a world view. S. J. Behrman, Leslie Corsa, Jr. *and* Ronald Freedman, *eds*. Ann Arbor, University of Michigan Press, 1969. p. 413–434.

Potter, Robert G. *and* Roger C. Avery. Use-effectiveness of contraception. *In* Measuring the Effect of family planning programs on fertility. C. Chandrasekaran *and* Albert I. Hermalin, *eds*. Liège, International Union for the Scientific Study of Population for Development Centre of Organization for Economic Co-operation and Development, 1975. p. 133–162.

Tietze, Christopher. Intra-uterine contraception; recommended procedures for data analysis. *Studies in family planning* (New York) 1:1–6, supplement, April 1967, no. 18.

Tietze, Christopher *and* Sarah Lewit. Recommended procedures for the statistical evaluation of intra-uterine contraception. *Studies in family planning* (New York) 4:35–42, February 1973.

Wolfers, David. The demographic effect of a contraceptive programme. *Population studies* (London) 23:111–141, March 1969.

———. Births averted. *In* Measuring the effect of family planning programs on fertility. C. Chandrasekaran *and* Albert I. Hermalin, *eds*. Liège, International Union for the Scientific Study of Population for Development Centre of Organization for Economic Co-operation and Development, 1975. p. 163–214.

Regression analysis

Duncan, Otis Dudley. Path analysis; sociological examples. *American journal of sociology* (Chicago) 72:1–16, July 1966.

Freedman, Ronald. A comment on "Social and economic factors in Hong Kong's fertility decline" by Sui-ying Wat and R. W. Hodge. *Population studies* (London) 27:589–595, November 1973.

Hermalin, Albert I. Regression analysis of areal data. *In* Measuring the effect of family planning programs on fertility. C. Chandrasekaran *and* Albert I. Hermalin, *eds*. Liège, International Union for the Scientific Study of Population for Development Centre of Organization for Economic Co-operation and Development, 1975. p. 245–300.

Johnston, J. Econometric methods. New York, McGraw-Hill, 1963.

Schultz, T. Paul. The effectiveness of population policies; alternative methods of statistical inference. Santa Monica, California, The Rand Corporation, 1971.

Wat, Sui-ying *and* R. W. Hodge, Social and economic factors in Hong Kong's fertility decline. *Population studies* (London) 26:455–464, November 1972.

Experimental designs

Bang, S. Assessment of the demographic impact of a five-year fertility control program in rural Korea. *In* International Population Conference, London, 1969. Liège, International Union for the Scientific Study of Population, 1971. v.2. p. 1085–1090.

Johnson, J. T., Tan Boon Ann *and* Leslie Corsa. Assessment of family planning programme effects on births; preliminary results obtained through direct matching of birth and programme acceptor records. *Population studies* (London) 27:85–96, March 1973.

Okada, L. M. The use of matched pairs in the evaluation of the District of Columbia, Department of Public Health birth control program. *In* Proceedings of the Social Statistics Section, 1967. Washington, American Statistical Association, 1968. p. 206–211.

———. Use of matched pairs in evaluation of a birth control program. *Public health reports* (Washington) 84:445–450, May 1969.

Population Council. India; the Singur study. *Studies in family planning* (New York) 1:1–4, July 1963, no. 1.

———. Korea; the Koyang study. *Studies in family planning* (New York) 1:7–9, December 1963, no. 2.

Wells, H. Bradley. Matching studies. *In* Measuring the effect of family planning programs on fertility. C. Chandrasekaran *and* Albert I. Hermalin, *eds*. Liège, International Union for the Scientific Study of Population for Development Centre of Organization for Economic Co-operation and Development, 1975. p. 215–244.

Yang, Jae Mo. Fertility and family planning in rural Korea. *In* Proceedings of the World Population Conference, Belgrade, 30 August–10 September 1965. v. 2. Selected papers and summaries; fertility, family planning, mortality. p. 309–312.
Sales No.: 66.XIII.6.

Couple-years of protection

Adil, Enver. Measurement of family planning progress in Pakistan. *Demography* (Washington) 5:659–655, 1968, no. 2.

Bean, Lee L. *and* William Seltzer. Couple-years of protection and births prevented; a methodological examination. *Demography* (Washington) 5:947–959, 1968, no. 2.

Mauldin, W. Parker. Births averted by family planning programs. *Studies in family planning*. (New York) 1:1–7, August 1968, no. 33.

Siddiqui, K. A. Personnel performance and reporting system. *In* Seminar on Family Planning. Karachi, Regional Co-operation for Development, 6–8 April 1966.

United Nations. Economic Commission for Asia and the Far East. Assessment of acceptance and effectiveness of family planning methods. Report of an expert group meeting, Bangkok, 11–21 June 1968. (Asian Population Studies Series, 4)
Sales No. E.69.II.F.15.

Wishik, Samuel M. Indexes for measurement of amount of contraceptive practice. Paper presented at the Seminar on Evaluation of Family Planning Programmes, Bangkok, 24 November–12 December 1969. (POP/ESFP/10)

Wishik, Samuel M. *and* K. H. Chen. The couple-year of protection; a measure of family planning program output. New York, Columbia University, International Institute for the Study of Human Reproduction, 1973. (Manuals for Evaluation of Family Planning and Population Programs, 7)

Zafar, S. Acceptance level for various contraceptive methods. *In* Seminar on Family Planning. Karachi, Regional Co-operation for Development, 6–8 April 1966.

Component projection approach

A study on the effectiveness of sterilizations in reducing the birth rate. *By* M. Alfred Haynes *and others*. *Demography* (Washington) 6:1–11, February 1969.

Lee, B. M. *and* John Isbister. The impact of birth control programs on fertility. *In* Family planning and population programs; a review of world development. Bernard Berelson *and others, eds.* Chicago, University of Chicago Press, 1966. p. 737–758.

Mauldin, W. Parker. Births averted by family planning programs. *Studies in family planning* (New York) 1:1–7, August 1968, no. 33.

Miller, Peter C., L. T. Lillian *and* R. J. Lapham. Fertility reduction in an MCH/family planning program; a model for projection. *Studies in family planning* (New York) 6:2–16, January 1975.

Nortman, Dorothy. Births averted by the post-partum program; a methodology and some estimates and projections. *In* Post-partum family planning; a report on the international program. G. I. Zatuchni, *ed.* New York, McGraw-Hill, 1970. p. 133–166.

———. Hypothetical illustration of demographic impact of a post-partum program. Appendix B. *In* Howard C. Taylor, Jr. *and* Bernard Berelson. Comprehensive family planning based on maternal-child health services; a feasibility study for a world program. *Studies in family planning* (New York) 2:50–54, February 1971.

United Nations. Economic and Social Commission for Asia and the Pacific. Some techniques for measuring the impact of contraception; an aid to target setting. Bangkok, 1974. (Asian Population Studies Series, 18, E/CN.11/1119)

Vallin, Jacques. Planning familial et perspective de population en Tunisie, 1966–1975. *Revue tunisienne de sciences sociales* (Tunis) 5:71–88, January 1968.

Venkatacharya, K. A model to estimate births averted due to IUCDs and sterilizations. *Demography* (Washington) 8:491–505, November 1971.

Simulation models

Barrett, J. C. Use of a fertility simulation model to refine measurement techniques. *Demography* (Washington) 8:481–490, November 1971.

———. A Monte Carlo simulation of reproduction. *In* Biological aspects of demography. William Brass, *ed.* London, Taylor and Francis, 1971. p. 11–31.

Clague, Alice S. *and* Jeanne C. Ridley. The assessment of three methods of estimating births averted. *In* Computer simulation in human population studies. Bennett Dyke *and* Jean W. MacCluer, *eds.* New York, Academic Press, 1973. p. 329–382.

The evaluation of four alternative family planning programs for POPLAND, a less developed country. *By* A. V. Rao *and others*. *In* Computer simulation in human population studies. Bennett Dyke *and* Jean W. MacCluer, *eds.* New York, Academic Press, 1973. p. 261–304.

Holmberg, Ingvar. Fecundity, fertility and family planning. I. Application of demographic micromodels. Göteborg, University of Göteborg, 1970. (Demographic Institute reports, 10)

Horvitz, D. G. POPSIM, a demographic simulation model. *In* International Population Conference, London, 1969. Liège, International Union for the Scientific Study of Population, 1971. v. I. p. 95–106.

Hyrenius, Hannes *and* I. Adolfsson. A fertility simulation model. Göteborg, University of Göteborg, 1964. (Demographic Institute reports, 2)

Jacquard, Albert. La reproduction humaine en régime malthusien. *Population* (Paris) 22:897–920, septembre-octobre 1967.

Lachenbruch, P. A., M. C. Sheps *and* A. M. Sorant. Applications of POPREP, a modification of POPSIM. *In* Computer simulation in human population studies. Bennett Dyke *and* Jean W. MacCluer, *eds.* New York, Academic Press, 1973. p. 305–328.

Menken, Jane. Simulation studies. *In* Measuring the effect of family planning programs on fertility. C. Chandrasekaran *and* Albert I. Hermalin, *eds.* Liège, International Union for the Scientific Study of Population for Development Centre of Organization for Economic Co-operation and Development, 1975. p. 351–380.

On the apparent subfecundity of non-family planners. By Jeanne Clare Ridley and others. *Social biology* (New York) 16:24–28, March 1969.

Perrin, Edward B. *and* Mindel C. Sheps. Human reproduction; a stochastic process. *Biometrics* (Raleigh, North Carolina.) 20:28–45, March 1964.

Potter, Robert G. Births averted by contraception; an approach through renewal theory. *Theoretical population biology* (New York) 1:251–272, November 1970.

———. Renewal theory and births averted. *In* International Population Conference, London, 1969. Liège, International Union for the Scientific Study of Population, 1971. v. 1. p. 145–150.

———. Description of ACCOFERT II. Providence, R.I., and Ann Arbor, Mich.; Brown University and University of Michigan, April 1971.
Mimeographed.

———. Births averted by induced abortion; an application of renewal theory. *Theoretical population biology* (New York) 3:62–86, March 1972.

———. Additional births averted when abortion is added to contraception. *Studies in family planning* (New York) 3:53–59, April 1972.

Potter, Robert G. *and* James M. Sakoda. A computer model of family building based on expected values. *Demography* (Washington) 3:450–461, 1966, no. 2.

Ridley, Jeanne C. *and* Mindel C. Sheps. An analytic simulation model of human reproduction with demographic and biological components. *Population studies* (London) 19:297–310, March 1966.

Sheps, Mindel C. Application of probability models to the study of patterns of human reproduction. *In* Public health and population change; current research issues. Mindel C. Sheps *and* Jeanne C. Ridley, *eds*. Pittsburgh, University of Pennsylvania Press, 1965. p. 307–332.

———. Contribution of natality models to program planning and evaluation. *Demography* (Washington) 3:445–449, 1966, no. 2.

Sheps, Mindel C. *and* Edward B. Perrin. Changes in birth rates as a function of contraceptive effectiveness; some applications of a stochastic model. *American journal of public health* (Washington) 53:1031–1046, July 1963.

Sheps, Mindel C., Jane A. Menken *and* Annette P. Radick, Probability models for family building; an analytical review. *Demography* (Washington) 6:161–183, May 1969.

Tietze, Christopher *and* John P. Bongaarts. Fertility rates and abortion rates: simulations of family limitations. *Studies in family planning* (New York) 6:114–120, May 1975.

Venkatacharya, K. Some implications of susceptibility and its application in fertility evaluation models. *Sankhya* (Calcutta) 32:41–54, series B, June 1970, no. 1–2.

———. Reduction in fertility due to induced abortions; a simulation model. *Demography* (Washington) 9:339–352, August 1972.

POTENTIAL FERTILITY

Brass, William. Assessing the demographic effect of a family planning programme. *Proceedings of the Royal Society of Medicine* (London), 63:29–31, November 1970.

Chandrasekaran, C., D. V. R. Murty *and* K. Srinivasan. Some problems in determining the number of acceptors needed in a family planning programme to achieve a specified reduction in the birth rate. *Population studies* (London) 25:303–308, July 1971.

Davis, Kingsley *and* Judith Blake. Social structure and fertility; an analytical framework. *Economic development and cultural change* (Chicago) 4:211–235, April 1956.

Henry, Louis. La fecondité naturelle; observations—théorie—résultats. *Population* (Paris) 16:625–636, octobre-décembre 1964.

Lee, B. M. *and* John Isbister. The impact of birth control programs on fertility. *In* Family planning and population programs; a review of world development. Bernard Berelson *and others, eds*. Chicago, University of Chicago Press, 1965. p. 737–758.

Mauldin, W. Parker. Births averted by family planning programs. *Studies in family planning* (New York) 1:1–7, August 1968, no. 33.

On the apparent subfecundity of non-family planners. *By* Jeanne Clare Ridley *and others*. *Social biology* (Chicago, Ill.) 16:24–28, March 1969.

Perrin, Edward B. *and* Mindel C. Sheps. Human reproduction; a stochastic process. *Biometrics* (Raleigh, North Carolina) 20:28–45, March 1964.

Potter, Robert G. A technical appendix on procedures used in manuscript "Estimating births averted in a family planning program". Prepared for Major Ceremony V, University of Michigan Sesquicentennial Celebration, 1 June 1967.

———. Estimating births averted in a family planning program. *In* Fertility and family planning; a world view. S. J. Behrman, Leslie Corsa, Jr. *and* Ronald Freedman, *eds*. Ann Arbor, University of Michigan Press, 1969. p. 413–434.

Presser, Harriet B. The role of sterilization in controlling Puerto Rican fertility. *Population studies* (London) 23:343–361, November 1969, table 9.

Republic of Korea. Ministry of Health and Social Affairs. National intra-uterine contraception report. Seoul, Planned Parenthood Federation of Korea, 1967.

Ross, John A. Cost of family planning programs. *In* Family planning and population programs. Bernard Berelson *and others, eds*. Chicago, University of Chicago Press, 1966. p. 759–778.

United Nations. Economic Commission for Asia and the Far East. Assessment of acceptance and effectiveness of family planning methods. Report of an expert group meeting, Bangkok, 11–21 June 1968. (Asian Population Studies Series, 4)
Sales No. E.69.II.F.15.

Wolfers, David. Determinants of birth intervals and their means. *Population studies* (London) 22:253–262, July 1968.

———. The demographic effects of a contraceptive programme. *Population studies* (London) 23:111–140, March 1969.

———. The estimation of potential fertility for family planning evaluation. *Proceedings of the Royal Society of Medicine* (London), 63:41–44, 1970, no. 11.

———. Some problems in calculating births averted. *In* International Population Conference, Liège, 1973. Liège, International Union for the Scientific Study of Population, 1974. v. 2. p. 233–245.

———. Births averted. *In* Measuring the effect of family planning programs on fertility. C. Chandrasekaran *and* Albert I. Hermalin, *eds*. Liège, International Union for the Scientific Study of Population for Development Centre of Organization for Economic Co-operation and Development, 1975. p. 163–214.

DATA REQUIREMENT PROBLEMS

Barrett, J. C. *and* W. Brass. Systematic and chance components in fertility measurement. *Population studies* (London) 28:473–493, November 1974.

Blalock, Hubert M. Multiple indicators and the causal approach to measurement error. *American journal of sociology* (Chicago) 75:264–272, September 1969.

———. Estimating measurement errors using multiple indicators and several points in time. *American sociological review* (Albany, New York) 35:101–112, February 1970.

Caldwell, J. C. *and* A. A. Igun. An experiment with census-type age enumeration in Nigeria. *Population studies* (London) 25:287–302, July 1971.

Chai, John J. Correlated measurement errors and the least squares estimator of the regression coefficient. *Journal of the American Statistical Association* (Washington) 66:478–483, 1971, no. 335.

Chandrasekaran, C., D. V. R. Murty *and* K. Srinivasan. Some problems in determining the number of acceptors needed in a family planning programme to achieve a specified reduction in the birth rate. *Population studies* (London) 25:303–308, July 1971.

Chidembaram, V. C. Use of KAP survey data to ascertain changes in fertility in areas where it is high. Paper presented at the Technical Meetings on Methods of Analysing Fertility Data for Developing Countries, Budapest, 14–25 June 1971. (E/CN.9/AC.12/R.10)

Coale, Ansley J. The design of an experimental procedure for obtaining accurate vital statistics. *In* Proceedings of the International Population Conference, New York, 1961. London, International Union for the Scientific Study of Population, 1963. v. 2. p. 372–375.

Cochran, W. G. Sampling techniques. New York, Wiley, 1961. 330 p.

———. Some effects of errors of measurement of multiple correlation. *Journal of the American Statistical Association* (Washington) 65:22–34, 1970, no. 329.

El-Badry, M. A. Failure of enumerators to make entries of zero; errors in recording childless cases in population censuses. *Journal of the American Statistical Association* (Washington) 56:909–924, December 1961, no. 296.

Fox, Karl A. Intermediate economic statistics. New York, Wiley, 1968. 568 p.

Freedman, Ronald. A comment on "Social and economic factors in Hong Kong's fertility decline" by Sui-ying Wat and R. W. Hodge. *Population studies* (London) 27:589–595, November 1973.

Gupta, P. B. *and* C. R. Malaker. Fertility differential with level of living and adjustment of fertility, birth and death rates. *Sankhya* (Calcutta) series B, 25:23–48, November 1963, no. 1–2.

Harman, H. H. Modern factor analysis. 2.ed. rev. Chicago, University of Chicago Press, 1967. 474 p.

Hansen, Morris H., William N. Hurwitz *and* William G. Madow. Sample survey methods and theory; v. 1, Methods and applications. New York, Wiley, 1953. 638 p.

Hansen, Morris H., William N. Hurwitz *and* M. A. Bershad. Measurement errors in census and surveys. *Bulletin of the International Statistical Institute* (Tokyo) 38:359–374, 1961, no. 2.

Hauser, Philip M. Family planning and population programs; a book review article. *Demography* (Washington) 4:397–414, 1967, no. 4.

Hofsten, E. Births variations in populations which practise family planning. *Population studies* (London) 25:315–326, July 1971.

Johnston, J. Econometric methods. New York, McGraw-Hill, 1963.

Kish, L. *and* M. R. Frankel. Inference from complex samples. *Journal of the Royal Statistical Society* (London), 36:1–37, series B, 1974, no. 1.

Kish, L., N. K. Namboodiri *and* K. Pillai. The ratio bias in surveys. *Journal of the American Statistical Association* (Washington) 57:863–876, December 1962, no. 300.

Murthy, M. N. Assessment and control of non-sampling errors in censuses and surveys. *Sankhya* (Calcutta) 25:263–282, series B, December 1963, no. 3–4.

Potter, J. B. The validity of measuring change in fertility by analysing birth histories obtained in surveys. Doctoral dissertation, Princeton, New Jersey, Princeton University, 1975.

Ravenholt, R. T. *and* H. Frederiksen. Numerator analysis of fertility patterns. *Public health reports* (Washington) 83:449–458, June 1968.

Ross, John A. Evaluating demographic control programs. *In* Social change and economic growth. Paris, Development Centre of Organization for Economic Co-operation and Development, 1967. p. 31–56.

Schoenberg, R. Strategies for meaningful comparison. In *Sociological methodology,* 1972. Herbert L. Costner, *ed.* San Francisco, Calif., Jossey-Bass, 1972. p. 1–35.

Schuessler, K. Ratio variables and path models. *In* Structural equation models in the social sciences. A. S. Goldberger *and* Otis D. Duncan, *eds.* New York, Seminar Press, 1973. p. 201–228.

Seltzer, William. Some results from Asian Population Growth Studies. *Population studies* (London) 23:395–406, November 1969.

———. Measurement of accomplishment; the evaluation of family planning efforts. *Studies in family planning* (New York) 1:9–16, May 1970.

Seltzer, William *and* R. S. Fand. A note on the annual variability of the crude birth rate. *In* Proceedings of the Social Statistics Section, 1973. Washington, American Statistical Association, 1974. p. 386–391.

Siegel, Jacob S. Development and accuracy of projections of population and households in the United States. *Demography* (Washington) 9:51–68, February 1972.

Siegel, Paul *and* Robert W. Hodge. A causal approach to the study of measurement error. *In* Methodology in social research. H. M. Blalock *and* Ann B. Blalock, *eds.* New York, McGraw-Hill, 1968. p. 28–59.

Som, R. K. Recall lapse in demographic enquiries. Bombay, Asian Publishing House, 1973. 212 p.

———. A manual of sampling techniques. London, Heinemann, 1973, 384 p.

Stouffer, S. A. Measurement in sociology. *American sociological review* (Albany, New York) 18:591–597, December 1953.

Taves, M. J. An experimental design to preserve randomization in social experiments. *American sociological review* (Albany, New York) 18:90–96, February 1953.

United Nations. Department of Economic and Social Affairs. Manual IV; Methods of estimating basic demographic measures from incomplete data. p. 7–30 and 31–40.
Sales No. E.67.XIII.2.

United Nations. Economic Commission for Asia and the Far East. Assessment of acceptance and effectiveness of family planning methods. Report of an expert meeting, Bangkok, 11–21 June 1968. (Asian Population Studies Series, 4)
Sales No. E.69.II.F.15.

Wiley, David E. *and* James A. Wiley. The estimation of measurement error in panel data. *American sociological review* (Albany, New York) 35:112–117, February 1970.

Wolfers, David. The demographic effects of a contraceptive programme. *Population studies* (London) 23:111–140, March 1969.

———. Some problems in calculating births averted. *In* International Population Conference, Liège, 1973. Liège, International Union for the Scientific Study of Population, 1974. v. 2. p. 233–245.

Interaction problems

Duncan, Otis Dudley. Path analysis; sociological examples. *American journal of sociology* (Chicago, Ill.) 72:1–16, July 1966.

———. Partials, partitions, and paths. *In* Sociological methodology, 1970. Edgar F. Borgatta and G. Bohrnstedt, eds. San Francisco, Calif., Jossey-Bass, 1970, p. 38–47.

———. Introduction to structural equation models. New York, Academic Press, 1975.

Heise, D. R. Problems in path analysis and causal inference. *In* Sociological methodology, 1969. Edgar F. Borgatta, *ed.* San Francisco, Calif., Jossey-Bass, 1969. p. 38–73.

Hermalin, Albert I. Regression analysis of areal data. *In* Measuring the effect of family planning programs on fertility. C. Chandrasekaran *and* Albert I. Hermalin, *eds*. Liège, International Union for the Scientific Study of Population for Development Centre of Organization for Economic Co-operation and Development, 1975. p. 245–300.

Kitagawa, Evelyn M. Components of a difference between two rates. *Journal of the American Statistical Association* (Washington) 50:1168–1174, December 1955, no. 272.

Wright, S. The method of path coefficients. *Annals of mathematical statistics* 5:162–215, September 1934.

Uncontrolled variables

Alwin, Duane F. *and* R. C. Tessler, Causal models, unobserved variables and experimental data. *American journal of sociology* (Chicago, Ill.) 80:58–86, July 1974.

Blalock, H. M., Jr. Making causal inferences for unmeasured variables from correlations among indicators. *American journal of sociology* (Chicago, Ill.) 69:53–62, July 1963.

Duncan, Otis Dudley. Introduction to structural equation models. New York, Academic Press, 1975.

Hauser, Philip M. *and* A. S. Goldberger. The treatment of unobservable variables in path analysis. *In* Sociological methodology, 1971. Herbert L. Costner, *ed.* San Francisco, Calif., Jossey-Bass, 1971. p. 81–117.

APPLICATION OF METHODS OF MEASURING THE IMPACT OF FAMILY PLANNING PROGRAMMES ON FERTILITY: THE CASE OF KARNATAKA STATE, INDIA*

*K. Srinivasan***

INTRODUCTION

Methods of measuring the impact of family planning programmes on fertility have been developed only relatively recently, mostly within the past 15 years. This is because family planning programmes, as endeavours organized directly by Governments or with the support of governmental agencies, are of recent origin; nearly all of them have been implemented within the past two decades. India has the distinction of having been the first Government to formulate a demographic goal and to implement a family planning policy, having taken this step as early as 1951 in conjunction with its first five-year developmental plan. The 1960s and early 1970s witnessed the launching of national programmes of family planning by an increasing number of countries; and as of mid-1975, 63 countries of the world had national programmes of family planning run by governmental departments or by voluntary agencies supported by the Governments.[1]

Most of these countries with national programmes of family planning have also stipulated demographic goals in terms of specified reduction in population growth rates or crude birth rates within a given period of time. They have also set targets of family planning acceptors to be recruited by the programme. The acceptor targets are usually selected so as to be consistent with the desired demographic goal, though considerations of operational feasibility are also taken into account. In general, family planning programmes are looked upon as the prime movers in reducing fertility rates and crude birth rates to desired levels. However, in recent years, especially after the World Population Conference held in Bucharest in August 1974, it has been increasingly questioned whether family planning programmes in fact have been the major instruments for altering fertility patterns, though they might have influenced the marital fertility rates. The crude birth rate in any population is the result of the interactions between numerous demographic, social and biological factors. In the context of the modernization process that is currently taking place in many developing countries, these factors are themselves undergoing change. In such circumstances, the problem of determining the extent to which family planning programmes have contributed to fertility change assumes a crucial significance, both as an interesting area of scientific inquiry and because guidelines are needed as to the effectiveness of further investment and support by Governments and international agencies in family planning programmes vis-à-vis other measures which can also be expected to influence fertility.

The methods for measuring the fertility effects of family planning programmes vary widely in their conceptual schemes, complexity, assumptions involved, data requirements and estimation procedures. Consequently, it can be expected that application of different methods to the same situation may yield different results. In cases where the application of different methods to the same population or geographical area gives similar results, faith in the result is reinforced and the validity of the findings increased. On the other hand, when different methods are applied to different populations; or, for example, if the area in question has been affected by large-scale in- or out-migration, considerable caution has to be exercised in interpreting the validity of the results. A detailed analysis of the results obtained from the application of various methods to the same population can be expected to throw light on the strengths and limitations of various methods, the validity of results obtained and the applicability of a method under different levels of data availability and reliability; and to suggest, one hopes, the nature of future research in this direction.

Objectives of the study

The present study was undertaken at the request of the Population Division of the Department of Economic and Social Affairs of the United Nations Secretariat, as one of three country studies on the application of the methods of measuring the effect on fertility of family planning programmes. The country studies were commissioned with a view to identifying:

(*a*) Problems that arise when an evaluation method is put to use in specified circumstances;

(*b*) Comparison of the results obtained by the different methods and an analysis of the probable reasons for whatever differences are found to exist.

---

* The original version of this paper appeared as document ESA/P/AC.7/2.

** Director, Population Centre, Bangalore, India.

[1] Dorothy Nortman, *Population and Family Planning Programs: A Factbook*, Reports on Population/Family Planning, No. 2, 7th ed. (New York, The Population Council, 1975).

The emphasis of the country studies is, therefore, primarily on the problem areas of method application and on the comparative validity of the results.

The United Nations Secretariat has categorized the currently available methods as follows, recognizing the fact that they are neither mutually exclusive nor completely comprehensive: (*a*) standardization approach; (*b*) trend analysis (fertility projection approach); (*c*) experimental designs; (*d*) couple-years of protection; (*e*) component projection approach; (*f*) analysis of the reproductive process; (*g*) regression analysis (including path analysis); (*h*) simulation models.

A succinct description of each of the above-mentioned methods is given in the background paper[2] prepared by the United Nations Secretariat and hence is not attempted in this paper. This article presents the problems, results and comparative analysis of the findings obtained from the application of different methods to the population of Karnataka State in India for the period 1961–1971. That period was chosen primarily because the most recent population censuses in India were conducted in March 1961 and April 1971, and data on population structure and demographic characteristics could be obtained from those two censuses. Also, the period 1961–1971 witnessed intensified family planning activity in the state, and the results of the application of different methods would cast considerable light on the effect of the programme of family planning on fertility. Although the period of study was chosen to provide maximum advantage from the point of view of data availability, certain essential information could not be obtained and those data which were available were of poor quality. These were major problems faced at every stage of analysis.

One of the consistent findings observed in this case study was that the current methods of assessing the impact of programmes on fertility call for a wide range of information on the population as a whole, not just the people who come into contact with the programme; and changes in these population characteristics over time impede the effective application of these methods in developing countries. All methods, except experimental designs (in which a programme is introduced into one [the experimental] unit and withheld from a matched [control] unit, and results between the two are compared over time), serious assumptions on the dynamics of the base population appear necessary and unless steps are taken to validate these assumptions on the basis of empirical evidence, results obtained from the application of the methods may be misleading. Before proceeding with actual application of the methods to the evaluation, it appears appropriate to describe briefly the population and the programme taken for analysis in this case study.

[2] "Methods of measuring the impact of family planning programmes on fertility: problems and issues" (ESA/P/AC.7/1). See part one of the present publication.

THE STUDY AREA AND THE PROGRAMME

Area and the people

The State of Karnataka was formed in November 1956 by the States Reorganisation Act passed by the Indian Parliament in that year, integrating five subregions, Bombay-Karnataka, Old Mysore, Hyderabad-Karnataka, Madras-Karnataka and Coorg, on the basis of the fact that the language spoken by the majority of the population in each of those subregions was Kannada. After the integration, for administrative purposes, the state was divided into four divisions, Bangalore, Mysore, Belgaum and Gulbarga, which were further subdivided into 19 districts, the first two divisions having five and six districts, respectively, and the last two having four districts each. Though the language of Kannada is a unifying element among all the subregions, there are significant differences in the socio-economic conditions of the population in the five subregions.

The total land area of the state is 190,000 square kilometres and the enumerated population in the 1971 census was 29.3 million. The state ranks sixth among the states in India, in terms of land area, and eighth in terms of population, having 5.4 per cent of the population of the country as a whole. About 7.1 million people, or 24 per cent of the total population of the state, live in urban areas (defined as habitation clusters with more than 5,000 population); and one third of this population lives in the three metropolitan cities of Bangalore, Mysore and Hubli-Dharwar. According to the 1971 census, the crude literacy rates (a literate is defined as a person able to read and write in one language) was 31.5 per cent and only slightly higher than the national average of 29.5 per cent. The proportion of the workers aged 15–59 engaged in non-agricultural activities was 31.2, compared with 24.4 per cent in the country as a whole. A low level of living and income characterize the economy of the state. In 1968–1969, about 95 per cent of the households in the rural areas and 90 per cent of those in urban areas lived on a monthly expenditure of less than Rs 120/– (about $15). For the country as a whole, this figure was around Rs 100/– ($12). Thus from the point of view of extent of urbanization, level of literacy and economic condition, the state is almost at the average level in India, with conditions marginally better than the average.

The population of the area presents a wide diversity of social and cultural patterns within the Indian context. The people in the Bombay-Karnataka subregion, which was part of Bombay presidency before the states were reorganized, follow the traditions and cultural habits of Maharashtrians, and are closer to the northern Indian culture. The people of the subregion of Hyderabad-Karnataka, which was part of the Islamic princely state of Hyderabad before 1956, have been influenced considerably by the Islamic culture and a dominance of northern Indian style of living. The people of Old Mysore and Madras-Karnataka, which constitute about 60 per cent of the population of the state, follow the traditions and cultural values of southern India in dress, eating habits, art and music. The people of the subregion of Coorg, which was centrally administered prior to 1956, had been westernized in their living style owing to their close association with the British over a considerable period of time prior to independence and are more modern in their outlook and habits than the other four subregions. For the state as a whole, 86.5 per cent of the population are Hindus, 10.6 per cent are Moslems, 2.1 per cent are Christians and the other 0.8 per cent adhere to a variety of other religions. However, the religious composition of the population varies slightly from subregion to subregion. The proportion of Moslems in Hyderabad-Karnataka is 15.9 per cent, whereas in Mysore-Karnataka, it is only 8.1 per cent. Similarly, in terms of languages spoken by the population, a significant proportion of the population (22.1 per cent) know, in addition to Kannada, one or more of the languages Tamil, Telugu, Marathi, Urdu and Hindi; but in India as a whole, bilingualism is only 10.1 per cent. Thus, in all aspects, the state is one of the most cosmopolitan in India, with an intermixture of cultural patterns of the south and the north, different religious groups and languages.[3] Summary profiles of the population and the socio-economic characteristics of the state, and of the country as a whole for comparative purposes, are given in table 1.

Population growth

At the censuses of 1971, 1961 and 1951, the inhabitants of the state numbered, respectively, 29.3 million, 23.6 million and 19.4 million (adjusted for the present area). The average annual growth rate during 1951–1961 was 1.96 per cent; during 1961–1971, it was 2.15 per cent. For the country as a whole, the growth rates during the two decades were, respectively, 1.78 and 2.40 per cent. Thus, although Karnataka had a higher growth rate than the national average in 1951–1961, it had a lower than average growth rate in 1961–1971. The reason for this is that the state had achieved a much lower death rate in 1951–1961 than the rest of the country; and although the death rate declined only marginally during the 1960s, the birth rate also began its downward trend. In any year, both the birth and death rates were lower than the national average. The 1950s and 1960s were characterized by marked improvements in medical and public health care for the people through the implementation of national programmes for eradication of malaria, plague and smallpox, and through considerable progress in the organization of primary health centres in rural areas. There was a considerable increase in the number of medical and paramedical personnel employed per thousand population during those two decades. Though these improvements in preventive and curative programmes were executed as a part of the national strategy in all the states of India, Karnataka State had a greater ability to absorb the benefits of the programme even in the 1950s, because of a better developed pre-existing health infrastructure, availability of training institutions for the supply of medical and paramedical personnel and better conditions in terms of communication and over-all development. Consequently, the death rates in the state had always remained lower than the national average. The rate of decline in the death rate during 1951–1961 was also higher than the national decline. In short, the state had always lower birth and death rates than the average figures prevailing in the country; and since 1960, the state growth rates have been less than those for the country as a whole.

Reliable data on fertility, mortality and migration in Karnataka State are not available. The registration of births and deaths is grossly deficient both in coverage and in quality, under-registration amounting to nearly 50 per cent. Consequently, the data from the registration system cannot be used for any analysis of levels and trends in fertility and mortality.

A more reliable source of information on birth and death rates is the Sample Registration System[4] operated by the Government of India throughout the country. A second source for obtaining data on vital rates is the National Sample Survey conducted by the Government of India at periodic intervals. These are multipurpose surveys which include questions on income, expenditure occupation etc. on a sample of households in rural and urban areas; and in a few selected rounds, data on fertility and mortality have also been compiled. The information on birth rates, gross reproduction rates and death rates compiled from available data from the Sample Registration system and the National Sample Survey for rural and urban areas of Karnataka State are provided in table 2.

From table 2, it may be seen that the birth rates in rural Karnataka fell from 40.3 in 1958–1959 to 32.8 in 1972; the death rate dropped from 15.4 to 14.3 during the same period. The figures on birth and death rates computed from the Sample Registration Scheme from 1966 onward show considerable annual fluctuations and it is hard to discern any trends from these rates. This may be due in part to the inherent deficiency in the scheme itself, possibly in matching births obtained from the survey and registration or omission of some events in both the sources, and possibly due also to a real fluctuation in the fertility and mortality levels of the population. This point needs further analysis and is dealt with separately in relevant subsections of this paper.

Family planning programme

As early as 1930, two official family planning clinics were opened in Karnataka by the Government, one at

[3] D. M. Nanjudappa, *Surplus Rural Manpower and Economic Development in Mysore* (Dharwar, Karnataka University, 1968).

[4] India, Registrar General, Vital Statistics Division, *Measures of Fertility and Mortality in India,* Sample Registration System Analytical Series, No. 2 (1972).

TABLE 1. KARNATAKA STATE, A POPULATION PROFILE IN COMPARISON WITH INDIA AS A WHOLE, 1971

Cultural subregion	*Number of districts*	*Population (thousands)*	*Density per square kilometre*	*Literacy rate (literates per 100 population)*			*Percentage urban population*	*Religion*				*Per capita income (rupees)*	*Percentage of population able to speak a language other than their mother tongue*[a]	*Percentage of labour force aged 15-19 engaged in non-agricultural activities*
				Persons	*Male*	*Female*		*Hindu*	*Moslem*	*Christian*	*Other*			
Central Karnataka	1	380	92	44.3	50.2	37.8	15.5	84.6	12.0	3.4	0.0	1 218	32.8	28.8
Madras Karnataka	2	3 060	167	36.7	45.9	27.7	22.8	81.1	12.0	6.4	0.5	636	33.1	36.7
Hyderabad Karnataka	3	3 980	112	19.5	29.9	8.9	16.2	82.3	15.9	1.6	0.2	538	17.8	25.6
Bombay Karnataka	4	7 600	139	33.4	45.4	20.8	23.8	84.9	12.0	1.0	2.1	463	17.7	28.8
Mysore Karnataka	9	14 280	180	32.4	41.7	22.6	27.4	89.7	8.1	1.8	0.4	544	23.1	32.8
All Karnataka	19	29 300	153	31.5	41.6	21.0	24.3	86.5	10.6	2.1	0.8	540	22.1	31.2
All India	356	547 950	167	29.5	39.5	18.7	19.9	82.7	11.2	2.6	3.5	530	10.1	24.4

Source: Unless otherwise noted, data are from the 1971 population census.

[a] Data from 1961 census.

Note:

Period	*Geometrical growth rate*	
	Karnataka State	*India*
1951-1961	1.96	1.78
1961-1971	2.15	2.40

TABLE 2. FERTILITY AND MORTALITY LEVELS, 1958-1972

	Index year									
	1958-1959[a]	*1960-1961*[b]	*1963-1964*[c]	*1966-1967*[d]	*1967-1968*[d]	*1968*[d]	*1969*[d]	*1970*[e]	*1971*[e]	*1972*[e]
Birth rate *(live births per 1,000 population)*										
Rural	40.3	...	35.59	33.01	34.5	33.7	34.1	35.0	34.6	32.8
Urban	...	33.59	31.81	...	...	29.8	28.9	27.8	25.3	28.0
						(Based on half year)				
Combined	...	...	...	...	...	...	...	33.0	31.7	31.5
Gross reproduction rate										
Rural	2.5	...	...	...	...	2.18[f]	2.27[f]	...	2.2[d]	2.2[d]
Urban	...	2.23	...	...	...	...	...	...	1.48[d]	1.75[d]
Combined	...	...	...	...	...	...	...	...	...	...
Death rate *(per 1,000 population)*										
Rural	15.4	...	10.79	14.3	14.5	13.3	15.4	14.2	14.0	14.3
Urban	...	8.20	7.70	...	...	9.0	9.5	10.3	7.2	8.7
						(Second half)				
Combined	...	...	...	...	...	...	...	13.1	12.1	12.8

[a] A. K. De and R. K. Som, *Fertility and Mortality Rates in India, Fourteenth Round, July 1958-June 1959,* National Sample Survey, Report No. 76 (New Delhi, Cabinet Secretariat, 1963).

[b] India, Cabinet Secretariat, *Tables with Notes on the Fertility and Mortality Rates in Urban Areas of India, Sixteenth Round, August 1960-July 1961,* National Sample Survey, Report No. 180 (New Delhi, 1971).

[c] India, Cabinet Secretariat, *Tables with Notes on Differential Fertility and Mortality Rates in India, Eighteenth Round, February 1963-January 1964,* National Sample Survey, Report No. 175 (New Delhi, 1970).

[d] Karnataka State, Bureau of Economics and Statistics, *A Report on the Sample Registration System in Karnataka, 1971-1972,* Sample Registration System Report Series, No. 1 (1974).

[e] India, Registrar General, Vital Statistics Division, *Sample Registration Bulletin,* vol. 9, Nos. 1 and 2 (January and April 1975).

[f] India, Registrar General, Vital Statistics Division, *Measures of Fertility and Mortality in India,* Sample Registration System Analytical Series, No. 2 (New Delhi, 1972).

Vanivilas hospital at Bangalore and the other at Cheluvamba hospital at Mysore. These clinics were set up with the objective of providing family planning advice to couples for the purposes of spacing. Prior to the 1950s, family planning services were provided as a part of social service to women to enable them to have babies by choice and not by chance, and to reduce the incidence of illegally induced abortion and its consequent ill effects. During the first five-year plan in 1951–1956, the family planning programme was introduced as a part of the government policy to achieve not only social welfare goals but the demographic objectives of reducing the birth rates to certain desired levels. There was a shift from viewing family planning as a purely social welfare measure for the protection of the health and well-being of women to a national demographic policy or fertility control measure.

In Karnataka State, as a part of the First Five-Year Plan, a family planning training centre was set up at Ramanagaram in 1952; and in the subsequent years, a number of family planning clinics were opened. The allotment of national funds to the programme increased dramatically from plan to plan: although only Rs. 580,000 were spent on the programme during the Second Five-Year Plan (1956–1961), the amount increased to Rs. 6,630,000 during the third plan and to Rs. 73,470,000 in the fourth plan (1966–1971). The expenditure in the programme during the year 1974/75 was Rs. 22.0 million. During 1962/63, in the third plan period, there was a change of strategy in the programme from one of clinic approach to extension approach. Under the latter method, the message of family planning was to be carried to every eligible couple, and the contraceptive services were to be provided in a socially and psychologically acceptable manner.

Accordingly, there were increased investments in personnel, and the number of family planning acceptors through programme channels also increased steadily. As of 31 March 1975, there had been approximately 450,000 vasectomies, 295,000 tubectomies, 315,000 insertions of intra-uterine devices (IUDs) and 401,000 users of conventional contraceptives, such as condoms and diaphragms.[5] Data on the annual numbers of acceptors of family planning methods in the state during the period 1956–1974 are given in table 3.

DATA FOR THE STUDY

General data problems

The data needed for the application of the different methods of measuring the impact of family planning programmes on fertility can be categorized under the following four headings:

(1) Population structure, including age, sex, marital status distributions and changes in these distributions over time during the period of analysis;

[5] Karnataka, State Family Planning Bureau, Directorate of Health and Family Planning Services, *Family Planning Programme in Karnataka: Progress at a Glance* (1975). For additional information, see R. N. Bhaskar, *Family Planning in Karnataka* (Bangalore, Institute for Social and Economic Change, 1975).

TABLE 3. FAMILY PLANNING ACCEPTORS THROUGH PROGRAMME SOURCES, BY METHOD, 1956-1974

Year	Number of acceptors of sterilization: Vasectomy	Tubectomy	Total sterilizations	Intra-uterine device insertions	Estimated users of conventional contraceptives
1956	853	725	1 578	...	...
1957	920	996	1 916	...	...
1958	567	765	1 332	...	...
1959	953	791	1 744	...	...
1960	1 434	990	2 424	...	...
1961	2 535	1 143	3 678	...	...
1962	4 282	1 592	5 874	...	...
1963	8 254	1 843	10 097	...	...
1964	17 783	3 267	15 050	...	...
1965	17 695	3 219	20 914	57 079	...
1966	41 355	2 602	43 957	84 750	25 684
1967	89 729	5 666	95 395	52 820	35 523
1968	86 877	9 748	96 625	25 092	39 817
1969	40 459	19 200	59 659	14 718	44 122
1970	17 387	29 873	47 260	10 732	47 901
1971	12 548	41 723	54 271	11 326	46 302
1972	68 947	51 185	120 132	12 898	47 300
1973	35 075	49 941	85 016	11 430	46 947
1974	7 595	52 244	59 839	11 530	49 699

Source: Karnataka, State Family Planning Bureau, Directorate of Health and Family Planning Services, *Family Planning Programme in Karnataka: Progress at a Glance* (1975).

(2) Fertility and mortality levels and trends in age-specific fertility rates and life-table values, annually, for the period under investigation. Information on the fertility levels prevailing prior to the commencement of the programme is also needed for estimation of potential fertility;

(3) Annual numbers and characteristics of family planning acceptors within the programme and outside it, from the beginning of the programme through the period of analysis. The characteristics include method of acceptance, age parity distribution at the time of acceptance, age/method-specific continuation rates, use-effectiveness of methods and fertility rates of the acceptors that would have prevailed in the absence of contraception. Data are also needed on the overlap of periods of post-partum amenorrhoea with contraceptive use, incidence of secondary sterility, extent of substitution of methods and extent of acceptance and use of traditional and non-traditional family planning methods obtained outside of the programme;

(4) Information on the social and economic characteristics of the population is also needed for control in the analysis. Such factors as levels of urbanization, literacy rates; proportion of population, especially women, employed in non-agricultural activities; and income have been found to be strongly associated with fertility differentials among groups; and to the extent to which the analysis of the relationship of programme acceptance to fertility can control for these factors, it governs the extent to which the net effects of the programme can be assessed.

The data available for Karnataka State for the period 1961–1971 in each of the above-listed categories had to be compiled from different sources and were of varying coverage and quality; they posed the problems in the application of any method. The major data sources were the censuses of 1961 and 1971, data collected through the National Sample Survey and Sample Registration Scheme, and the family planning service statistics system. The data from the official records of birth and death registers had to be totally ignored because of gross deficiencies in coverage and quality. The censuses, although they provided a variety of information on the population, were grossly defective in respect of a key set of variables that is essential in the evaluation of the impact of the programme. One of these variables is age distribution. As an illustration of the errors in age reporting in the censuses, a chart of the unsmoothed frequency distribution by individual ages for the state as obtained in the 1971 census is given below. Smoothing procedures have to be resorted to for any meaningful analysis. These procedures in themselves have an effect on the ultimate results of any programme evaluation. For some of the variables, the data needed were just not available; and some procedures had to be devised for estimating these variables through indirect methods from available data, or values of these variables had to be borrowed from other populations similar to that of Karnataka State. An example of this type of adjustment is the age distribution of acceptors and method-specific continuation rates. There are some factors for which no data are available, either for the state or for similar populations, and any study has to accommodate to this gap. Among these factors are the pre-acceptance fertility of acceptors, generally used as an index of their future potential fertility; the extent of non-programme acceptance, the incidence of secondary sterility; and efforts of substitution of traditional methods of contraception by modern methods, which are of greater significance in Indian culture.

Faced with these limitations of data, certain attempts were made to assess the quality of data and to adopt appropriate corrective measures to improve the quality; and wherever necessary and unavoidable, to borrow information from sources outside the State. These efforts are briefly described below.

Population distribution

As may be seen from the figure, the age data from the two censuses reflect severe response errors, especially digital preferences. The proportion of the population reporting their age in multiples of five in the 1971 census was 44 per cent. It is also likely that there was under-enumeration of some section of the population, as well as double counting in a few instances. These latter types of errors are not taken into account in this analysis, under the assumptions that they may be of small magnitude, with the incidence being the same in the two censuses. With regard to the smoothing of the age distribution, two procedures were adopted: the 11-point moving-average method[6] and Coale's method.[7]

Moving-average method

In the 11-point moving-average method, between the ages 8 and 67, the population in each age x was first obtained as an average of the 11 years $x - 5$ to $x + 5$; and these were grouped in five-year age intervals with multiples of five as the central points, such as 8–12, 13–17 and 18–22. These totals in the age groups were again smoothed by a three-point moving-weighted-averaged method; population in age group (L) is corrected as $P_L^1 = 1/4P_{L-1} + 1/2P_L + 1/4P_{L+1}$. These were again redistributed to individual ages using Sprague multipliers and regrouped into conventional five-year age groups. For ages 67 years and over, smoothed distributions were obtained by assuming that the second-order differences above that age remained constant. The numbers in the younger ages, 0–7 years, were estimated using the formula $y = A + Bx + CH^x$ (where x is age and y is the number in that age) and adopting a least-square method. This procedure was adopted independently for males and females.

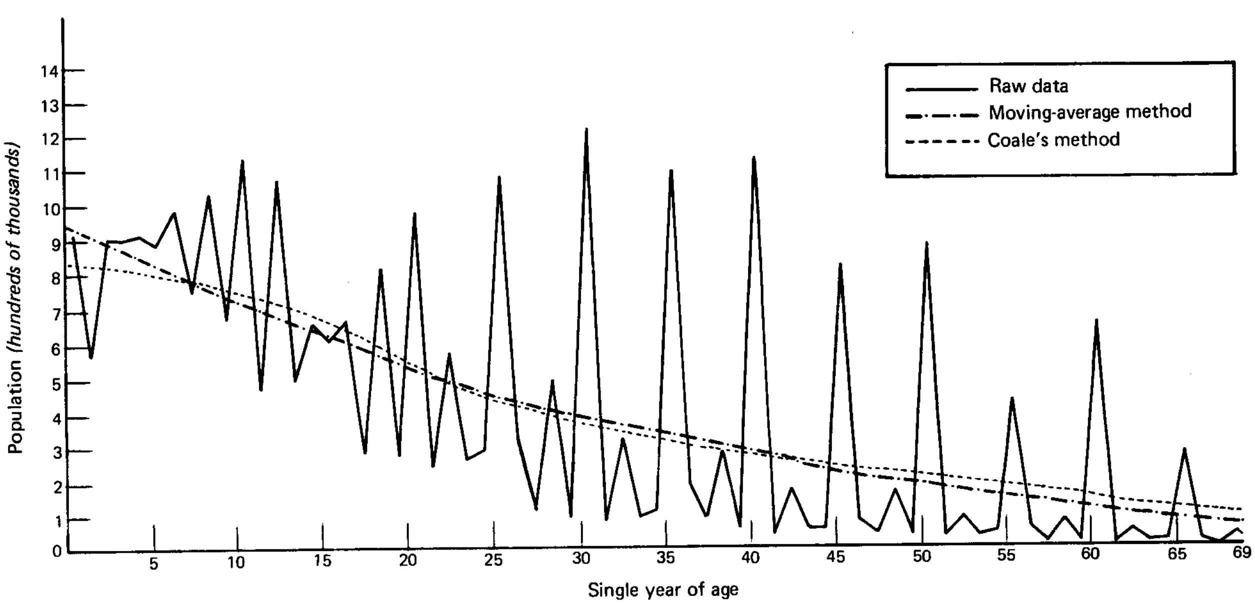
After the smoothed age distribution for each group were obtained, those distributions were allocated to rural and urban areas and to different marital status groups in the same proportions as observed in the unsmoothed data, and adjusted proportionately in such a way that the totals in rural and urban areas and in different marital groups were the same as those enumerated in the census population.

Coale's method

The second procedure that was adopted for smoothing was the one recommended by Coale, using the quasi-stable population theory. In this procedure, the observed age distributions in the censuses are actually ignored and only the totals enumerated are used in conjunction with an approximate idea of the gross reproduction rate (GRR) and the expectation of life (e_0^0) to work out the age distribution of the population on the basis of formulae developed by Coale. In his approach, Coale recommends the use of estimates of GRR and expectation of life (e_0^0) derived by the appli-

[6] *Census of India, 1961,* Paper No. 1 of 1963 (Delhi, 1963). See also S. P. Jain, "Smoothed 1961 census age distribution", *Demography India* (New Delhi), vol. I, No. 1 (October 1972).

[7] Ansley J. Coale, "Constructing the age distribution of a population recently subject to declining mortality", *Population Index*, vol. 37, No. 2 (April-June 1971), pp. 75–82. See also Ansley J. Coale and Paul Demeny, *Regional Model Life Tables and Stable Populations* (Princeton, N.J., Princeton University Press, 1966).



Raw and smoothed population distributions of Karnataka State, based on 1971 census

cation of quasi-stable population theory for the purposes of estimating the correction factors to be applied to model stable age distributions approximating the given population. However, in the present exercise, the values of GRR and e_0^0 for 1961 and 1971 used were those assumed by the Bureau of Economics and Statistics for official projection purposes.

The smoothed distributions obtained using the two procedures described above are also charted in the figure, for comparative purposes. The actual population distributions obtained after smoothing for 1961 and 1971 for rural, urban and combined populations by the two smoothing procedures are given in annexed tables 15–18.

It was decided at the outset to adopt two smoothing techniques, because it was considered that the procedures adopted for smoothing could yield different results in the estimation of impact on fertility of the family planning programme. The two smoothing procedures adopted are conceptually very different; in the moving-average method, which is strictly an algebraical exercise, based on the census enumerated age distribution of the population, no assumptions of the fertility and mortality levels and the past trends in those levels are involved; in the second method, Coale's method, the observed age distribution in the census is virtually ignored and resort is had to stable population theory, and age distribution of approximate stable populations are corrected on the basis of estimates of fertility and mortality and trends of those parameters in the past. What is the nature of differences that could arise in the results on the impact on fertility of family planning programmes, if these two age distributions of the population obtained by two different smoothing rules were used in the various calculations? To attempt to answer that question, each of the methods of evaluating the programme impact was applied to the two population age distributions.

To give a comparative idea of the results obtained from the use of the two population age distributions, table 4 shows the average birth rates, gross reproduction rates and death rates estimated for the period 1961–1971 by the application of the forward-survival procedure, the quasi-stable population theory[8] and the reverse-survival method from the 0–9 age group in 1971. It can be seen from table 4 that the estimates of birth rates obtained for 1961–1971 vary widely, depending both upon the smoothing procedure adopted as well as the estimating technique used. Since the application of Coale's procedure presupposes stability over time in the fertility rates of a population and it has been documented from more than one source that the fertility levels in India have been declining since 1966, the validity of any results obtained from the application of a stable or quasi-stable population theory could be questioned. As such, the results based on the moving-average age distributions without any assumptions on the stability of the birth rates over time could be taken as relatively more valid to the context in Karnataka State after 1966.

Fertility and mortality measures

As mentioned earlier, fairly reliable data on levels and trends in fertility and mortality for Karnataka State could be obtained only from the sample surveys conducted at different times by the National Sample Survey and also from the Sample Registration System of the Government of India. The Sample Registration System provides information on the birth and death rates from 1966–1967, generally on the basis of half-yearly figures as well as annual figures. The National Sample Survey provided data on the vital rates for the years 1958–1959, 1960–1961 and 1963–1964. The rates compiled from these two different sources have been provided in table 2. It is interesting to note that even as early as 1958–1959 the birth rate for the rural areas of Karnataka State was 40.3. The Mysore population study conducted in 1951–1952[9] under the joint auspices of the Government of India and the United Nations also provided an estimate of 40.0 for the old Mysore State. Thus, it appears that fertility in Karnataka State even prior to the commencement of any organized large-scale family planning activity was comparatively low, being on the order of 40 per 1,000

[8] *Manual IV. Methods of Estimating Basic Demographic Measures from Incomplete Data* (United Nations publication, Sales No. 67. XIII.2).

[9] *The Mysore Population Study* (United Nations publication, Sales No. 61. XIII.3). See also Mysore State, Bureau of Economics and Statistics, *Population Projections for Mysore, 1972–1986* (1973).

TABLE 4. VITAL RATES FOR 1961-1971, ESTIMATED BY VARIOUS METHODS

Age distribution used	Estimating method: Forward-survival ratio		Quasi-stable population theory			Reverse-survival ratio		Geometrical growth rate
	Birth rate	Death rate	Birth rate	Death rate	Gross reproduction rate	Birth rate	Death rate[a]	
Raw age distributions	41.87	20.36	42.71	21.20	2.885	42.01	20.50	2.151
Smoothed by moving-average method (S_1)	41.80	20.39	44.48	22.97	3.095	41.94	20.43	2.151
Smoothed by Coale's method (S_2)	35.29	13.78	38.43	16.92	2.687	34.69	13.18	2.151

Source: Method S_2 based on procedure given in Ansley J. Coale, "Constructing the age distribution of a population recently subject to declining mortality", *Population Index*, vol. 37, No. 2 (April-June 1971), pp. 75–82.

[a] Death rates calculated by subtracting growth rate from birth rate.

population. It has been estimated that the biologically maximum level of crude birth rate that can be sustained over any considerable periods of time is on the order of 60–65 per 1,000 population. Thus, the fact that the pre-modern, pre-family planning fertility levels in India were much lower than the biological potential and below the levels prevailing even in 1975 in many of the developing countries indicates the effects of various cultural practices, social norms and traditions in dampening the fertility levels. Cultural practices that had such a lowering effect on fertility were the restrictions on widow remarriage, long periods of abstinence within marriage due to religious reasons or beliefs, the transfer of pregnant women to their mother's home for confinement and the practice whereby the women remained in their maternal home for considerable periods of time even after delivery and long periods of breast-feeding of children. With modernization, these traditional checks are being gradually loosened and the potential fertility of couples is in a state of flux. In the absence of any data for the state at any earlier point of time, the age-specific fertility patterns obtained from the National Sample Surveys in 1958–1959 for rural areas and in 1960–1961 for urban areas were taken as the natural fertility levels prevailing in the population of the state in the absence of any large-scale family planning efforts. These patterns agree quite closely to the patterns observed in the Mysore population study. The age-specific fertility rates and marital fertility rates computed from various sources for the years around 1961 and 1971 are given below in table 5.

The estimates of the birth rates obtained from the sample registration data for the year 1968 indicate moderate annual fluctuations with a general declining trend. These fluctuations may be due in part to errors in the collection of data in the sample registration system, including errors of matching, omission of events in both sources and errors of reference period. They can also include some real fluctuations caused by the interaction of fertility-reducing effects of modern methods of contraception adopted by some couples with the fertility-increasing effects caused by the decline in application of traditional checks due to modernizations, mentioned earlier.

To study the trends in these rates after 1968, linear regressions were fitted to the rural and urban rates over a five-year period, 1968–1972. Separate regressions based on data for the first half-year and the second half-year over these years were also fitted. These rates and regressions are given in table 6.

On the whole, birth rates have been declining at a rate of 0.53 per annum since 1968, 0.52 points in urban areas and 0.43 points in rural areas. When one considers the data separately for the first half-year and the second half-year, different trends are observed. Based on the data for the first half of the year, an increasing trend is observed for the rural areas; but based on the second half year, a steep declining trend is noticed. This indicates the possible presence of some errors of coverage in the sample registration data, wherein events of the first half of the year have been erroneously classified in the second half of the year in the earlier years of the Sample Registration Scheme. The presence of other types of errors, including omissions of vital events, cannot be ruled out and possibly the magnitude of such errors differ from year to year. It is worth-while noting that the beta coefficients are not statistically significant at the 10 per cent level mainly because of the smallness of the sample.

In general, it is believed that the rates based on the Sample Registration System somewhat underestimate the true picture. However, in view of the fact that no precise data are available to quantify the extent of under-registration and also in the light of the fact that the quality of the information in the Sample Registration System has improved over the years, the rates for 1971 are presumed to represent the fertility and mortality picture prevailing among the population of the state. Thus, although the data on the fertility and mortality levels for the year 1961 were obtained from the National Sample Survey, the data for 1971 were obtained from the Sample Registration System. These data, given in table 5, were taken as the basis for further evaluative analysis.

Family planning programme data

The data on the number of acceptors of various methods of family planning are available from the service statistics records published annually by the Directorate of Health and Family Planning. Though exact figures are available on sterilizations and IUD insertions, the data on conventional contraceptives, such as condoms, diaphragms, jellies and foam tablets, are not equally reliable. With regard to conventional contraceptives, the usual procedure adopted by the Government of India is to convert the quantity of the total contraceptives distributed into number of person-years of use, by dividing the number of condoms and foam tablets distributed by 72, which is supposed to represent the average frequency of coitus per couple per year, by 2 for each diaphragm fitted and by 7 for each tube of vaginal jelly distributed. Thus, the conventional contraceptives distributed are directly converted into couple-years of use rather than into the number of acceptors of these methods. Though this method facilitates the estimation of the number of births averted, there are serious assumptions involved with regard to the extent to which the contraceptives distributed have been actually used, the average coital frequency among the users and the effectiveness of these methods during the period of use. Unfortunately, no data are available for empirical verifications of these assumptions.

Although for some years the numbers of acceptors of family planning methods were available on a fiscal-year basis (from 1 April to 31 March), for the rest, the data were available on a calendar-year basis. With some simplistic assumptions of uniform distribution of

TABLE 5. AGE-SPECIFIC FERTILITY RATES FOR ALL WOMEN AND FOR MARRIED WOMEN, SELECTED PERIODS

	1951			1958-1959		1960-1961		1961[a]		1971			1971[b]		
	MASFR			ASFR	MASFR[c]	ASFR	MASFR[d]	ASFR	MASFR	ASFR			MASFR		
Age group	*Rural*	*Urban*	*Combined*	*Rural*	*Rural*	*Urban*	*Urban*	*Combined*	*Combined*	*Rural*	*Urban*	*Combined*	*Rural*	*Urban*	*Combined*
15-19	279.8	283.0	280.6	204.6	290.0	129.3	247.9	186.0	281.7	114.6	65.3	101.0	219.5	188.4	213.1
20-24	315.4	261.2	302.3	274.0	294.7	224.5	263.8	262.2	287.8	234.8	149.7	211.1	263.2	194.7	246.1
25-29	309.8	213.0	288.0	257.1	275.5	229.9	250.8	250.9	270.0	232.8	190.7	223.8	247.0	207.2	238.9
30-34	190.5	193.0	191.1	160.5	181.9	184.8	206.3	165.7	187.2	165.5	132.8	157.8	181.4	144.8	172.7
35-39	172.2	107.0	157.6	109.5	133.9	107.3	127.5	109.0	132.5	129.0	82.9	118.0	146.8	93.0	133.9
40-44	59.4	42.1	55.5	32.2	46.6	33.3	46.5	32.4	46.6	49.7	33.8	46.2	63.4	42.7	58.8
45-49	...	...	...	15.8	26.5	10.9	17.9	14.8	24.7	16.7	16.0	16.5	23.8	22.6	23.6
TFR (TMFR).	6.6	5.5	6.4	5.3	6.3	4.6	5.8	5.1	6.2	4.7	3.4	4.4	5.7	4.5	5.4

Sources: For 1951, *The Mysore Population Study* (United Nations publication, Sales No. 61.XIII.3); for 1958-1959, A. K. De and R. K. Som, *Fertility and Mortality Rates in India, Fourteenth Round, July 1958-June 1959*, National Sample Survey, Report No. 76 (New Delhi, Cabinet Secretariat, 1963); for 1960-1961, India, Cabinet Secretariat, *Tables with Notes on the Fertility and Mortality Rates in Urban Areas of India, Sixteenth Round, August 1960-July 1961*, National Sample Survey, Report No. 180 (New Delhi, 1971); for 1971, Karnataka State, Bureau of Economics and Statistics, *A Report on the Sample Registration System in Karnataka*, 1971-1972, Sample Registration System Report Series, No. 1 (1974).

[a] Rural rates for 1958-1959 and urban rates for 1960-1961 were combined to obtain 1961 rates.

[b] Marital age-specific fertility rate calculated using married proportion of 1971 population smoothed by the moving-average method.

[c] Age-specific fertility rate converted into marital age-specific fertility rate using the married population of 1961, smoothed by the moving-average method.

[d] Age-specific fertility rate (MS) calculated using married proportion (urban) of 1961 population smoothed by the moving-average method.

Notes: ASFR = age-specific fertility rate;
MASFR = marital age-specific fertility rate;
TFR = total fertility rate;
TMFR = total marital fertility rate.

TABLE 6. KARNATAKA STATE: TRENDS IN CRUDE BIRTH RATES, BASED ON SAMPLE REGISTRATION DATA, 1968-1972

Birth rates based on:	*1968*	*1969*	*1970*	*1971*	*1972*	*Regression equation*
First half-year						
Rural	29.8	31.6	31.9	33.6	30.8	$y = 30.74 + 0.40x$
Urban	...	27.4	25.9	24.2	27.0	$y = 26.850 - 0.29x$
Second half-year						
Rural	38.1	36.8	38.5	35.2	35.0	$y = 38.28 - 0.78x$
Urban	29.8	30.3	29.4	26.5	29.3	$y = 30.02 - 0.48x$
Combined						
Rural	33.7	34.1	35.0	34.6	32.8	$y = 35.20 - 0.43x$[a]
Urban	...	28.9	27.8	25.3	28.0	$y = 28.8 - 0.52x$
Combined	...	32.84[b]	33.0	31.7	31.5	$y = 33.59 - 0.53x$

[a] Based on only four parts, beginning in 1969.

[b] Estimated figures using the rural/urban proportion at the 1971 census. The origin of all the equations fitted is taken as 1968.

acceptance over the months within a year, all data were converted to calendar-year acceptance figures. These are the figures provided in table 3.

The age distribution of the acceptors was available from the data published by the Directorate of Health and Family Planning Services on the basis of analysis of a sample of the acceptor records from the year 1969 onward. Since no information was available for the earlier years, it was assumed that the age distributions of the acceptors observed for each method for the year 1969 were also applicable to the figures in the earlier years. The age distributions of acceptors of different methods for the different years are given in table 7.

With regard to the continuation rates, no follow-up studies appear to have been conducted in Karnataka State, and it was necessary to apply data from other studies. The rates that have been published[10] on a national basis for IUDs from the survey conducted by the National Institute of Family Planning (NIFP) in 1970–1971 were adopted with minor modifications. A small-scale follow-up study of IUD acceptors in an industrial establishment at Bangalore revealed that the IUD continuation rates were lower than those reported in the study mentioned above. It was also surmised from discussions that in the national study the initial immediate drop-outs, i.e., within one week, or expulsions of the IUDs had not been included in the computation of the continuation rates. Consequently, the a value for the continuation function ae^{-rt} was obtained from the follow-up study done at Gandhigram,[11] and using these a values, the annual discontinuation rate r was recalculated from the NIFP study. The values of a and r thus derived and used in this analysis are also given in table 7.

With regard to sterilizations, the annual discontinuations that were assumed were due to mortality of either the husband or the wife, and the possibility of the woman having completed age 49 (assumed to be the age after which fertility becomes zero). For the conventional contraceptives, as the data provided an estimate of the couples' years of use for different years, these data were directly converted into the births averted assuming that the effectiveness of the method during use is 75 per cent (i.e., one couple-year of protection of conventional contraceptives equals 75 per cent of effective years of use).

Socio-economic data

Data on socio-economic and family planning programme input variables are essential as control factors for any analysis of programme impact on fertility. Since a major consideration in programme planning is population size or areal unit, such an analysis must be macro or areal in nature. The convenient administrative areal unit for such an analysis in the Indian setting is the Primary Health Centre (PHC) in rural areas and the census block in urban areas. Though, with considerable efforts through compilation from various documents, information could be obtained on certain socio-economic factors, such as literacy rate, proportion of population employed in non-agricultural activities and proportion of villages and houses electrified, and on such programme factors as inputs of personnel and money in each area, no data are available from any existing sources on fertility and mortality rates at the PHC level or even at the district level. Without knowledge of the dependent variable, no areal or regression analysis is possible for the state. However, an attempt is being made to relate programme acceptance to other areal variables; and from the special survey being conducted by the Population Centre,[12] regression of fertility or fertility change on other factors would become

[10] P. S. Mohapatra and others, *A Four-Year Follow-up Study of Intra-Uterine Contraceptive Device Acceptors*, Series 19 (New Delhi, National Institute of Family Planning, 1973).

[11] A. Muthaiah, "Termination rates and other contraceptive use of IUD acceptors in Arthoor Block", *Bulletin of the Gandhigram Institute of Rural Health and Family Planning*, vol. 5, No. 2 (December 1970).

[12] P. H. Reddy and others, "Fertility, mortality and demand for family planning: a longitudinal study in progress", *Population Centre Newsletter* (Bangalore, India), vol. 1, No. 4 (May-June 1975).

TABLE 7. ESTIMATED PERCENTAGE DISTRIBUTION [a] OF ACCEPTORS BY AGE AND METHODS USED, AND INTRA-UTERINE DEVICE CONTINUATION RATES BY AGE OF WOMEN

Age group	Up to 1969				1970				1971				Continuation rates for intra-uterine device [c]
	Vasectomy	Tubectomy	Intra-uterine device	Conventional contraceptives [b]	Vasectomy	Tubectomy	Intra-uterine device	Conventional contraceptives [b]	Vasectomy	Tubectomy	Intra-uterine device	Conventional contraceptives [b]	
15-19	1.99	0.24	2.14	5.43	1.10	0.27	2.88	5.43	1.20	0.30	4.22	5.43	$a = 0.96$
20-24	7.70	6.48	18.84	25.96	9.50	7.99	19.21	25.96	10.01	8.66	22.19	25.96	$r = 0.4158$
25-29	22.40	28.86	34.65	40.53	22.70	31.27	30.42	40.53	24.22	33.13	31.43	40.53	$a = 0.81$, $r = 0.2524$
30-34	31.01	36.24	28.22	19.36	30.70	35.37	27.80	19.36	32.53	34.44	25.70	19.36	$a = 0.81$, $r = 0.2160$
35-39	22.33	22.36	13.08	7.18	23.10	20.54	16.01	7.18	21.02	18.53	12.95	7.18	$a = 0.99$
40-44	14.57	5.82	3.07	1.06	12.90	4.56	3.88	1.06	11.02	4.94	3.21	1.06	$r = 0.3756$
45-49	0.0	0.0	0.0	0.48	0.0	0.0	0.0	0.48	0.0	0.0	0.0	0.48	

[a] For vasectomies and condoms (among the conventional contraceptives), age data relate to age of the client's wife; for all other methods, the age is that of the client herself.

[b] Data on conventional contraceptives are from the State Family Planning Bureau of Karnataka. For users of conventional contraceptives, it was assumed that 75 per cent drop out immediately and that the data on users are for that year only.

[c] Continuation rates were derived from the studies of the National Institute of Family Planning at Gandhigram.

possible. For the present report, no such analysis could be attempted.

ANALYSIS OF PROGRAMME IMPACT

The data compiled from various sources and adjusted to the extent possible for response errors, as described above, were used in estimating the impact on fertility of family planning programmes in Karnataka State, during the period 1961–1971. The following methods were applied to estimate the impact:

(*a*) Standardization approach;
(*b*) Trend analysis (fertility projection approach);
(*c*) Couple-years of protection;
(*d*) Component projection approach;
(*e*) Experimental designs (matching studies);
(*f*) Simulation models.

The results derived from the application of each of those methods are briefly discussed below.

Standardization approach

The indirect standardization technique was adopted for estimating the effects of changes in the age distribution, marital status, rural-urban composition and marital fertility on the crude birth rates between 1961 and 1971. The age-specific fertility rates of 1971 were applied to the population distributions of 1961, and the estimates of birth rates, standardized separately and jointly for the factors of age, marital status and rural-urban composition, were derived. The results are presented in table 8.

TABLE 8. RESULTS OF APPLICATION OF INDIRECT STANDARDIZATION PROCEDURES IN ANALYSING FERTILITY CHANGE IN KARNATAKA STATE, 1961-1971

	Population distributions of 1961 and 1971 smoothed by	
	Method S_1 [a]	*Method* S_2 [b]
Crude birth rate; 1961	39.6	38.0
Crude birth rate; 1971	32.0	30.6
Amount of decline, 1961-1971	7.6	7.4
Percentage change in crude birth rate, 1961-1971	19.3	19.4
1971 crude birth rate, standardized for 1961 age distribution	32.3	30.9
1971 crude birth rate, standardized for 1961 age and marital status	34.3	33.8
1971 crude birth rate, standardized for 1961 age, marital status and rural/urban residence	34.4	34.0
Percentage change in crude birth rate, 1961-1971, attributable to:		
Change in age	4.6	4.9
Age and marital status	31.0	43.6
Age, marital status and rural/urban residence	32.4	45.9
Percentage change in marital fertility	67.6	54.1

Source: Method S_2 based on procedure given in Ansley J. Coale, "Constructing the age distribution of a population recently subject to declining mortality", *Population Index,* vol. 37, No. 2 (April-June 1971), pp. 75-82.

[a] Moving-average method.

[b] Coale's method.

The procedure adopted for separating the effects of various factors is the same as that adopted by Freedman and others[13] in the analysis of the fertility decline in Hong Kong. The age-specific fertility schedules for 1961 and 1971 were those obtained from the National Sample Survey and the Sample Registration Scheme, respectively, as presented in table 5. However, two sets of population distributions were used, the first obtained by application of the moving-average method (S_1) and the second by Coale's method (S_2), mainly in order to study the differences in the smoothing procedures on the results obtained.

It can be seen from the table that when using the population smoothed by the moving-average procedure (S_1), age distribution changes accounted for 4.6 per cent of the change in the crude birth rate, age and marital status changes caused 31.0 per cent and age-marital status and rural-urban composition changes jointly accounted for 32.4 per cent of the crude birth-rate decline. The balance of 67.6 per cent of the change in the crude birth rate can thus be attributed to changes in the marital fertility of the population. On the other hand, if one uses the population figures smoothed by Coale's method (S_2), the changes in marital fertility during 1961–1971 accounted for a lower percentage of change in the crude birth rate, 54.1 per cent.

Thus, it may be seen that differences in the smoothing procedures could lead to different estimates of birth-rate changes attributable to various factors. For reasons stated above, the application of Coale's method, based on stable population theory, appear to have serious limitations in the analysis of fertility trends in the short run; and, hence, greater validity can be placed in the results obtained on the basis of population smoothed by the moving-average method. This statement is relevant for all the findings in this paper.

Trend analysis

The population of 1961 was projected, year by year, up to 1971, under two sets of fertility assumptions: (1) that the age-specific fertility rates remained at the level noticed in 1961; and (2) that the age-specific fertility rates declined linearly from 1961 to the 1971 levels observed in the Sample Registration Scheme. As in the standardization technique, two sets of results were obtained based on the two sets of population distributions for 1961, smoothed by the moving-average method (S_1) and by Coale's method (S_2). The mortality assumptions for each of the methods were based on

[13] Ronald Freedman and others, "Hong Kong's fertility decline, 1961–68", *Population Index,* vol. 36, No. 1 (January-March 1970), pp. 3–18. See also Ronald Freedman, "A comment on 'Social and Economic Factors in Hong Kong's Fertility Decline' by Sui-ying Wat and R. W. Hodge", *Population Studies* (London), vol. XXVII, No. 3 (November 1973), pp. 589–595.

the expectation of life derived by the reverse-survival procedure, for that particular method by comparing appropriate cohorts between 1961 and 1971, and using the procedure outlined in *Manual IV.*[14] The mortality levels estimated from the two sets of population distributions are given above in table 4. The total number of births estimated to have occurred under different assumptions are provided below in table 9. The total

TABLE 9. ESTIMATES OF TOTAL BIRTHS THAT OCCURRED AND OF BIRTHS AVERTED DURING 1961-1971, OWING TO THE FERTILITY DECLINE

		Population distributions of 1961 and 1971 smoothed by	
		Method S_1 [a]	*Method* S_2 [b]
Estimated births (1961-1971) assuming continuation of 1961 fertility pattern	B_1 ...	11 551 900	11 464 500
Estimated births that occurred during 1961 to 1971 taking into account the observed fertility decline	B_2 ...	10 434 700	10 323 400
Estimated births averted because of fertility decline	B_3 ...	1 117 200	1 141 100
	$\frac{B_3}{B_2} \times 100$ ...	10.7	11.1

Source: Method S_2 based on procedure given in Ansley J. Coale, "Constructing the age distribution of a population recently subject to declining mortality", *Population Index,* vol. 37, No. 2 (April-June 1971), pp. 75-82.

[a] Moving-average method.

[b] Coale's method.

number of births, assuming constant fertility from 1961 to 1971, are estimated at 11,550,000 with the population distribution based on the moving-average method; (S_1) and at 11,460,000 based on Coale's method (S_2). The births that have been averted because of the fertility decline were estimated as 1,120,000 based on the S_1 population distribution and at 1,140,000 on the S_2 distribution. The births averted because of fertility decline as a proportion of the total births that occurred during 1961 and 1971 is found to be 10.7 per cent according to S_1, and 11.1 per cent according to S_2. The latter method of smoothing the population slightly over-estimated the proportion of births averted because of the fertility decline, based on the component projection method; but, on the basis of standardization technique, the opposite effect obtained. Of course, the differences in this case are not significant.

Couple-years of protection[15]

The births that were averted by the number of acceptors of various methods of family planning prior to 1971 were estimated for each year, beginning in 1961. Based on assumptions of continuation rates for each method, as described earlier, the number of couple-years of protection contributing to the saving of births for each calendar year was estimated. This estimate of couple-years of use was made for each age group separately; and assuming that the age-specific fertility rates for 1961 would have continued in the absence of contraception (i.e., potential fertility of acceptors is the same as the average marital fertility rates of all women in that age group), the estimates of births averted in each year because of the family planning programme were derived. For working out the couple-years of protection and births averted, the acceptance data available from 1956 were used. The results obtained for each family planning method, based on the two population distributions, S_1 and S_2, are presented in table 10.

It can be seen that although the total fertility decline between 1961 and 1971 prevented approximately 1,120,000 births according to S_1 distribution (and about 1,140,000 according to S_2 distribution) the proportion of those births averted because of acceptance of family planning methods from within the programme was 37.4 per cent in S_1 (and 36.2 per cent in S_2). Although the estimate of births averted because of the fertility decline has consistently increased over the years, the proportion of those accountable by family planning increased each year up to 1968; thereafter, a declining trend is discerned. Thus, since 1969, the non-programme effects or indirect effects of the programme have begun to play an increasingly significant role in fertility decline in the state. In general, both the estimates of births averted by fertility decline and the proportion accountable to direct programme acceptance are slightly higher by S_1 distribution than by S_2 distribution. During 1971, the total births averted because of fertility decline was 219,800, taking S_1 distribution; and 209,200, taking S_2 distribution, and the proportion of those attributable to family planning practice in the earlier years was 36.4 per cent by S_1 and 37.7 per cent by S_2. It is significant to note that only one third of the decline in marital fertility could be attributable to recorded family planning acceptance within the programme.

Component projection approach

The component projection method first developed by Lee and Isbister[16] in estimating the impact of family planning programmes on fertility is almost similar to

[14] *Manual IV. Methods of Estimating Basic Demographic Measures from Incomplete Data.*

[15] Samuel M. Wishik and K. H. Chen, *The Couple-year of Protection: A Measure of Family Planning Program Output,* Manuals for Evaluation of Family Planning and Population Programs, No. 7 (New York, Columbia University, International Institute for the Study of Human Reproduction, 1973). See also W. Parker Mauldin, "Births averted by family planning programs", *Studies in Family Planning,* vol. 1, No. 33 (August 1968).

[16] B. M. Lee and John Isbister, "The impact of birth control programs on fertility", in Bernard Berelson and others, eds., *Family Planning and Population Programs. A Review of World Development* (Chicago, University of Chicago Press, 1966), pp. 737–758.

Table 10. Births averted by different family planning methods, 1961-1971, based on continuation rates given in table 7 and potential fertility represented in table 5

	Method[a]	1961	1962	1963	1964	1965	1966	1967	1968	1969	1970	1971	Total
Vasectomy	S_1	907	1 425	2 406	3 954	6 199	10 880	21 530	35 066	42 898	45 563	42 392	213 221
	S_2	897	1 411	2 379	3 908	6 123	10 760	21 310	34 659	42 343	42 877	41 754	208 421
Tubectomy	S_1	907	1 080	1 300	1 660	2 180	2 446	2 992	4 164	6 530	10 680	16 775	50 714
	S_2	894	1 062	1 280	1 635	2 078	2 407	2 947	4 103	6 438	10 531	16 584	49 959
Intra-uterine device	S_1	...	...	...	...	5 165	15 502	21 904	21 535	18 277	14 846	12 218	109 447
	S_2	...	...	...	...	5 143	15 425	21 775	21 382	18 175	14 597	12 067	108 564
Conventional contraceptives	S_1	...	...	...	...	...	4 736	6 550	7 341	8 135	8 832	8 537	44 131
	S_2	...	...	...	...	...	4 736	6 550	7 341	8 135	8 832	8 537	44 131
Total	S_1	1 814	2 505	3 706	5 614	13 544	33 564	53 976	68 106	75 840	79 921	79 922	417 513
	S_2	1 791	2 473	3 659	5 543	13 544	33 328	52 582	67 485	75 091	76 837	78 942	411 175
Estimated births averted because of decline in marital fertility	S_1	...	17 500	35 800	54 900	75 000	95 200	117 900	141 600	166 600	192 900	219 800	1 117 200
	S_2	...	17 100	35 600	55 600	77 300	100 500	123 100	147 200	172 300	198 700	209 200	1 136 600
Percentage decline because of family planning programme	S_1	...	14.3	10.3	10.2	18.1	35.3	44.9	48.1	45.5	41.4	36.4	37.4
	S_2	...	14.5	10.3	10.0	17.3	31.7	42.7	45.8	43.6	38.7	37.7	36.2

Source: Method S_2 based on procedure given in Ansley J. Coale, "Constructing the age distribution of a population recently subject to declining mortality", *Population Index*, vol. 37, No. 2 (April-June 1971), pp. 75-82.

[a] Population distribution in 1961 and 1971 smoothed by: method S_1, moving-average; S_2, Coale's method.

that for couple-years of protection, described above, except for the assumption with respect to potential fertility of acceptors. The number of acceptors of different family planning methods prior to 1971 are survived to mid-1970 by applying the appropriate mortality schedules and the method-specific continuation rates. The number of couples using contraceptives by the middle of 1970 is estimated for each age group separately. This was multiplied by the potential fertility of acceptors obtained from the results of matching of the fertility of acceptors and non-acceptors over a three-year period prior to acceptance. The details of the matching study are given in the next section. It was found that the pre-acceptance fertility of acceptors has been roughly 15 per cent higher than that of the control group, though there was variation among different age groups between the matches and control. Assuming the higher potential fertility for the acceptors, the births averted during 1971 were estimated and found to be on the order of 90,000. The results are provided in table 11. Subtracting the births averted, the estimates of age-specific fertility rates in 1971 were derived, and it is interesting to note that the fertility rates derived by assumptions of the two population-smoothing procedures (S_1 and S_2) are very nearly the same. The percentage reduction in marital fertility because of family planning, during the year 1971, has been estimated to be 41 per cent based on S_1 and 43 per cent based on S_2. The estimate of the total fertility rate in 1971 thus obtained using the component projection method and taking into account the impact of the family planning programme, works out to very nearly the same value with a fertility rate of 4.6 per woman, while the total fertility figures computed from the Sample Registration Scheme data for 1971 was 4.4. Details regarding births averted by each method in each age group in 1971 and the percentage of total births averted attributable to the family planning programme are given in table 11.

Experimental designs

Data for the pre-acceptance and post-acceptance fertility of acceptors of different family planning methods, compared with control groups matched for age, parity and open birth interval, were compiled from the information collected in the Longitudinal Sample Survey of households, currently being conducted by the Population Centre in selected villages of Bangalore city, Bangalore rural districts and Kolar districts. This large-scale sample study envisages collection of data on fertility, mortality, family planning practice and related social and programmatical factors from 5,200 households selected according to scientific sampling procedures from the five districts of the project area. At this writing, the survey had been completed in a sample of 1,200 households (700 in Bangalore city, 250 in Bangalore rural districts and 250 in Kolar districts) and the information collected from these households was available for analysis. From these schedules, data were compiled on the age, parity and open birth interval of family planning acceptors of 1969, 1970 and 1971 at the time of acceptance; their pre-acceptance and post-acceptance fertility over a three-year period and the fertility of non-acceptors over the same duration matched for age (within the same five-year age group); parity (the same parity 0, 1, 2, 3, 4, 5, 6+), and open birth interval (less than 12 months and over 12 months). The total number of acceptors during 1969, 1970 and 1971 as obtained from these samples of 1,200 households was only 35 and the number of matches selected was 105 (three match to one acceptor). The pre-acceptance and post-acceptance fertility rates for the acceptors and for the matches, computed separately for vasectomies, tubectomies and IUD cases, are given in table 12.

The results from this table have to be used with a considerable amount of caution as the data are based on a very small sample. It can be seen that the pre-

TABLE 11. IMPACT OF THE FAMILY PLANNING PROGRAMME AS MEASURED BY THE COMPONENT PROJECTION APPROACH

	Estimated number of couples using respective methods during mid-1970 who entered the conventional age group during 1971					*Number of births averted*					*Estimated age-specific fertility rate in 1971*	
Age group	*Vasectomy*	*Tubectomy*	*Intra-uterine device*	*Conventional contraceptives*	*Total*	*Vasectomy*	*Tubectomy*	*Intra-uterine device*	*Conventional contraceptives*	*Total*	*Method S_1* [a]	*Method S_2* [b]
15-19	2 066	78	497	1 419	4 140	640	31	210	549	1 430	185.0	185.0
20-24	11 335	2 227	4 855	7 547	25 964	3 589	897	2 096	2 824	9 406	254.7	254.2
25-29	36 251	10 727	15 024	12 994	74 996	10 768	4 055	6 085	4 561	25 469	227.4	225.7
30-34	69 447	18 369	25 522	8 148	121 486	14 303	4 815	7 168	1 983	28 269	135.4	134.2
35-39	78 669	17 530	22 251	3 323	121 773	11 469	3 253	4 423	573	19 718	84.0	84.4
40-44	66 218	10 320	9 458	788	86 784	3 391	673	660	48	4 772	25.1	25.7
45-49	32 045	4 203	2 316	207	38 771	870	145	86	7	1 108	12.7	13.0
TOTAL	296 031	63 454	79 923	34 506	473 914	45 030	13 869	20 728	10 545	90 172	4.62[c]	4.60[c]

	Number of births averted in 1971 because of change of marital fertility	*Percentage reduction in marital fertility because of family planning programme*
Method S_1 [a]	219 800	41.0
Method S_2 [b]	209 200	42.5

Source: Method S_2 based on procedure given in Ansley J. Coale, "Constructing the age distribution of a population recently subject to declining mortality", *Population Index*, vol. 37, No. 2 (April-June 1971), pp. 75-82.

[a] Moving-average method.
[b] Coale's method.
[c] Number of children per woman.

TABLE 12. PRE-ACCEPTANCE AND POST-ACCEPTANCE FERTILITY OVER A THREE-YEAR PERIOD

Family planning method	Pre-acceptance fertility: Acceptors Number	Pre-acceptance fertility: Acceptors Average	Pre-acceptance fertility: Matches Number	Pre-acceptance fertility: Matches Average	Ratio of acceptor's fertility to that of the matches	Post-acceptance fertility: Acceptors Number	Post-acceptance fertility: Acceptors Average	Post-acceptance fertility: Matches Number	Post-acceptance fertility: Matches Average	Ratio of acceptor's fertility to that of the matches	Percentage decline in fertility: Acceptors	Percentage decline in fertility: Matches
Vasectomy	6	500.00	18	388.89	1.29	6	0	18	277.78	0	100	28.57
Tubectomy	7	523.81	21	333.33	1.57	7	0	21	269.84	0	100	19.05
Intra-uterine device	6	444.44	18	277.78	1.60	6	111.19	18	203.70	0.55	75	26.67
Nirodh (condom)	16	458.33	48	368.06	1.25	16	187.50	48	263.89	0.71	39.09	28.30
TOTAL	35	476.18	105	349.20	1.36	35	104.78	105	257.14	0.41	63.30	26.36

Source: Partial data from the Longitudinal Sample Survey being conducted by the Population Centre, Bangalore, India.

acceptance fertility of acceptors, for all the methods combined, was 36 per cent higher than the non-acceptors; and the post-acceptance fertility of the same group over a three-year period following the acceptance was 59 per cent lower than that of the matches. It is also interesting to note that the fertility of matches, that is, those who reported not to have used any family planning method, has itself declined by 26 per cent over a three-year period. This may be partially because of the selection of cases, both acceptors and matches, on the basis of a recent birth within the last one or two years; as a consequence, the probability of their giving birth to a child in the next two- to three-year period can be expected to be lower even without any contraception. However, in the absence of any data based on a larger sample for the state, these figures were used in the computation of births averted by the component projection method described in the earlier sections. It is also to be noted that the post-acceptance fertility of acceptors and matches given in the table refer mostly to a period after 1971 and, consequently, it can be surmised that the fertility decline after 1971 has been greater than that estimated before 1971.

Simulation models

On assumptions concerning fertility, mortality, foetal loss and post-partum amenorrhoea that are very similar to the conditions prevailing in Karnataka State, Venkatacharya[17] has computed, using a micro-simulation model, the annual birth probabilities of an acceptor of a family planning method in successive years after acceptance for vasectomy, tubectomy and IUD, specified by the age group at acceptance. These birth probabilities have been computed taking into account specifically the susceptible state of the wife at the time of acceptance. For example, all the tubectomies in the state have been performed on post-partum women immediately after delivery; and with a long period (about a year) of amenorrhoea following a live birth, practically no birth would be prevented in the next one-year period. On the other hand, since the IUDs are inserted after making sure that the woman had resumed menstruation after the last childbirth, the fecundability of the woman is higher and the births prevented in the next one-year period for an IUD are much higher. Since the state of the woman at the time of the vasectomy of her husband is not known from the records, it was assumed that the birth prevention probabilities of a vasectomy case is an average of that of an IUD and a tubectomy case. The probabilities for IUD and tubectomy as worked out by Venkatacharya are given in table 13.

Using the birth probabilities and the number of acceptors of tubectomy, IUD and vasectomy reported in Karnataka State from 1961 onward, and the age distributions of these acceptors as given in table 7, the births averted per annum by each of these methods during 1961–1971 are given in table 14.

It is interesting to compare the results obtained by the application of simulation procedures with those obtained by the couple-years of protection given in table 10. The births averted by IUDs by applying the simulation birth probabilities from 1965–1971 work out to 116,179, while by couple-years-of-protection concept given in table 10 they work out to 109,447 (S_1 distribution). For tubectomies, the figures are 22,783 and 50,714, respectively. The reason that the simulation procedure has underestimated the births averted in this case is that it takes into account the post-partum amenorrhoea after tubectomy. For vasectomy, the figures for the two methods are again close, being 195,396 according to simulation procedure and 213,221 according to couple-years of protection.

SUMMARY AND CONCLUSIONS

This report presents the results of application of six methods of measuring the impact of family planning programmes on fertility in Karnataka State, India. The methods applied were standardization approach; trend analysis (fertility projection approach); couple-years of protection; component projection approach; experimental designs (matching studies); and simulation models. The methods developed by Potter[18] and Wol-

[17] K. Venkatacharya, "A computer model to estimate births averted due to IUCD and sterilizations", Bombay, International Institute of Population Studies, 1970 (mimeographed).

[18] Robert G. Potter, "Estimating births averted in a family planning program", in S. J. Behrman, Leslie Corsa, Jr., and Ronald

TABLE 13. PROBABILITY OF A WOMAN'S GIVING BIRTH (NOT USING ANY FAMILY PLANNING METHOD) WITH INITIAL SUSCEPTIBLE STATUS OF AN INTRA-UTERINE DEVICE USER OR SALPINGECTOMIZED FEMALE IN EACH CALENDAR YEAR SINCE THE INITIAL POINT OF TIME

Age group	Age-specific marital fertility rate	Birth probabilities by matrix: Intra-uterine device 1	2	3	4	5	Salpingectomy 1	2	3	4	5
15-19	0.184	0.026	0.299	0.278	0.238	0.268	0.000	0.019	0.171	0.324	0.274
20-24	0.302	0.049	0.471	0.314	0.247	0.301	0.000	0.033	0.259	0.400	0.278
25-29	0.323	0.057	0.544	0.289	0.243	0.360	0.000	0.038	0.298	0.421	0.261
30-34	0.257	0.032	0.390	0.265	0.198	0.277	0.000	0.024	0.206	0.332	0.226
35-39	0.155	0.012	0.182	0.181	0.143	0.152	0.000	0.010	0.095	0.186	0.160
40-44	0.075	0.004	0.070	0.059	0.054	0.053	0.000	0.004	0.036	0.078	0.077

Source: K. Venkatacharya, "A computer model to estimate births averted due to IUCD and Sterilization", Bombay, International Institute of Population Studies, 1970 (mimeographed).

fers[19] for estimating births averted, which relate the extent to which the fecundable state of a woman is lengthened because of contraception to the birth intervals, could not be applied because of paucity of necessary data. Also, it was not possible to apply any regression analysis on the basis of areal data in the state, in view of the fact that information on fertility levels and trends for administrative units within the state, necessary for such analysis, was not available.

The results derived from the application of the six methods relate to the period 1961–1971 in Karnataka State. This period was chosen because the two most recent censuses were conducted in 1961 and 1971, and could provide the necessary basic demographic and socio-economic data for the population of the state, and also because the official family planning programme in the state had received a considerable impetus in terms of increased financial and organizational inputs during that period. The data necessary for the application of the various evaluation methods were derived from these two censuses, from National Sample Surveys conducted in 1958–1959 and 1960–1961, from the Sample Registration Scheme figures on fertility and mortality available from 1968 onward; and from the family planning service statistics system, which provided data on the number of acceptors of various family planning methods and the age distribution of the acceptors. In addition, data on certain essential parameters, such as continuation rates of methods, were borrowed from a few special studies conducted in other states in India.

Though the censuses did provide data on various aspects of the population of the state in 1961 and 1971, certain essential particulars, such as the age distribution recorded, were of poor quality and revealed considerable response errors. Smoothing of the age distribution of the population was a prerequisite of any further analysis. It was assumed, *a priori*, that the procedures adopted for smoothing might, in themselves, have effects on the results of programme impact; and, consequently, two different smoothing procedures were adopted in order to determine the extent of variation caused by such procedure on the final results. The two smoothing procedures were the moving-average method (S_1) and Coale's method (S_2). Analysis of the results obtained by the application of various methods of programme impact revealed that the smoothing procedure does have a significant effect on the results of the impact of the programme, especially at the first stage where the total birth-rate change has to be partitioned into two parts: one due to the structural changes in the population, such as changes in age-sex-marital status; and the second part due to changes in marital fertility.

On the other hand, smoothing procedures do not appear to influence to the same extent the results with regard to the proportion of changes in marital fertility attributable to the family planning programme. The other important observations that emanate from a comparative analysis of the results obtained from the

Freedman, eds., *Fertility and Family Planning: A World View* (Ann Arbor, Mich., University of Michigan Press, 1969), pp. 413–434.

[19] David Wolfers, "The demographic effects of a contraceptive programme", *Population Studies*, vol. XXIII, No. 1 (March 1969), pp. 111–141.

TABLE 14. APPLICATION OF THE RESULTS OF COMPUTER SIMULATION

Family planning method	Births averted: 1961	1962	1963	1964	1965	1966	1967	1968	1969	1970	1971	Total
Vasectomy	38	476	1 248	2 688	4 699	7 832	14 972	29 045	42 002	46 999	45 397	195 396
Tubectomy	0	25	221	571	929	1 417	1 914	2 227	2 925	4 623	7 931	22 783
Intra-uterine device.....	...	...	...	...	851	14 646	26 827	25 332	20 192	16 062	12 269	116 179
TOTAL	38	501	1 469	3 259	6 479	23 895	43 713	56 604	65 119	67 684	65 597	334 358

Note: The number of births averted has been calculated from the data on acceptors in 1961 and thereafter, whereas in table 10, acceptor data for 1956 and subsequent years have been used.

application of the various methods are as follows. Unless otherwise specified, the findings given below are generally based on populations smoothed by the moving-average method (S_1):

1. The birth rate in the state has declined by about 7.6 points during the 10-year period 1961–1971; about two thirds of this decline can be attributed to changes in marital fertility;

2. The births averted because of the fertility decline during 1961–1971 constituted 11 per cent of the total births that occurred during the same period;

3. About 37 per cent of the births that were averted during 1961–1971 because of the fertility decline can be attributed to the acceptance and use of various methods of family planning from within the programme in the state prior to 1971;

4. Forty-one per cent of the reduction in marital fertility in the year 1971 can be attributed to family planning acceptors from within the programme prior to that year;

5. There is a considerable decline in the post-acceptance fertility of acceptors of different methods of family planning compared with their pre-acceptance fertility levels. On the other hand, even among the control group, there is a significant decline in fertility over the same period of time;

6. The analysis reveals that both programme factors and non-programme factors have played a significant role in the decline of fertility in the state during 1961–1971, the indirect effects of the programme and non-programme effects accounting for almost two thirds of the decline in marital fertility.

7. The application of stable population theory to smooth the age distribution or for estimation of fertility and mortality levels has serious limitations in the Indian situation after 1961. Further research in this aspect appears very necessary.

Annex

KARNATAKA STATE: SMOOTHED POPULATION DISTRIBUTIONS, 1961 AND 1971

TABLE 15. POPULATION DISTRIBUTION SMOOTHED BY MOVING-AVERAGE METHOD, 1961

(Hundreds)

	Rural population														
	Male					*Female*					*Total*				
Age group	*Single*	*Married*	*Widowed*	*Divorced or separated*	*Total*	*Single*	*Married*	*Widowed*	*Divorced or separated*	*Total*	*Single*	*Married*	*Widowed*	*Divorced or separated*	*Total*
0-9	27 541				27 541	27 419				27 419	54 960				54 960
10-14	9 920	157	4	1	10 082	9 301	1 348	15	5	10 669	19 221	1 505	19	6	20 751
15-19	7 875	735	12	2	8 624	2 547	6 391	76	43	9 057	10 422	7 126	88	45	17 681
20-24	4 616	2 944	45	15	7 620	331	7 331	157	66	7 885	4 947	10 275	202	81	15 505
25-29	1 651	5 377	88	33	7 149	95	6 410	295	70	6 870	1 746	11 787	383	103	14 019
30-34	454	5 741	135	39	6 369	62	5 205	563	70	5 900	516	10 946	698	109	12 269
35-39	202	5 250	170	37	5 659	42	4 015	797	58	4 912	244	9 265	967	95	10 571
40-44	132	4 449	271	36	4 888	41	2 834	1 180	48	4 103	173	7 283	1 451	84	8 991
45-49	92	3 700	323	29	4 144	27	2 025	1 323	30	3 405	119	5 725	1 646	59	7 549
50-54	69	2 823	401	24	3 317	23	1 214	1 565	18	2 820	92	4 037	1 966	42	6 137
55-59	54	2 121	394	16	2 585	18	804	1 459	12	2 293	72	2 925	1 853	28	4 878
60-64	32	1 456	411	11	1 910	12	398	1 466	7	1 883	44	1 854	1 877	18	3 793
65-69	27	924	324	8	1 283	10	240	1 162	4	1 416	37	1 164	1 486	12	2 699
70+	29	1 083	584	10	1 706	12	154	1 524	3	1 693	41	1 237	2 108	13	3 399
TOTAL	52 694	36 760	3 162	261	92 877	39 940	38 369	11 582	434	90 325	92 634	75 129	14 744	695	183 202
	Urban population														
0-9	7 532				7 532	7 365				7 365	14 897				14 897
10-14	2 989	15	–	–	3 004	2 806	154	2	1	2 963	5 795	169	2	1	5 967
15-19	2 974	101	2	1	3 078	1 402	1 554	13	9	2 978	4 376	1 655	15	10	6 056
20-24	2 110	696	8	3	2 817	317	2 100	36	15	2 468	2 427	2 796	44	18	5 285
25-29	726	1 493	19	5	2 243	89	1 835	63	15	2 002	815	3 328	82	20	4 245
30-34	195	1 751	31	6	1 983	33	1 460	121	16	1 630	228	3 211	152	22	3 613
35-39	69	1 585	35	5	1 694	22	1 106	173	12	1 313	91	2 691	208	17	3 007
40-44	39	1 343	57	5	1 444	15	752	272	9	1 048	54	2 095	329	14	2 492
45-49	25	1 028	61	3	1 117	11	533	326	7	877	36	1 561	387	10	1 994
50-54	17	766	80	4	867	8	312	375	5	700	25	1 078	455	9	1 567
55-59	12	539	75	2	628	6	210	354	2	572	18	749	429	4	1 200
60-64	7	358	87	1	453	10	104	350	2	466	17	462	437	3	919
65-69	5	218	62	1	286	8	65	255	2	330	13	283	317	3	616
70+	7	254	125	1	387	10	39	370	1	420	17	293	495	2	807
TOTAL	16 707	10 147	642	37	27 533	12 102	10 224	2 710	96	25 132	28 809	20 371	3 352	133	52 665

TABLE 15. (*continued*)

	Combined population														
	Male					*Female*					*Total*				
Age group	*Single*	*Married*	*Wid-owed*	*Divorced or sepa-rated*	*Total*	*Single*	*Married*	*Wid-owed*	*Divorced or sepa-rated*	*Total*	*Single*	*Married*	*Wid-owed*	*Divorced or sepa-rated*	*Total*
0-9	35 073				35 073	34 784				34 784	69 857				69 857
10-14	12 909	172	4	1	13 086	12 107	1 502	17	6	13 632	25 016	1 674	21	7	26 718
15-19	10 849	836	14	3	11 702	3 949	7 945	89	52	12 035	14 798	8 781	103	55	23 737
20-24	6 726	3 640	53	18	10 437	648	9 431	193	81	10 353	7 374	13 071	246	99	20 790
25-29	2 377	6 870	107	38	9 392	184	8 245	358	85	8 872	2 561	15 115	465	123	18 264
30-34	649	7 492	166	45	8 352	95	6 665	684	86	7 530	744	14 157	850	131	15 882
35-39	271	6 835	205	42	7 353	64	5 121	970	70	6 225	335	11 956	1 175	112	13 578
40-44	171	5 792	328	41	6 332	56	3 586	1 452	57	5 151	227	9 378	1 780	98	11 483
45-49	117	4 728	384	32	5 261	38	2 558	1 649	37	4 282	155	7 286	2 033	69	9 543
50-54	86	3 589	481	28	4 184	31	1 526	1 940	23	3 520	117	5 115	2 421	51	7 704
55-59	66	2 660	469	18	3 213	24	1 014	1 813	14	2 865	90	3 674	2 282	32	6 078
60-64	39	1 814	498	12	2 363	22	502	1 816	9	2 349	61	2 316	2 314	21	4 712
65-69	32	1 142	386	9	1 569	18	305	1 417	6	1 746	50	1 447	1 803	15	3 315
70+	36	1 337	709	11	2 093	22	193	1 894	4	2 113	58	1 530	2 603	15	4 206
TOTAL	69 401	46 907	3 804	298	120 410	52 042	48 593	14 292	530	115 457	121 443	95 500	18 096	828	235 867

TABLE 16. POPULATION DISTRIBUTION SMOOTHED BY COALE'S METHOD, 1961

(*Hundreds*)

	Rural Population														
	Male					*Female*					*Total*				
Age group	*Single*	*Married*	*Wid-owed*	*Divorced or sepa-rated*	*Total*	*Single*	*Married*	*Wid-owed*	*Divorced or sepa-rated*	*Total*	*Single*	*Married*	*Wid-owed*	*Divorced or sepa-rated*	*Total*
0-9	27 259				27 259	25 957				25 957	53 216				53 216
10-14	10 171	163	4	1	10 339	10 433	1 288	12	4	11 737	20 604	1 451	16	5	22 076
15-19	7 948	755	12	2	8 717	2 757	5 895	58	39	8 749	10 705	6 650	70	41	17 466
20-24	4 624	3 003	45	15	7 687	361	6 822	122	60	7 365	4 985	9 825	167	75	15 052
25-29	1 632	5 410	86	33	7 161	106	6 102	234	65	6 507	1 738	11 512	320	98	13 668
30-34	439	5 664	129	39	6 271	72	5 095	459	67	5 693	511	10 759	588	106	11 964
35-39	191	5 059	158	36	5 444	51	4 134	684	58	4 927	242	9 193	842	94	10 371
40-44	124	4 246	250	34	4 654	52	3 101	1 076	51	4 280	176	7 347	1 326	85	8 934
45-49	87	3 583	301	28	3 999	37	2 361	1 285	35	3 718	124	5 944	1 586	63	7 717
50-54	69	2 831	388	24	3 312	34	1 512	1 624	23	3 193	103	4 343	2 012	47	6 505
55-59	56	2 223	398	17	2 694	27	1 041	1 575	15	2 658	83	3 264	1 973	32	5 352
60-64	35	1 603	436	12	2 086	19	512	1 571	9	2 111	54	2 115	2 007	21	4 197
65-69	30	1 093	369	9	1 501	16	307	1 239	5	1 567	46	1 400	1 608	14	3 068
70+	29	1 127	586	11	1 753	18	199	1 643	3	1 863	47	1 326	2 229	14	3 616
TOTAL	52 694	36 760	3 162	261	92 877	39 940	38 369	11 582	434	90 325	92 634	75 129	14 744	695	183 202
	Urban population														
0-9	7 456				7 456	6 972				6 972	14 428				14 428
10-14	3 062	16	—	—	3 078	3 063	147	1	1	3 212	6 125	163	1	1	6 290
15-19	2 998	104	2	1	3 105	1 477	1 434	10	8	2 929	4 475	1 538	12	9	6 034
20-24	2 112	713	8	3	2 836	337	1 955	28	14	2 334	2 449	2 668	36	17	5 170
25-29	716	1 507	19	5	2 247	96	1 748	50	14	1 908	812	3 255	69	19	4 155
30-34	188	1 733	29	6	1 956	36	1 429	99	15	1 579	224	3 162	128	21	3 535
35-39	64	1 533	33	5	1 635	26	1 139	148	12	1 325	90	2 672	181	17	2 960
40-44	37	1 285	52	5	1 379	18	824	247	10	1 099	55	2 109	299	15	2 478
45-49	24	998	57	3	1 082	15	621	316	8	960	39	1 619	373	11	2 042
50-54	17	771	77	4	869	11	388	389	6	794	28	1 159	466	10	1 663
55-59	12	567	75	2	656	10	272	380	3	665	22	839	455	5	1 321
60-64	8	396	93	1	498	14	133	374	2	523	22	529	467	3	1 021
65-69	6	259	71	1	337	12	84	271	2	369	18	343	342	3	706
70+	7	265	126	1	399	15	50	397	1	463	22	315	523	2	862
TOTAL	16 707	10 147	642	37	27 533	12 102	10 224	2 710	96	25 132	28 809	20 371	3 352	133	52 665

TABLE 16. (*continued*)

	Combined population														
	Male					Female					Total				
Age group	Single	Married	Wid-owed	Divorced or sepa-rated	Total	Single	Married	Wid-owed	Divorced or sepa-rated	Total	Single	Married	Wid-owed	Divorced or sepa-rated	Total
0-9	34 715				34 715	32 929				32 929	67 644				67 644
10-14	13 233	179	4	1	13 417	13 496	1 435	13	5	14 949	26 729	1 614	17	6	28 366
15-19	10 946	859	14	3	11 822	4 234	7 329	68	47	11 678	15 180	8 188	82	50	23 500
20-24	6 736	3 716	53	18	10 523	698	8 777	150	74	9 699	7 434	12 493	203	92	20 222
25-29	2 348	6 917	105	38	9 408	202	7 850	284	79	8 415	2 550	14 767	389	117	17 823
30-34	627	7 397	158	45	8 227	108	6 524	558	82	7 272	735	13 921	716	127	15 499
35-39	255	6 592	191	41	7 079	77	5 273	832	70	6 252	332	11 865	1 023	111	13 331
40-44	161	5 531	302	39	6 033	70	3 925	1 323	61	5 379	231	9 456	1 625	100	11 412
45-49	111	4 581	358	31	5 081	52	2 982	1 601	43	4 678	163	7 563	1 959	74	9 759
50-54	86	3 602	465	28	4 181	45	1 900	2 013	29	3 987	131	5 502	2 478	57	8 168
55-59	68	2 790	473	19	3 350	37	1 313	1 955	18	3 323	105	4 103	2 428	37	6 673
60-64	43	1 999	529	13	2 584	33	645	1 945	11	2 634	76	2 644	2 474	24	5 218
65-69	36	1 352	440	10	1 838	28	391	1 510	7	1 936	64	1 743	1 950	17	3 774
70+	36	1 392	712	12	2 152	33	249	2 040	4	2 326	69	1 641	2 752	16	4 478
TOTAL	69 401	46 907	3 804	298	120 410	52 042	48 593	14 292	530	115 457	121 443	95 500	18 096	828	235 867

Source: Smoothing method based on procedure given in Ansley J. Coale, "Constructing the age distribution of a population recently subject to declining mortality", *Population Index*, vol. 37, No. 2 (April-June 1971), pp. 75-82.

Note: for the purpose of smoothing, male e_o^o was assumed to be 47.5; female e_o^o, 45.0; gross reproduction rate, 2.8; and sex ratio at birth, 107.

TABLE 17. POPULATION DISTRIBUTION SMOOTHED BY MOVING-AVERAGE METHOD, 1971
(*Hundreds*)

	Male					Female					Total				
Age group	Single	Married	Wid-owed	Divorced or sepa-rated	Total	Single	Married	Wid-owed	Divorced or sepa-rated	Total	Single	Married	Wid-owed	Divorced or sepa-rated	Total
Rural population															
0-9	32 809				32 809	32 511				32 511	65 320				65 320
10-14	13 872	125	3	—	14 000	13 001	984	11	3	13 999	26 873	1 109	14	3	27 999
15-19	10 292	613	10	1	10 916	5 016	5 611	89	28	10 744	15 308	6 224	99	29	21 660
20-24	5 795	2 867	29	10	8 701	755	8 071	163	58	9 047	6 550	10 938	192	68	17 748
25-29	1 889	5 935	64	29	7 917	151	7 607	245	69	8 072	2 040	13 542	309	98	15 989
30-34	495	6 458	99	35	7 087	74	6 497	480	70	7 121	569	12 955	579	105	14 208
35-39	192	6 166	132	36	6 526	47	5 270	620	62	5 999	239	11 436	752	98	12 525
40-44	147	5 413	199	32	5 791	36	4 015	1 023	50	5 124	183	9 428	1 222	82	10 915
45-49	87	4 558	247	31	4 923	26	2 940	1 195	32	4 193	113	7 498	1 442	63	9 116
50-54	61	3 647	337	27	4 072	20	1 830	1 594	24	3 468	81	5 477	1 931	51	7 540
55-59	37	2 846	326	22	3 231	25	1 241	1 490	22	2 778	62	4 087	1 816	44	6 009
60-64	30	2 047	372	13	2 462	11	632	1 586	10	2 239	41	2 679	1 958	23	4 701
65-69	19	1 378	310	8	1 715	8	372	1 214	6	1 600	27	1 750	1 524	14	3 315
70+	28	1 639	661	14	2 342	12	297	2 068	5	2 382	40	1 936	2 729	19	4 724
TOTAL	65 753	43 692	2 789	258	112 492	51 693	45 367	11 778	439	109 277	117 446	89 059	14 567	697	221 769
Urban population															
0-9	9 683				9 683	9 507				9 507	19 190				19 190
10-14	4 329	14	—	—	4 343	3 950	113	1	—	4 064	8 279	127	1	—	8 407
15-19	4 254	97	1	1	4 353	2 657	1 419	9	6	4 091	6 911	1 516	10	7	8 444
20-24	3 231	714	6	2	3 953	760	2 683	32	14	3 489	3 991	3 397	38	16	7 442
25-29	1 136	1 851	14	5	3 006	146	2 525	59	14	2 744	1 282	4 376	73	19	5 750
30-34	282	2 296	27	5	2 610	58	2 030	111	14	2 213	340	4 326	138	19	4 823
35-39	95	2 085	30	5	2 215	30	1 675	162	13	1 880	125	3 760	192	18	4 095
40-44	55	1 783	47	5	1 890	19	1 146	271	13	1 449	74	2 929	318	18	3 339
45-49	35	1 458	54	3	1 550	13	851	330	10	1 204	48	2 309	384	13	2 754
50-54	23	1 066	72	3	1 164	10	515	420	6	951	33	1 581	492	9	2 115
55-59	16	778	71	2	867	7	360	416	3	786	23	1 138	487	5	1 653
60-64	10	531	81	2	624	5	167	410	2	584	15	698	491	4	1 208
65-69	8	340	68	1	417	3	102	314	1	420	11	442	382	2	837
70+	9	397	145	1	552	4	81	526	1	612	13	478	671	2	1 164
TOTAL	23 166	13 410	616	35	37 227	17 169	13 667	3 061	97	33 994	40 335	27 077	3 677	132	71 221

TABLE 17. (*continued*)

Age group	Combined population														
	Male					Female					Total				
	Single	Married	Wid-owed	Divorced or sepa-rated	Total	Single	Married	Wid-owed	Divorced or sepa-rated	Total	Single	Married	Wid-owed	Divorced or sepa-rated	Total
0-9	42 492				42 492	42 018				42 018	84 510				84 510
10-14	18 201	139	3	—	18 343	16 951	1 097	12	3	18 063	35 152	1 236	15	3	36 406
15-19	14 546	710	11	2	15 269	7 673	7 030	98	34	14 835	22 219	7 740	109	36	30 104
20-24	9 026	3 581	35	12	12 654	1 515	10 754	195	72	12 536	10 541	14 335	230	84	25 190
25-29	3 025	7 786	78	34	10 923	297	10 132	304	83	10 816	3 322	17 918	382	117	21 739
30-34	777	8 754	126	40	9 697	132	8 527	591	84	9 334	909	17 281	717	124	19 031
35-39	287	8 251	162	41	8 741	77	6 945	782	75	7 879	364	15 196	944	116	16 620
40-44	202	7 196	246	37	7 681	55	5 161	1 294	63	6 573	257	12 357	1 540	100	14 254
45-49	122	6 016	301	34	6 473	39	3 791	1 525	42	5 397	161	9 807	1 826	76	11 870
50-54	84	4 713	409	30	5 236	30	2 345	2 014	30	4 419	114	7 058	2 423	60	9 655
55-59	53	3 624	397	24	4 098	32	1 601	1 906	25	3 564	85	5 225	2 303	49	7 662
60-64	40	2 578	453	15	3 086	16	799	1 996	12	2 823	56	3 377	2 449	27	5 909
65-69	27	1 718	378	9	2 132	11	474	1 528	7	2 020	38	2 192	1 906	16	4 152
70+	37	2 036	806	15	2 894	16	378	2 594	6	2 994	53	2 414	3 400	21	5 888
TOTAL	88 919	57 102	3 405	293	149 719	68 862	59 034	14 839	536	143 271	157 781	116 136	18 244	829	292 990

TABLE 18. POPULATION DISTRIBUTION SMOOTHED BY COALE'S METHOD, 1971
(*Hundreds*)

Age group	Male					Female					Total				
	Single	Married	Wid-owed	Divorced or sepa-rated	Total	Single	Married	Wid-owed	Divorced or sepa-rated	Total	Single	Married	Wid-owed	Divorced or sepa-rated	Total
Rural population															
0-9	30 346				30 346	29 565				29 565	59 911				59 911
10-14	13 967	97	2	—	14 066	14 770	850	8	2	15 630	28 737	947	10	2	29 696
15-19	11 060	508	7	1	11 576	5 889	5 012	62	25	10 988	16 949	5 520	69	26	22 564
20-24	6 703	2 554	23	9	9 289	898	7 307	115	51	8 371	7 601	9 861	138	60	17 660
25-29	2 285	5 529	52	27	7 893	185	7 049	177	62	7 473	2 470	12 578	229	89	15 366
30-34	606	6 093	82	33	6 814	93	6 216	358	65	6 732	699	12 309	440	98	13 546
35-39	233	5 781	109	33	6 156	63	5 326	489	60	5 938	296	11 107	598	93	12 094
40-44	179	5 100	164	29	5 472	52	4 324	860	52	5 288	231	9 424	1 024	81	10 760
45-49	110	4 466	212	30	4 818	40	3 409	1 081	36	4 566	150	7 875	1 293	66	9 384
50-54	83	3 828	311	28	4 250	32	2 292	1 559	29	3 912	115	6 120	1 870	57	8 162
55-59	55	3 260	327	24	3 666	45	1 679	1 574	28	3 326	100	4 939	1 901	52	6 992
60-64	48	2 550	406	15	3 019	22	910	1 782	14	2 728	70	3 460	2 188	29	5 747
65-69	33	1 874	369	11	2 287	16	583	1 485	9	2 093	49	2 457	1 854	20	4 380
70+	45	2 052	725	18	2 840	23	410	2 228	6	2 667	68	2 462	2 953	24	5 507
TOTAL	65 753	43 692	2 789	258	112 492	51 693	45 367	11 778	439	109 277	117 446	89 059	14 567	697	221 769
Urban population															
0-9	8 956				8 956	8 646				8 646	17 602				17 602
10-14	4 208	11	—	—	4 219	4 295	98	1	—	4 394	8 503	109	1	—	8 613
15-19	4 414	81	1	—	4 496	2 987	1 271	7	5	4 270	7 401	1 352	8	5	8 766
20-24	3 608	641	4	2	4 255	866	2 435	22	12	3 335	4 474	3 076	26	14	7 590
25-29	1 327	1 737	12	4	3 080	170	2 345	43	13	2 571	1 497	4 082	55	17	5 651
30-34	334	2 182	22	5	2 543	69	1 947	83	13	2 112	403	4 129	105	18	4 655
35-39	112	1 969	25	5	2 111	39	1 696	127	13	1 875	151	3 665	152	18	3 986
40-44	66	1 692	39	4	1 801	26	1 237	227	13	1 503	92	2 929	266	17	3 304
45-49	45	1 439	47	3	1 534	19	989	298	11	1 317	64	2 428	345	14	2 851
50-54	30	1 127	66	3	1 226	16	647	409	7	1 079	46	1 774	475	10	2 305
55-59	23	897	72	3	995	12	489	438	4	943	35	1 386	510	7	1 938
60-64	16	666	88	2	772	9	241	459	3	712	25	907	547	5	1 484
65-69	13	467	81	2	563	7	160	382	2	551	20	627	463	4	1 114
70+	14	501	159	2	676	8	112	565	1	686	22	613	724	3	1 362
TOTAL	23 166	13 410	616	35	37 227	17 169	13 667	3 061	97	33 994	40 335	27 077	3 677	132	71 221

TABLE 18. (*continued*)

Age group	Combined population														
	Male					Female					Total				
	Single	Married	Wid-owed	Divorced or sepa-rated	Total	Single	Married	Wid-owed	Divorced or sepa-rated	Total	Single	Married	Wid-owed	Divorced or sepa-rated	Total
0-9	39 302				39 302	38 211				38 211	77 513				77 513
10-14	18 175	108	2	—	18 285	19 065	948	9	2	20 024	37 240	1 056	11	2	38 309
15-19	15 474	589	8	1	16 072	8 876	6 283	69	30	15 258	24 350	6 872	77	31	31 330
20-24	10 311	3 195	27	11	13 544	1 764	9 742	137	63	11 706	12 075	12 937	164	74	25 250
25-29	3 612	7 266	64	31	10 973	355	9 394	220	75	10 044	3 967	16 660	284	106	21 017
30-34	940	8 275	104	38	9 357	162	8 163	441	78	8 844	1 102	16 438	545	116	18 201
35-39	345	7 750	134	38	8 267	102	7 022	616	73	7 813	447	14 772	750	111	16 080
40-44	245	6 792	203	33	7 273	78	5 561	1 087	65	6 791	323	12 353	1 290	98	14 064
45-49	155	5 905	259	33	6 352	59	4 398	1 379	47	5 883	214	10 303	1 638	80	12 235
50-54	113	4 955	377	31	5 476	48	2 939	1 968	36	4 991	161	7 894	2 345	67	10 467
55-59	78	4 157	399	27	4 661	57	2 168	2 012	32	4 269	135	6 325	2 411	59	8 930
60-64	64	3 216	494	17	3 791	31	1 151	2 241	17	3 440	95	4 367	2 735	34	7 231
65-69	46	2 341	450	13	2 850	23	743	1 867	11	2 644	69	3 084	2 317	24	5 494
70+	59	2 553	884	20	3 516	31	522	2 793	7	3 353	90	3 075	3 677	27	6 869
TOTAL	88 919	57 102	3 405	293	149 719	68 862	59 034	14 839	536	143 271	157 781	116 136	18 244	829	292 990

Source: Smoothing method based on procedure given in Ansley J. Coale, "Constructing the age distribution of a population recently subject to declining mortality", *Population Index*, vol. 37, No. 2 (April-June 1971), pp. 75-82.

Note: For the purpose of smoothing, male $\mathring{e}_0$ was assumed to be 50.0; female $\mathring{e}_0$, 48.0; gross reproduction rate, 2.5; and sex ratio at birth, 107.

APPLICATION OF METHODS OF MEASURING THE IMPACT OF FAMILY PLANNING PROGRAMMES ON FERTILITY: THE CASE OF TUNISIA*

*Yolande Jemai*** and Hedi Jemai****

The trend towards a decline in fertility among Tunisian women appears to have been confirmed during the past 10 years and therefore to indicate a definite change in the reproductive behaviour of couples, but it is still difficult to isolate the economic and social factors that have resulted in the replacement in the Moslem Arab society of the traditional attitude that favoured large families by a voluntary joint choice of family planning by both spouses.

In order to make a useful contribution to a knowledge of the mechanics of this transition, an effort must be made to identify the relative impact of spontaneous factors and deliberately induced factors on the fall in fertility, if one is to assess the impact of the policy adopted by the Tunisian authorities within the framework of the over-all economic and social development policy of the country.

A number of methods have been proposed for measuring the impact of a family planning programme on fertility, and the purpose here is to test them in the specific case of Tunisia and to try to identify their defects and the methodological difficulties involved in applying them.

In view of the need for uniformity in making any comparison of methods, this assessment has been made for the period 1966–1971 and for Tunisia as a whole. It should be pointed out, however, that this period has been selected for reasons of availability of demographic data, but that it is not the most favourable period for an assessment of the success of the Tunisian programme, as all evidence indicates that awareness of the family planning programme among the population in general has occurred chiefly since 1973, when the National Family Planning and Population Office (ONPFP) was established.

This analysis, which is guided largely by the methodological aspects of the question, essentially takes the form of applying simultaneously over space and time five methods of measuring the impact of family planning programmes on fertility: standardization approach; trend analysis (fertility projection approach); experimental designs; couple-years of protection (CYP); component projection approach. The present report does not cover analysis of the reproductive process and regression analysis.

Before describing the results of the application of these methods and the criticisms suggested by them, however, a brief summary is given of past and recent trends in fertility in Tunisia (some characteristic indicators for Tunisia are also given in annex I).

Past and recent fertility trends in Tunisia

Past trends

Prior to introduction of the national family planning programme in 1964

According to a number of estimates,[1] the population of Tunisia was rising slightly prior to 1921, after which date the first censuses of the native population were carried out. It is generally believed that from 1860 to 1921, the average annual rate of growth was no more than 1 per cent because of high mortality, due to epidemics and deplorable economic and social conditions, which masked the effects of the high birth rate which is traditional among Moslem Arab societies.

From 1921 to 1946, that trend continued; and it was not until after the Second World War that the annual rate of growth approached 2 per cent and continued to rise steadily until 1966, when it reached 2.8 per cent, the highest level ever recorded (see table1).

The progress achieved in the health field since independence is the main cause of the sudden fall in general mortality (27–28 per 1,000 in 1945; 9.5 per 1,000 in 1974),[2] which thus revealed the effects of natality. The high fertility level was, and still is, closely connected with a social and cultural context which requires, in Arab countries, that matrimonial customs conform to a number of rules codified by Islam: very young age at marriage; stability of unions; the prizing of male progeny; and natural behaviour which excludes con-

* The original version of this paper appeared as document ESA/P/AC.7/3.

** Office national du planning familial et de la population, Tunis.

*** Centre d'études de recherches économiques et sociales, Tunis University.

[1] See, in particular, Mahmoud Seklani, *La population de la Tunisie* (Paris, Comité international de coordination des recherches nationales de démographie, 1974). See also Mahmoud Seklani, "La population de la Tunisie, situation actuelle et évolution probable jusqu'en 1986", *Population* (Paris), vol. 16, No. 3 (July-September 1961), pp. 473–504.

[2] Office national du planning familial et de la population, *Statistiques de planning familial; indicateurs démographiques,* No. 6, second quarter (Tunis, 1975).

TABLE 1. POPULATION GROWTH IN TUNISIA, CENSUS YEARS, 1921-1975

Census year	*Crude data as of date of census*	*Corrected data*	*Average annual rate of growth (percentage)*
1921	2 093 939	2 000 000	1.2
1926	2 159 708	2 120 000	1.32
1931	2 410 692	2 260 000	1.46
1936	2 608 313	2 426 000	1.79
1946	3 230 952	2 891 000	2.0
1956	3 783 169	3 519 000	3.7
1966	4 533 351	4 589 000	2.8
1975	5 572 229	...	...

Sources: Tunisia, Institut national de la statistique; and for average annual rate of growth, Hedi Jemai, "Tunisie–Madagascar: démographie comparée", *Revue tunisienne de sciences sociales* (forthcoming).

traceptive practices. All these factors have contributed to causing the Tunisian birth rate, which was never among the highest in the world, to reach a dangerous level from the point of view of the requirements of balanced economic and social development.

According to official figures,[3] in 1960, the corrected crude birth rate was 45.7 per 1,000 (191,810 births); in 1964, just before the introduction of the official family planning programme, the crude birth rate reached its highest level, 46.2 per 1,000 (206,046 births) and the general fertility rate (for women aged 15–54 years) was about 207 per 1,000.

From the introduction of family planning until 1970

During the period from the introduction of family planning in 1964 up to 1970, the responsible authorities changed their position from an attitude favouring birth control to state direction in population matters.

Ever since independence, they had been aware of the problem and had taken a number of measures to alleviate it, either directly or indirectly:

(*a*) The Code of Personal Status, promulgated on 13 August 1956, granted women equality of civil status (regulations governing marriage and divorce);

(*b*) In 1961, the French Act of 1920, which prohibited the sale and distribution of contraceptives, was abrogated;

(*c*) In 1964, the minimum age for marriage was raised to 17 years for girls and 20 years for boys;

(*d*) In 1965, the first measure liberalizing abortion authorized the practice of social abortion on certain conditions (minimum of five children).

The decisive step in favour of population policy, however, was taken in 1964, with the initiation of the pilot family planning experiment following the encouraging results of a Knowledge/Attitude/Practice (KAP) survey of the population, which was found to be predominantly in favour of birth control. After two years of work and an analysis of the results, the success of the experiment opened the way for the national family planning programme.

The method that had been most prevalent until then was the intra-uterine device (IUD); however, following a speech by President Bourguiba in favour of population growth and propaganda against that method, there was a falling-off in the activities of the programme. This situation led in 1968 to organizational reforms: resumption of the work of the mobile teams; implementation of a post-partum and post-abortum programme; establishment of a Family Planning Department within the Ministry of Public Health; and most important, wide distribution of pills and condoms. Following these improvements, activities made progress until 1971. At the statistical level, the separate phases described above are illustrated by the data in table 2.

The official figures indicate that during the same period, the general fertility rate (15–54 years) fell from approximately 207 per 1,000 in 1964 to 193 per 1,000 in 1966 and 164 per 1,000 in 1970. According to the fourth economic and social development plan (1973–1976), this fall corresponds to an average annual decline of 5.1 births per 1,000 women of reproductive age, made up of: 1.8 births per 1,000 women of reproductive age averted by the effect of the age structure; and 3.3 births per 1,000 women of reproductive age averted by the specific effect of family planning.

A number of demographers have analysed the impact of various factors on the observed fertility decline between 1966 and 1970;[4] all reached the conclusion that, at most, one third of the decline was caused by family planning, and that age structure and matrimonial status together had an effect of the same magnitude. Accordingly, approximately one third of the fall is thought to have been due to the effect of eco-

[3] Data referred to in this paper as "official figures" were obtained, unless otherwise indicated, from Institut national de la statistique, Tunis.

[4] See, for example, Jacques Vallin, "Limitation des naissances en Tunisie: efforts et résultats", *Population* (Paris), vol. 26, special number (March 1971), pp. 181–204.

TABLE 2. RESULTS OF THE ACTIVITIES OF THE FAMILY PLANNING PROGRAMME, 1964-1970

Year	*Primary insertion of intra-uterine devices*	*New acceptors of the pill*	*Tubal ligation*	*Social abortion*	*New consultations*	*Total consultations*
1964	1 154	—	293	—	6 160	12 620
1966	12 077	350	766	1 396	16 176	41 517
1968	9 304	4 780	1 627	2 246	20 432	67 986
1970	9 638	9 959	2 531	2 705	35 362	184 419

Source: Tunisia, Office national du planning familial et de la population.

nomic and social development, the growth of internal and external migration after 1966, increased school attendance of girls (58.49 per cent in 1971/72) and awareness that the risk of infant mortality had fallen.

Recent trends in fertility: 1970–1975

The trend that had emerged in the preceding period was strengthened, although the decline in fertility might well have been considered a purely circumstantial phenomenon, given the characteristics of the age structure of the population and the reduced impact of the 1964 Act on the age of marriage.

For the period 1964–1974 as a whole, that is to say, after 10 years of the family planning programme, the estimated birth rates and fertility rates were as shown in table 3.

TABLE 3. CRUDE BIRTH RATES AND GENERAL FERTILITY RATES, 1964-1974

Year	*Corrected crude birth rate (live births per 1 000 population)*	*General fertility rate (per 1 000 women aged 15-54 years)*
1960	45.7	195
1961	45.4	...
1962	44.2	...
1963	44.6	...
1964	46.2	207
1965	43.5	190
1966	43.8	193
1967	40.8	178
1968	40.3	175
1969	40.7	177
1970	38.2	164
1971	36.8	157
1972	39.3	165
1973	37.5	156
1974	35.7	149

Sources: For corrected crude birth rate in 1960, A. Marcoux, "Evolution générale et mouvements saisonniers des naissances en Tunisie de 1956 à 1968", *Revue tunisienne de sciences sociales,* vol. 7, No. 20 (March 1970), pp. 173-214; for general fertility rate in 1960, 1964 and 1965, Jacques Vallin and R. J. Lapham, "Place du planning familial dans l'évolution récente de la natalité en Tunisie", *Revue tunisienne de sciences sociales,* vol. 6, Nos. 17-18 (1966), pp. 379-414; for 1974 data, Tunisia, Office national du planning familial et de la population, *Statistiques de planning familial,* No. 5, second quarter (Tunis, 1975). All other data obtained from Institut national de la statistique, Tunis.

Accordingly, from 1970 onward, the crude birth rate fell below 40 per 1,000 and the general fertility rate (women aged 15–54 years) below 170 per 1,000.

At the regional level, there are still disparities between the northern governorates and the coastal regions, in which contraception is already widely practised, and the central and southern governorates, whose population still strongly favours large families. In 1970, the official extreme figures for regional crude birth rates were as follows:

	Live births per 1 000 population
Gafsa	41
Kairouan	39
Medenine	41
Bejá	34
Jendouba	33
Le Kef	34
Sfax	33
Tunis	33

These differences are also attributable to the relative impact of the rural environment and the level of economic and social development. Thus far, the population of rural areas has been little affected by new modes of thinking, and their economic and social conditions are very close to the traditional model.

Three models of fertility behaviour have been identified[5] in contemporary society:

(a) *Primal fertility in rural areas:* families are large (from seven to eight children), and age at last birth is high (40–41.5 years), whereas age at marriage is low (18–19 years);

(b) *Advanced fertility pattern found in urban areas among the more affluent social categories:* final family size is no more than three children and contraception is widely practised;

(c) *Transitional fertility pattern of rural population having recently migrated to Tunis:* age at marriage and size of family approach those of rural areas, but age at last birth has fallen by one or two years. Women appear to adopt family planning in the last years of reproductive life.

Having analysed the results of this recent research, the authorities of the family planning programme currently have as their objective the further development of information and education, but with more specific emphasis on agricultural and working-class groups and on integrating population matters into secondary education. Since 1971, new structures have been set up. After the failure of the National Institute for Family Planning and Maternal and Child Welfare, an Act of 1973 created the National Family Planning and Population Office, which has carried out a complete reorganization of services and activities. Quantitative targets have been set for each method and in terms of births to be averted with a view to achieving in 2001 age-specific fertility rates equal to those of Italy in 1970 (see tables 4–7). Tables 6 and 7 provide evidence of the growth of these activities since 1973, which is confirmed by such encouraging results as: rate of protection per 100 women of reproductive age; 8.47 per 100 at the end of 1973 and 10.06 per 100 at the end of 1974;[6] rate of coverage of objectives, 103 per cent in 1974 and 113 per cent in 1975; and 392 family planning centres in operation in 1974.[7]

Another factor in this success was the complete liberalization of social abortion after 26 September 1973, and the extension of the practice of tubal liga-

[5] M. B'Chir and others, "L'influence sur le taux de fécondité du statut et du rôle de la femme dans la société tunisienne", *Revue tunisienne de sciences sociales,* vol. 10, No. 32–35 (1973), pp. 103–159.

[6] L. Behar, "Taux de protection par le planning familial en 1974 et 1975", *Bulletin de documentation de l'ONPFP* (Tunis, Office national du planning familial et de la population, 1975).

[7] Yolande Jemai, "L'évolution du nombre de centres de planning familial en activité de 1970 au 30 juin 1974", Tunis, Office national du planning familial et de la population, 1974 (mimeographed).

TABLE 4. TARGET FERTILITY RATES FOR TUNISIA IN 2001, BY AGE GROUP

(*Rates per 1,000*)

Age group	*Tunisia, 1971*	*Target 2001 = Italy 1970*
15-19	46	22
20-24	273	104
25-29	321	156
30-34	287	108
35-39	214	58
40-44	102	18
45-49	30	2
50-54	9	—
Crude reproduction rate	3.11	1.14

Source: Tunisia, Institut national de la statistique, *Perspectives d'évolution de la population, 1971-2001* (Tunis, 1972).

tion, but the most popular contraceptive method remains the IUD.[8]

In 1973, a survey of the continuation of contraceptive methods (IUD and pill)[9] gave the following results:

(*a*) Contraception begins late, since the average age

[8] See, for example, M. Ayad, "Les caractéristiques des acceptantes de DIU de 1970 à 1974", April 1975 (mimeographed); and A. Marcoux, "Les utilisatrices du programme de planning familial à la fin de l'année 1973", Tunis, Population Council, February 1974 (mimeographed).

[9] Tunisia, Office national du planning familial et de la population, *Enquête nationale sur la continuation des méthodes contraceptives, 1973:* vol. I. *Présentation et méthodologie;* vol. II. *Exploitation et résultats* (Tunis, 1974–1975).

TABLE 5. TARGETS AND ACTIVITIES OF FAMILY PLANNING PROGRAMMES, 1971-1981

(a) *Number of births to be averted per annum by the family planning programme, 1971-1981*

Year	*Births to be averted*	*Year*	*Births to be averted*
1971	12 000	1977	33 750
1972	15 500	1978	37 500
1973	19 000	1979	41 500
1974	22 500	1980	45 500
1975	26 250	1981	49 500
1976	30 000		

(b) *Former and revised targets for births to be averted, 1976 and 1977*

Births to be averted	*1976*	*1977*
Former target	30 000	33 750
New target	34 300	45 200
Difference	+ 4 300	+ 11 450

(c) *Family planning operations to be carried out, 1975-1976*

Year	*Tubal ligation*	*Social abortion*	*Intra-uterine device*	*Pill*	*Secondary method*
1975	12 000	12 500	25 200	17 800	3 400
1976	12 000	17 350	32 600	24 700	3 700

Source: Tunisia, Office national du planning familial et de la population, *Programme d'activités de l'ONPFP de 1974 à 1977* (Tunis, 1974).

TABLE 6. RESULTS OF ACTIVITIES OF THE NATIONAL FAMILY PLANNING PROGRAMME, 1971-1975

Year	*Primary insertion of intra-uterine device*	*New acceptors of the pill*	*Tubal ligation*	*Social abortion*	*New consultations*	*Total consultations*
1971	12 381	11 778	2 280	3 197	40 360	239 916
1972	13 250	12 026	2 453	4 621	43 665	246 675
1973	16 790	11 194	4 964	6 547	43 840	241 355
1974	19 084	10 795	10 757	12 427	50 901	256 984
1975[a]	(9 917)	(7 709)	(6 503)	(7 833)	(30 927)	(151 714)

Source: Tunisia, Office national du planning familial et de la population, *Statistiques des activités du programme de planning familial de 1971 à 1973* (Tunis, 1974).

[a] Result of the first quarter.

TABLE 7. BIRTHS AVERTED COMPARED WITH TARGET FOR 1971-1975

	1971	*1972*	*1973*	*1974*	*1975*
Target	12 000	15 500	19 000	22 500	26 250
Achieved	13 330	15 515	17 288	23 117	29 720
$R = \frac{\text{Achieved}}{\text{Target}}$	1.11	1.00	0.91	1.03	1.13

Source: Tunisia, Office national du planning familial et de la population, *Statistiques de planning familial; indicateurs de planning familial,* No. 6, second quarter (Tunis, 1975).

of acceptors is over 30 years (pill, 31 years; IUD, 32 years);

(*b*) Parity among female contraceptive users is higher than that observed elsewhere in the world (pill, 4.4; IUD, 4.75);

(*c*) The educational level of acceptors is very low: 75 per cent are illiterate;

(*d*) Continuation rates are about the average observed elsewhere in the world: pill, 56 per cent after six months, 42 per cent after one year; IUD, 84 per cent after six months, 75 per cent after one year;

(*e*) Protection is less effective for acceptors of the pill than for acceptors of IUD. The pregnancy rate of pill acceptors is not very different from that of the female population as a whole and 25 per cent of acceptors of IUD are pregnant within two years of acceptance, as against 50 per cent of acceptors of the pill.

It has been claimed that the decline in fertility could not continue unless very ambitious family planning targets were achieved. It looks as if they have indeed been achieved: the percentage of births averted by family planning compared with births registered (5 per cent correction) was 4.9 per cent in 1970, but appears to have been 11.5 per cent in 1974. Is there any justification for stating that the decline in the birth rate is largely due to the impact of the national family planning programme in 1974–1975, while the existing age structure is in itself an impediment to the continuation of past trends? It is hoped that the application of the methods proposed for this study will answer that question.

APPLICATION OF THE METHODS TO TUNISIA[10]

Standardization approach

The estimates

In order to determine the relative impact of all the factors involved, the base year selected was 1966, a census year and a year in which fertility was very close to natural fertility.

The official data of the National Statistical Institute (INS) at Tunis were used to show the number of females in the 15–54 age groups from 1966 to 1971; the corrected figures for births registered in 1966 and 1971; and the general fertility rates (for women aged 15–54 years) and age-specific rates for 1966 and 1971.

The application to all women aged 15–54 years in 1971 of the general fertility rate of 1966 gives the number of births which would have occurred among those women if fertility conditions had remained unchanged. Thus, by applying to the female population aged 15–54 in 1971 (whose numbers are estimated at 1,229,300) the 1966 general fertility rate of 193 per 1,000, one obtains 237,255 theoretical births. In fact, births registered by the civil registry authorities, with a 5 per cent correction,[11] numbered 192,959 (see table 8). It is therefore considered that 44,296

[10] This report does not discuss in detail the characteristics of these methods, which are described in the background paper entitled "Methods of measuring the impact of family planning programmes on fertility: problems and issues" (ESA/P/AC.7/1); see part one of the present publication.

[11] The statistics of the civil registry authorities since 1960 are very valuable, as it is generally believed that the rate of registry coverage is 95 per cent for births and 75 per cent for deaths. Such gaps as exist are largely due to non-declaration of the birth of a girl or of the death of a very young child in rural areas. For 1966, the birth figures have not been corrected, because Institut national de la statistique believes the rate of coverage to be 100 per cent for that year.

TABLE 8. DATA USED IN CALCULATING HYPOTHETICAL BIRTHS IN 1971 ON BASIS OF GENERAL FERTILITY RATE IN 1966

	1966	*1971*
Births registered	206 730	183 311
Corrected number of births	206 730	192 959
Women aged 15-54 (mid-year)	1 071 300	1 229 300
Proportion of married women aged 15-54, per 100	71.8	66.7
Married women aged 15-54 (mid-year)	769 460	818 752
General fertility rate *(per 1,000 women aged 15-45)*	193	157
General legitimate fertility rate *(per 1,000 women aged 15-45)*	268.6	234

Sources: For number of women aged 15-54 years in 1971, Tunisia, Institut national de la statistique, *Perspectives d'évolution de la population, 1971-2001* (Tunis, 1972); proportion of married women in 1971, estimate of A. Marcoux, "Sur les facteurs de l'évolution passée et future des naissances en Tunisie", Tunis, Population Council, 1972 (mineographed). Other data taken from Tunisia, Civil Register; and Institut national de la statistique, *Recensement général de la population et des logements du 3 mai 1966,* new ed. (Tunis, 1973).

Note:

Hypothetical births in 1971 on basis of general fertility rate in 1966: 1 299 300 × 193	237 255
Births registered (corrected) in 1971	192 959
Births averted in 1971 *(standardized on the basis of 1966)*	44 296

(237,255 – 192, 959) births were averted by all the factors, i.e., 23 per cent of the corrected registered births in 1971, standardized on the basis of 1966.

This decline in births can be attributed to the three effects described below.

Effect of change in age structure of the female population

Had the general age-specific fertility rate remained constant between 1966 and 1971, there would have been 225,410 hypothetical births on the basis of the size of the female population in 1971 (table 9). Thus, the changes in the age structure of the female population aged 15–54 have in themselves had the effect of lowering the number of births by 11,845, or 26.74 per cent of the total births averted.

Effect of changes in legitimate age-specific fertility rates

It has been estimated that there would have been 214,789 hypothetical births in 1971 if the age-specific legitimate fertility rates had remained at the 1966 level; however, there were 192,959 births in 1971, or 21,830 fewer births, which represents 49.28 per cent of the total of births averted (see tables 10 and 11).

Effect of variations in structure by marital status

Had the proportions of married women remained constant in each age group between 1966 and 1971, the number of births estimated on the basis of constant general fertility and constant legitimate fertility rates would have been identical (illegitimate births, amounting to approximately 3 per cent in Tunisia, are not taken into account here). Thus, the difference between these figures is due solely to the effect of changes in these proportions between 1966 and 1971.

It is accordingly necessary to estimate the hypothetical births which would have taken place among married women in 1971 if legitimate age-specific fertility had remained at the 1966 level.

The difference between the total of hypothetical births in 1971 on the basis of general age-specific fertility in 1966 and the total of hypothetical births in 1971 on the basis of legitimate age-specific fertility in 1966 gives the births averted as the result of changes in structure by marital status, as follows:

$$225{,}410 - 214{,}789 = 10{,}621 \text{ births}$$

or $\dfrac{10{,}621}{44{,}296} = 23.98$ per cent of total births averted:

Hypothetical births in 1971 on basis of 1966 general age-specific fertility rate	225,410
Hypothetical births in 1971 on basis of 1966 legitimate age-specific fertility rate	214,789
Births averted by changes of structure by marital status	10,621

The results

The estimates given above can be summarized as follows:

Corrected figures for births registered in 1971	192,959
Total births averted in 1971 (by comparison with the base year 1966)	44,296

TABLE 9. ESTIMATE OF HYPOTHETICAL BIRTHS IN 1971, BASED ON GENERAL FERTILITY RATES IN 1966

Age group	1966		1971		
	Women aged 15-54 (thousands, at mid-year) (1)	General fertility rate (per 1 000 women aged 15-54) (2)	Women age 15-54 (thousands, at mid-year) (3)	General fertility rate (per 1 000 women aged 15-54) (4)	Hypothetical births based on 1966 age-specific rates (2) × (3)=(5)
15-19	198.8	73	291.8	41	21 300
20-24	153.3	296	190.3	255	56 330
25-29	152.3	350	143.8	309	50 330
30-34	147.1	316	144.8	283	45 760
35-39	131.3	236	141.0	210	33 280
40-44	105.4	114	126.1	102	14 380
45-49	95.6	31	101.1	26	3 130
50-54	87.5	10	90.4	9	900
TOTAL	1 071.3	193	1 229.3	157	225 410

Sources: Tunisia, Institut national de la statistique, *Recensement général de la population et des logements du 3 mai 1966*, new ed. (Tunis, 1973); for forecast of number of women aged 15-54 years in 1971, *idem, Perspectives d'évolution de la population, 1971-2001* (Tunis, 1972).

Note:

Hypothetical births in 1971 on the basis of the general fertility rate for 1966	= 237 255
Hypothetical births in 1971 on the basis of the age-specific general fertility rates in 1966	= 225 410
Births averted in 1971 by the change in the age structure of the female population (by comparison with 1966)	= 11 845

TABLE 10. DATA USED IN ESTIMATING HYPOTHETICAL BIRTHS IN 1971 ON BASIS OF LEGITIMATE FERTILITY RATES IN 1966

Age group	1966				1971			
	Women aged 15-54 (thousands, at mid-year)	*Proportion of married women (percentage)*	*Married women aged 15-54*	*Legitimate fertility rate (per 1 000)*	*Women aged 15-54 (thousands, at mid-year)*	*Proportion of married women[a] (percentage)*	*Married women aged 15-54*	*Legitimate fertility rate (per 1 000)*
15-19	198.8	18.5	36 778	400	291.8	12.0	35 016	342
20-24	153.3	71.1	108 996	410	190.3	68.7	130 736	371
25-29	152.3	88.5	134 785	398	143.8	85.5	122 949	362
30-34	147.1	92.4	135 920	342	144.8	92.1	133 361	307
35-39	131.3	91.8	120 533	260	141.0	92.7	130 707	226
40-44	105.4	87.9	92 647	129	126.1	89.8	113 238	113
45-49	95.6	81.8	78 201	38	101.1	84.2	85 126	32
50-54	87.5	70.4	61 600	12	90.4	74.8	67 619	13
TOTAL	1 071.3	71.8	769 460	268.6	1 229.3	66.7	818 752	234

Source: Tunisia, Institut national de la statistique.

[a] Proportions estimated by extrapolating the data for 1956 and 1966 and taking into account the raising of the age of marriage by a law of 1964. See A. Marcoux, "Sur les facteurs de l'évolution passée et future des naissances en Tunisie", Tunis, Population Council, 1972 (mimeographed).

Broken down as follows:

	Number	*Percentage*
Effect of age structure	11,845	27
Effect of structure by marital status	10,621	24
Effect of variations in legitimate fertility	21,830	49
	44,296	100

The diminished birth cohorts caused by epidemics and food shortages during the Second World War (1939–1945) largely explain the births averted by the effect of structure. Without these losses, the number of women aged 25–29 in 1971 would have been higher by 20,000; and they would have given birth to 6,180 children if their general fertility rate had been the same as that observed in the group aged 25–29 in 1970, or 309 per 1,000.

It should be noted that the effect of the diminished cohorts reaching their maximum fertility between 1969 and 1973 is due to the circumstances prevailing at the time. Beginning with 1975, the 25–29 year age group will be composed of the larger cohorts born after the war; furthermore, estimates have shown that the number of births averted by the effect of age structure after 1971 was declining.[12]

TABLE 11. ESTIMATE OF HYPOTHETICAL BIRTHS IN 1971, BASED ON LEGITIMATE FERTILITY RATES IN 1966

Age group	*Married women aged 15-54 years, 1971*	*Legitimate age-specific fertility rate, 1966 (per 1 000)*	*Hypthetical births in 1971 based on legitimate age-specific fertility rates for 1966*
15-19	35 016	400	14 006
20-24	130 736	410	53 602
25-29	122 949	398	48 934
30-34	133 361	342	45 609
35-39	130 707	260	33 984
40-44	113 238	129	14 608
45-49	85 126	38	3 235
50-54	67 619	12	811
TOTAL	818 752	268.6	214 789

Note:

Hypothetical births in 1971 based on legitimate age-specific fertility rates for 1966	214 789
Registered births (corrected figures) in 1971	192 959
Births averted in 1971 by changes in legitimate fertility (by comparison with 1966)	21 830

Nearly one fourth of the births averted have been due to the change in structure by marital status and are concentrated in the 15–19 age group. This situation results from the application of the law of 1964 raising the age of marriage for women to 17 complete years, but also from a number of factors tending to modify traditional structures as they affect marriage (school attendance of girls, growing urbanization, advancement of women in society in general).[13] Lastly, almost half the births averted are attributable to factors specifically affecting fertility as such, including the impact of the national family planning programme.

If the estimates of the Population Division of ONPFP are accepted as satisfactory, it will be seen that the activities of the official programme have in themselves had the effect of averting 13,330 births, or 30.09 per cent of the total of births averted, which coincides with the percentage generally acknowledged by various researchers. This estimate depends upon the validity of the method used in estimating the number of women protected and of births averted; however, account has not been taken of contraception practised by individuals independently of the government programme or of the indirect effects produced by official education and information programmes.

As to the other factors (social and economic), it is recognized that they must have helped to determine the new behaviour adopted by many couples, but any measurement of them is still a very rough estimate. The most important among them are the spread of education and particularly school attendance of girls (however, the increased attendance resulting from the raising of the minimum age of marriage for girls to 17

[12] M. Ayad, "La fécondité des Tunisiennes en mutation", Tunis, Office national du planning familial et de la population, 1975.

[13] L. Behar, "Evolution récente de la nuptialité en Tunisie", Tunis, Office du planning familial et de la population, 1975.

has not reduced appreciably the reproductive period of the lives of women in general); and migrations. Workers in foreign countries, who have been numerous since 1966, often live alone in the receiving country and the absence of the husband from Tunisia reduces the reproductive life to a corresponding extent; but in the majority of cases the husband returns each year and the total progeny of the couple is not reduced. However, internal migration has definitely had an important impact on the decline in fertility because the migrants leaving their villages for the towns often adopt urban behaviour, as has been verified by INS.[14]

As concerns the decline in mortality, the general death rate, as officially reported by INS with a correction of 30 per cent, dropped from 14.0 per 1,000 in 1966 to 95 in 1974. The decline in infant mortality has been much smaller, but the progress made in the field of health has increased the chances of a couple seeing their children reach adulthood. The other economic and social development factors have had an impact only in their combined action, which is all the more difficult to measure; individually, they have not been determinative. Nevertheless, through their influence on fertility they have accounted for about 20 per cent of the total of births averted, which is by no means negligible; and it would be worth while to determine more explicitly how this effect has been achieved, provided an appropriate method of doing so could be applied.

Appraisal of the method

The standardization approach has the advantage of being easy to apply and of showing clearly the impact of the different factors analysed in lowering fertility. It leads to results that are consistent with findings actually made, and it can thus be usefully taken into consideration by the authority in assessing the proportion of the decline in fertility which will take place spontaneously, thus relieving the burden on the family planning programme.

Nevertheless, it is this very simplicity which also constitutes the weakness of the method, for the three reasons discussed below.

First, it is assumed that the factors are additive and independent of one another, and some are favoured because they are easier to measure.

The fact is, however, that in Tunisia, the family planning programme is part of an over-all development model which favours small families and penalizes those which deviate from the model. All the economic and social measures on which the planning undertaken since 1962 is based are directed to a type of society where the traditional family cannot continue to exist and where the couple is obliged to change its behaviour and, in particular, its attitude towards procreation, exercising more control in this sphere if it wishes to share in and benefit from the advantages the model offers (education for the couple's children, home ownership, a motorcar, medical care etc.).

The family planning policy cannot be reduced to the activities of the programme properly so called, for all the legislative measures referred to earlier, including the law of 1964 raising the age of marriage, contribute to the possibility of making a controlling choice, at all levels, with regard to fertility.

Secondly, in addition to acknowledging the rigidity of this technique of analysis, one must also realize the fragility of the conclusions reached, depending upon the degree of reliability of the data used. Thus, for the present estimates, it was considered that under-registration of births amounted to 5 per cent (according to INS), whereas some sources maintain that under-registration in 1971 amounted to 6.9 per cent, which would mean that the number of births would be higher by 3,938 (196,897 instead of 192,959).

Furthermore, the figure for the female population in 1971 is the result of forecasts and of an estimate of the proportions of married women by extrapolation, the quality of which cannot be assessed until after publication of the results of the 1975 census. Lastly, the formula that has made it possible to relate hypothetical births to constant legitimate and general fertility is based on the assumption that the number of illegitimate births is negligible (2.5 per cent in 1971 and 3.3 per cent in 1966).

Thirdly, in taking 1966 as the base year, the primary considerations were the use of census data and the beginning of the official family planning activities, but where structure by age and by marital status is concerned, it is not certain that that choice was ideal. The fact is that by 1966 the effects of the law of 1964 were already being felt; and, in addition, a distortion occurred in nuptiality for the period 1964–1966, due to circumstances prevailing at that time.

Moreover, the present estimate assumes that the women who did not marry or were not born would have had the same fertility as the women who were actually married in 1971. The effects of structure by age and by marital status are undoubtedly over-estimated.

Trend analysis

The estimates

On the basis of the series of birth rates (corrected figures) for the period from 1956 to 1973,[15] the level and trend of fertility before the introduction of the programme were determined. Dealing in the same manner with the period after 1964, it was possible to

[14] Tunisia, Institut national de la statistique, *Enquête migration et emploi à Tunis, 1972–1973. Résultats*, Demographie Series, No. 4 (Tunis, 1974).

[15] For the period 1956–1959, the figures are those given in A. Marcoux, "La croissance de la population de la Tunisie: passé récent et perspectives", *Population* (Paris), vol. 26, special issue (March 1971), pp. 105–123; for 1960–1973, official figures published by Institut national de le statistique have been used.

measure the extent of the decline which, according to this method, is attributable to the family planning programme (see table 12).

The figures given below show first the estimates on the basis of which the trend of fertility from 1956 to 1963 was determined, and then the estimates from 1965 to 1971 (see also figure I):

CRUDE BIRTH RATE, 1956–1963; ESTIMATE OF TREND

x	y	x^2	y^2	xy
1	46.4	1	2,152.9	46.4
2	46.5	4	2,162.2	93.0
3	46.3	9	2,143.6	138.9
4	46.2	16	2,134.4	184.8
5	45.7	25	2,088.4	228.5
6	45.4	36	2,061.1	272.4
7	44.2	49	1,953.6	309.4
8	44.6	64	1,989.1	356.8
36	365.3	204	16,685.3	1,630.2

$n = 8$
$\bar{x} = 4.5$
$\bar{y} = 45.7$

$$a = \frac{\Sigma\, xiyi - n\,\bar{x}\,\bar{y}}{\Sigma\,(xi)^2\, n\,\bar{x}^2}$$

$$a = \frac{1{,}630.2 - 1{,}643.8}{240 - 162} = -13.6/42 = \boxed{-0.32}$$

$$b = \bar{y} - a\bar{x} = 45.7 - [(-0.32)\,(4.5)] = 47.14$$

$$\boxed{Y_1 = -0.32\ X_1 + 47.14}$$

CRUDE BIRTH RATE, 1965–1971; ESTIMATE OF TREND

x	y	x^2	y^2	xy
10	43.5	100	1,892.25	435.0
11	43.8	121	1,918.44	481.8
12	40.8	144	1,664.64	489.6
13	40.3	169	1,624.09	523.9
14	40.7	196	1,656.49	569.8
15	38.2	225	1,459.24	573.0
16	36.8	256	1,354.24	588.8
91	284.1	1,211	11,569.39	3,661.9

$n = 7$
$\bar{x} = 13$
$\bar{y} = 40.6$

$$a = \frac{3{,}661.9 - [(7)(13)(40.6)]}{1{,}211 - 7(13)^2} = \frac{[3{,}6619 - 3{,}694.6]}{1{,}211 - 1{,}183}$$

$$= \frac{-32.7}{28} = \boxed{-1.17}$$

$$b = \bar{y} - a\bar{x} = 40.6 - [-1.17(13)] = [55.81]$$

$$\boxed{Y_2 = -1.17\,X_2 + 55.81}$$

The results

With reference to the birth rate, the estimates and figure I show the following results:

(*a*) From 1956 to 1963, a slow decline of 0.32 point per annum;

(*b*) After the 1964 peak, resulting from the rush of marriages in 1963, a more rapid decline of 1.17 points per annum.

TABLE 12. CRUDE BIRTH RATES AND PERCENTAGE OF DECLINE IN SUCCESSIVE YEARS, 1956-1971

Year	Rate	Decline	Year	Rate	Decline
1956	46.4		1964	46.2	+3.6
1957	46.5	+0.2	1965	43.5	−5.8
1958	46.3	−0.4	1966	43.8	+0.7
1959	46.2	−0.2	1967	40.8	−6.8
1960	45.7	−1.0	1968	40.3	−1.2
1961	45.4	−0.7	1969	40.7	+1.0
1962	44.2	−2.6	1970	38.2	−6.1
1963	44.6	+0.9	1971	36.8	−3.7

Sources: For 1956-1959, A. Marcoux, "La croissance de la population de la Tunisie: passé récent et perspectives", *Population* (Paris), vol. 26, special issue (March 1971), pp. 105-123; for 1960-1971, official figures of Institut national de la statistique, Tunis.

Note: The year 1964 was exceptional: the peak in the birth rate registered that year (46.2 per 1 000) was actually a consequence of the rush of marriages in 1963, the year preceding that in which the age of marriage for women was fixed at 17 years. In estimates for measuring the percentage of the annual decline in the birth rate for the period 1956-1971, the rate for 1964 has been eliminated, because it distorted the general trend of the curve.

This method thus assumes that before the introduction of the family planning programme in 1964–1965, the decline was 0.32 point per annum; and the birth rate was influenced by urbanization, industrialization and all the other factors of economic and social development in general. It might have been expected that the trend would continue to be slow and linear in the following years; however, after family planning activities had begun and the circumstantial effects of the law of 1964 raising the age of marriage began to be felt, the decline accelerated, reaching 1.17 points per annum. Consequently, the increase in the decline of the crude birth rate must be attributed to the family planning programme.

Appraisal of the method

While noting that the interruption in the trend of the crude birth rate between 1956 and 1971 is very marked and that it can confidently be deduced that this change is due to the introduction of the family planning programme, it is nevertheless necessary, for a number of reasons, to be very cautious in making a quantified estimate of the impact of the programme.

First, the method of estimating the trend is very approximate and is based on a mathematical technique which is often challenged in its application to economic and social phenomena.

Secondly, the downward trend of the crude birth rate must be viewed in the light of trends in the general fertility rate, shown in table 13 and figure II.[16]

Until 1961, the rate of increase in fertility appears to have been steady because the female population of

[16] This comment is based on the analysis of A. Marcoux, "La croissance de la population de la Tunisie: passé récent et perspectives".

TABLE 13. TRENDS IN THE GENERAL FERTILITY RATE, 1958-1974
(*Rate per 1 000 women aged 15-54 years*)

Year	*General fertility rate*	*Year*	*General fertility rate*
1958	186.6	1966	193
1959	188.7	1967	178
1960	189.5	1968	175
1961	190.6	1969	177
1962	187.0	1970	164
1963	188.6	1971	157
1964	198.1	1972	165
1965	187.2	1973	156
		1974	149

Sources: For 1958-1965, rates estimated by A. Marcoux, "La croissance de la population de la Tunisie: passé récent et perspectives", *Population* (Paris), vol. 26, special issue (March 1971), pp. 105-123; for 1966-1971, official rates; see Tunisia, Institut national de la statistique, *Naissances, décès, marriages, divorces, 1970; statistiques détaillées,* Demographic Series, No. 5 (Tunis, 1974).

reproductive age was growing at a rate markedly lower than that of the population as a whole, owing to emigration and the arrival at reproductive age of the "diminished birth cohorts"; however, that factor had only a very slight effect on fertility, because women in the 15–19 age group have a low fertility rate. Thus, those two phenomena did not appreciably decrease the birth potential, which explains the rise in general fertility rates (15–54 years).

In a second phase, 1964–1966, there was a circumstantial fluctuation due to the rush of marriages in 1963 in anticipation of the law of 1964, leading to an exceptional increase in births and hence to particularly high fertility rates. The general trend in the period 1966–1971 was unquestionably downward for both fertility rates and birth rates. But to the effect of the family planning programme, which has certainly been considerable, must be added the "braking" effect on fertility exercised by the "diminished birth cohorts" upon reaching the age of high fertility (in 1970, they formed the 25–29 age group). Thus, trend analysis (the fertility projection method) provides a very crude means of determining trends in the period in question.

Lastly, unlike the standardization approach applied above, trend analysis does not permit estimation of the relative impact of the various factors contributing to the decline in fertility and attributes to the programme an impact considerably greater than it has had in reality. If one considers, as has already been stated, not only family planning activities *per se* but the aggregate of the indirect effects of the programme, and then the demographic policy of Tunisia, it can be seen that this method enables one to measure in a much more significant way the results of the policy implemented by the authorities.

As in the case of the standardization method, here too it is necessary to be able to judge the reliability of the data used. As concerns births, this aspect has already been commented upon. With respect to total population (see table 14), the INS estimates were used by interpolation of the censuses between 1956 and 1966 and by extrapolation for the period 1966–1971.[17]

[17] The forecasts of Institut national de la statistiques for the period 1971–2001 (drawn up in 1972) are based on several hypotheses:

(a) *Mortality:* according to United Nations studies made in countries at a level of development similar to that of Tunisia, life expectancy at birth is increasing by about 6 months per year;

(b) *Fertility:* it should drop steadily, reaching in 2001 the rates per age prevailing in Italy in 1970 (cited in table 4);

(c) *External migration:* it should gradually become negligible.

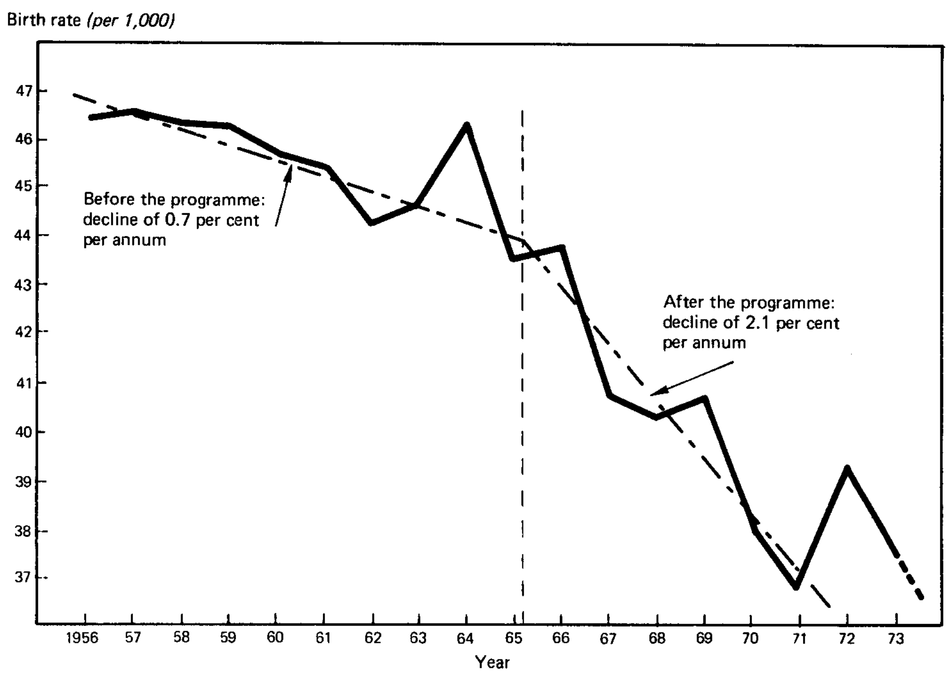


Figure I. Tunisia: trends in the crude birth rate, 1950–1973

Figure II. Tunisia: trend in the general fertility rate, 1958-1974

Experimental designs

Experimental group A

The experimental group (A) selected for the application of the experimental-design method consists of the women who were interviewed for a survey on "continuation of the principal contraceptive methods under the Tunisian Family Planning Programme" conducted by ONPFP in 1973.[18]

The purpose of the survey was to provide scientific answers to the questions that arise concerning the effectiveness and demographic impact of the programme and concerning some aspects of the actual operation of the programme.

Characteristics of the sampling procedure

The choices that governed the preparation of the sampling scheme were as follows:

(*a*) The persons about whom information was sought were women who had accepted the IUD or the pill at family planning programme centres;

(*b*) Of those women, the population to be interviewed would be restricted to those who had accepted after 1 January 1969 and at least a year before the survey, so that for all the respondents there would be data relating to a one-year continuation.

(*c*) The sample would be nation-wide, meaning that every acceptor, whatever centre she belonged to, would have the same probability of being included in the sample;

(*d*) The statistical results that the survey would attempt to obtain would be nation-wide;

(*e*) In order to be able to detect any disparities between the continuation rates for the two methods at a satisfactory level of significance, the sample to be surveyed would consist of 2,000 persons, comprising 1,000 IUD acceptors and 1,000 pill acceptors. The sampling rate would thus be 1/35.

When it came to drawing the sample, the question was what procedure should be chosen for making the selection from among the population of 69,906 acceptors registered, between January 1969 and April 1972, at 350 centres whose volume of activity for that

[18] Tunisia, Office national du planning familial et de la population, *Enquête nationale sur la continuation des méthodes contraceptives, 1973;* vol. I. *Présentation et méthodologie;* vol. II. *Exploitation et résultats.*

TABLE 14. TRENDS IN PROPORTION OF TOTAL POPULATION REPRESENTED BY FEMALE POPULATION, 1958-1971

Year	*Total population at 1 July (thousands)* (1)	*Female population aged 15-54 years (thousands)* (2)	*Proportion of women aged 15-54 years (percentage)* (2)/(1)
1958	4 040.0	1 002.0	24.8
1959	4 117.0	1 007.0	24.4
1960	4 198.5	1 013.0	24.1
1961	4 268.5	1 018.0	23.8
1962	4 335.0	1 027.0	23.7
1963	4 422.0	1 039.0	23.5
1964	4 523.0	1 050.0	23.2
1965	4 619.5	1 063.0	23.0
1966	4 717.5	1 071.3	22.7
1967	4 825.0	1 103.7	22.9
1968	4 928.0	1 135.2	23.0
1969	5 027.5	1 165.1	23.2
1970	5 126.5	1 194.3	23.3
1971	5 228.4	1 229.3	23.5

Source: Tunisia, Institut national de la statistique.

period varied from 1 to 4,868 acceptors. The following choices were made:

(*a*) The sample would be drawn in two stages: first a sample of centres; and then a sample of acceptors at each selected centre;

(*b*) The distribution of the 2,000 expected respondents would be 40 women per centre at 50 centres;

(*c*) Centres would be drawn with a probability proportionate to their "size" (number of acceptors during the reference period), and the same number of women (40) would be drawn at each selected centre, thus giving every acceptor an equal probability of being included in the sample;

(*d*) The method of drawing centres would allow for geographical stratification, with the aim of improving accuracy and obtaining some data by major geographical divisions.

Thus, the theoretical sample—the number of acceptors multiplied by the sampling fraction—would number 2,227. In fact, the actual sample comprised only 2,060 acceptors, or 92.5 per cent of the 2,227 expected.

However, it is important to mention that this sample, which forms the experimental group A, includes only acceptors of two principal methods, namely, the pill and IUD, the results for which are dealt with separately. This restriction on the composition of the group was necessitated by the lack of any statistical data concerning acceptors of other contraceptive methods (another survey is in progress on the characteristics of women who have undergone an abortion or tubal ligation).

Control groups

It was impossible to find in Tunisia a group of women who had not been affected by the national family planning programme and had socio-demographic characteristics similar to those of acceptors. In order to be able to apply the experimental method in spite of this problem, two alternative solutions were found.

First, the fertility of IUD and pill acceptors, as determined by the survey described above, was compared with the age-specific fertility rates for all married women for 1966 and 1971. Acceptors are, of course, included in this population, a fact which tends to over-represent the effects of the programme for the purpose of comparison with 1966 fertility and to under-represent its effects in the comparison with 1971 fertility.

The comparability of the two groups is difficult to analyse because for 1971 the socio-demographic characteristics of the married female population are not known and reference had to be made to the 1966 census figures. It may, however, be assumed that the changes were not very great. This control group is hereafter referred to as group B.

Secondly, the fertility of acceptors before and after acceptance also was compared, assuming that the control group, referred to as group C, consists of women who are not yet acceptors and comparing the fertility of the two groups at the same age.

Comparability of the groups

Comparison with control group B

On the basis of the data available both for experimental group A and for control group B, it was possible to select five comparison factors, which are discussed separately below.

Age distribution. For the purpose of comparing the age distribution of group A with that of group B, it was decided to match acceptors with married women—and not all women—according to the census figure for 1966 and an estimated figure for 1971 determined by application of the standardization approach (table 15). This decision was facilitated, first, by the fact that in Tunisia, acceptors are often, if not always, married; and, secondly, by the availability of this datum for both 1966 and 1971.

Before comparing these age distributions, it is important to bear in mind some data yielded by the "continuation" survey. According to the results of that survey,[19] the mean age at the time of acceptance is 30.8 years for IUD acceptors and 30.0 years for pill acceptors; the median age is 31.5 years for the former and 30.5 years for the latter.

These figures, when compared with observed behaviour in other countries, show that acceptance of contraception occurs rather late in Tunisia. Only Hong Kong (32.4 years in 1968), Morocco (33 in 1972) and the Republic of Korea (33.2 in 1966) have higher median ages than Tunisia for IUD. In the case of the pill, the only example of the same order is that of the Republic of Korea (34 years in 1969).

It must be said that it is quite common for the population of pill acceptors to be younger than that of women practising intra-uterine contraception. This situation is often due to the fact that oral contraceptives are more convenient for women who want to space births rather than prevent them altogether, and that the lower mean age of pill acceptors is associated, as is shown below, with lesser parity.

If one now compares the two distributions (see figure IV) for the experimental group with those for the control group in both 1966 and 1971, it can be seen that the former group differs from the latter in having a relative concentration of acceptances around the 30–34 age group. In any event, these distributions are seldom similar, which would indicate a comparable degree of acceptance of contraception at all ages.

Geographical distribution. For the purpose of comparing the geographical origin of experimental group A with that of control group B, lower and upper age limits of 20 and 44 years were set for women in the

[19] *Ibid.*, vol. II, p. 3.

TABLE 15. DISTRIBUTION, BY AGE GROUP, OF EXPERIMENTAL GROUP A, ACCORDING TO METHOD; AND CONTROL GROUP B, ACCORDING TO 1966 CENSUS AND 1971 ESTIMATE

(*Percentage*)

	Experimental group A according to method		*Control group B according to:*	
Age group	*Intra-uterine device*	*Pill*	*1966 census*	*1971 estimate*
15-19	1.9	2.7	5.0	4.7
20-24	18.0	19.3	15.5	17.4
25-29	21.5	24.0	19.8	16.4
30-34	25.2	25.6	19.8	17.8
35-39	21.6	17.5	17.3	17.4
40-44	7.4	6.5	12.7	15.0
45-49	1.9	1.3	9.9	11.3
Not stated	2.5	3.1	...	...
TOTAL	100.00 (1 136)	100.0 (924)	100.0 (690 453)	100.0 (751 133)

latter group (see table 16). The reason for this limitation was to avoid distorting the results, which would have vitiated the comparison, precisely because of the high percentage of underestimation (see table 15) of acceptors in group A in both the youngest (15–19) and the oldest (45–49) age groups.

Thus, the figures in table 16 show that there is a sizable over-representation of group-A women residing in the capital and its suburbs (29 per cent), compared with group-B women (17 per cent). This over-representation is greater (nearly double) in other urban areas (40 per cent in group A as against 21 per cent in group B). In the interior (rural) stratum, on the other hand, there is about a 50 per cent under-representation of women in the experimental group (31 per cent), compared with the control group, in which the proportion of women of rural origin is 62 per cent.

Parity by age group. The survey on the continuation of contraceptive methods showed a mean family size of 4.6 children still living at the time of acceptance in the case of IUD acceptors, compared with a mean of 4.3 for pill acceptors; the median figures were 4.5 and 4.2, respectively.

Compared with the 1966 census data (table 17), the mean parity of acceptors by age at the time of acceptance shows (see figure V) that such women (the experimental group) are distinctly more fertile—at the same age—than the average, this being particularly true after the age of 30.[20]

In order to corroborate these findings, a comparison was made of the distribution of women by number of children still living in the experimental group and the control group (table 18).

[20] A comparison between the two groups according to the number of live-born children was not possible owing to a lack of data for group A. However, it should be pointed out that deviations in mortality certainly could not cause such differences in the number of surviving children.

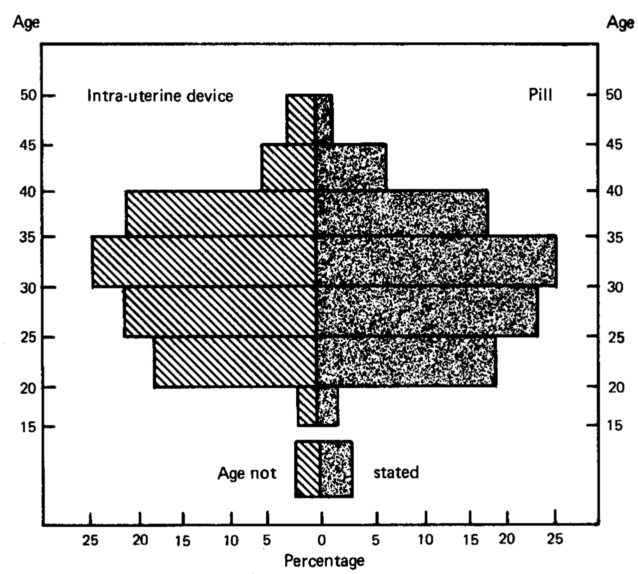


Figure III. Population pyramid of experimental group A, by contraceptive method

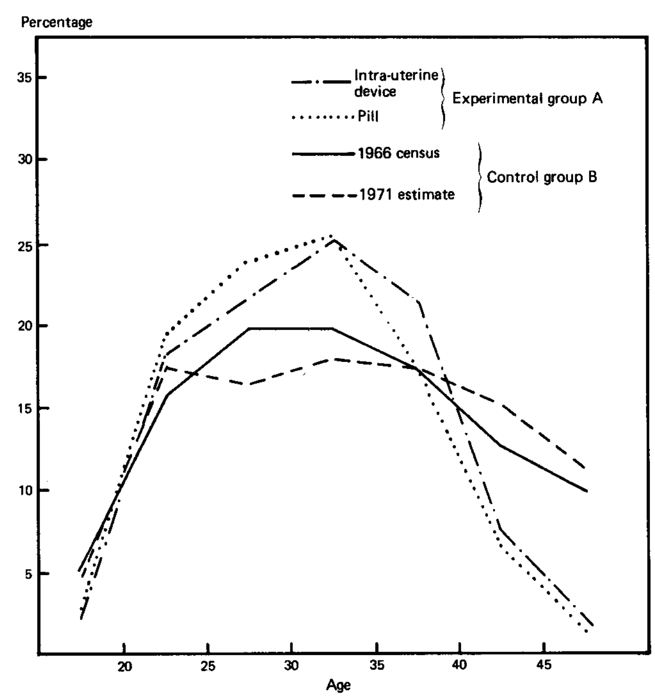


Figure IV. Distribution by age group of acceptors in experimental group A, according to method; and of married women in control group B under age 50 in 1966 (census) and 1971 (estimated)

TABLE 16. GEOGRAPHICAL DISTRIBUTION OF EXPERIMENTAL GROUP A AND CONTROL GROUP B (MARRIED WOMEN BETWEEN THE AGES OF 20 AND 44 IN 1966)

Geographical origin	Experimental group A		Control group B	
	Number	Percentage	Number	Percentage
Tunis and suburbs	600	29.0	95 394	17.0
Interior (urban)	826	40.0	124 876	21.0
Interior (rural)	634	31.0	366 113	62.0
TOTAL	2 060	100.0	586 383	100.0

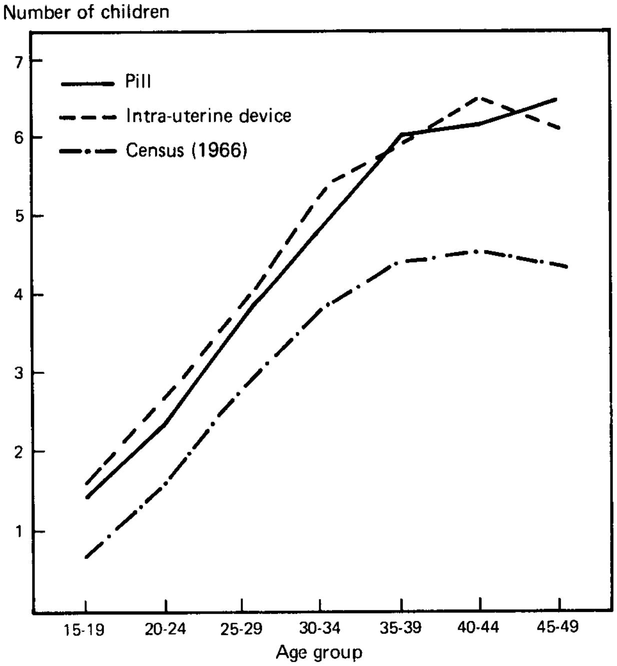


Figure V. Number of children still living, by age group of woman: comparison of experimental group A (by method) and control group B (1966 census)

TABLE 17. NUMBER OF CHILDREN STILL LIVING, BY AGE GROUP OF WOMEN, EXPERIMENTAL GROUP A AND CONTROL GROUP B

Age group	Experimental group A		Control group B
	Intra-uterine device	Pill	1966 census
15-19	1.60	1.40	0.66
20-24	2.47	2.30	1.60
25-29	3.88	3.67	2.78
30-34	5.30	4.88	3.80
35-39	5.88	6.0	4.40
40-44	6.48	6.18	4.57
45-49	6.08	6.40	4.48

Thus, tables 17 and 18 show that fertility is higher among contracepting women. It is especially after the fourth child that women decide to limit their offspring; and their fertility then exceeds—as a relative percentage—that of Tunisian women as a whole. This finding leads one to conclude that acceptance of contraception comes late in Tunisia, as previously noted in connexion with age distribution.

Level of education. For the purpose of analysing educational status in the two groups, four categories or levels of education were established. The first comprises women who have had no education and are therefore unable to read or write. The second category consists of women who received a minimal education either from the Koranic schools (*kuttabs*) or from the literacy campaign schools known as "schools of social education". The third category relates to women who completed all or part of the primary cycle of public education. The fourth category comprises all those who went beyond primary school. This classification permitted the comparison shown in table 19.

Within the experimental group itself, a slight difference can be seen, in that pill acceptors are, on average, a little better educated, one fourth of them having had at least some primary education, compared with one fifth of the IUD acceptors.

A comparison between women in the experimental group as a whole and those in the control group, on the other hand, shows that the former group is, on average, much better educated than those in the latter group, 25 per cent of contraceptors having had some formal education as against only 8 per cent of non-contraceptors. This is added evidence of the often-observed relationship between educational status and family planning practice.

Economic activity. The same observation applies to the economic activity of women, as shown in table 20.

In comparison with the control group, the acceptors in the experimental group have a higher proportion of economically active women (11.8 per cent, compared with 5.5 per cent in the control group). However, and contrary to the finding with respect to the level of

TABLE 18. DISTRIBUTION OF WOMEN IN EXPERIMENTAL GROUP A AND CONTROL GROUP B BY NUMBER OF SURVIVING CHILDREN

	Number of surviving children											
Percentage of women	0	1	2	3	4	5	6	7	8	9	10+	Not stated
Experimental group	0.4	6.7	12.9	14.4	16.9	16.1	13.7	8.5	5.0	2.1	1.5	1.8
Control group	9.7	12.6	14.3	14.6	16.0	12.5	8.0	5.4	3.0	1.4	0.8	1.1

TABLE 19. COMPARATIVE LEVELS OF EDUCATION OF EXPERIMENTAL GROUP A AND CONTROL GROUP B
(Percentage)

Level of education	Experimental group A			Control group B
	Intra-uterine device	Pill	Aggregate	1966 census
Illiterate	77.1	71.6	74.6	92.1
Kuttab or social education	1.8	3.4	2.6	0.4
Primary	15.7	19.1	17.3	4.3
Secondary, technical or higher	5.1	5.5	5.2	2.6
Other or not stated	0.3	0.4	0.3	0.6
TOTAL	100.0	100.0	100.0	100.0

TABLE 20. COMPARISON OF ECONOMIC ACTIVITY OF WOMEN IN EXPERIMENTAL GROUP A AND WOMEN IN CONTROL GROUP B
(Percentage)

Occupation	Experimental group A			Control group B
	Intra-uterine device	Pill	Total	1966 census
No occupation or housewife	86.9	89.4	88.0	94.0
Women economically active in the home or outside	12.8	10.4	11.8	5.5
Not stated	0.3	0.2	0.2	0.5
TOTAL	100.0	100.0	100.0	100.0

education within the experimental group, fewer acceptors of the pill (10.4 per cent) than acceptors of IUD (12.8 per cent) work.

Comparison with control group C

As mentioned above, control group C comprises the same women as experimental group A. The essential difference between them is that those in group C had never practised any method of contraception whereas group A comprises women who are acceptors of contraceptives.

In order to make this comparison between two fundamentally different states of the same group of women, it was assumed that women in, for example, the 25–29 age group agreed to practise contraception at an average age of 27.5, so that one can describe their fertility at "about age 25" (actually age 24.5) as "before acceptance". This result could then be set against that obtained for acceptors in the 20–24 age group, which would represent the fertility of this group at "about age 25" but "after acceptance".[21]

By using this five-year differential, one is able to compare two different states of the same group. The advantage of this method is, of course, that the socio-economic, cultural and demographic characteristics are identical because the same women are involved.

[21] The method of estimating fertility rates both before and after acceptance is set forth in the section on application and results of the method.

Fertility levels

Having defined the characteristics of each group, one may proceed to a comparison of the series of legitimate fertility rates by age group for experimental group A and control groups B and C.

Differences in fertility between experimental group A and control group B

Before a comparison is made of the fertility rates of the two groups, some comments are called for:

(a) *Regarding experimental group A:* only the series of rates after the second year of acceptance have been used. The reason for this choice is that a woman is not pregnant at the time of acceptance of contraception, which implies a selection process that is reflected in abnormally low fertility rates in the year following acceptance. Therefore, an average of each rate by age group was computed on the basis of the second, third and fourth years after acceptance. In table 21, this series of rates is compared to that obtained for group B;

(b) *Regarding control group B:* two series of age-specific rates were used. The first is that for married women in 1966, and the second is specific to married women in 1971. In order to make the subjects of the comparison homogeneous, the series of rates begin with that for the 20–24 age group, which therefore corresponds, in group A, to that for the group under 25 years. Furthermore and for the same purpose, an average rate was computed for the 40–44 and 45–49 age

TABLE 21. LEGITIMATE FERTILITY RATES FOR ACCEPTORS (*n* YEARS AFTER ACCEPTANCE) IN EXPERIMENTAL GROUP A AND FOR WOMEN IN GENERAL IN CONTROL GROUP B

(Rates per 1 000 women)

	Experimental group A									
	Intra-uterine device acceptors				*Pill acceptors*				*Control group B*	
Age group	*Second year*	*Third year*	*Fourth year*	*Average rate*	*Second year*	*Third year*	*Fourth year*	*Average rate*	*1966 census*	*1971 estimate*
Under 25	177	203	220	200	373	303	272	318	410	371
25-29	81	226	261	189	365	369	362	365	398	362
30-34	105	61	196	120	213	105	281	199	342	307
35-39	30	168	194	130	249	221	171	214	260	226
40 and over	66	45	52	54	41	45	[a]	...	83	72
All ages	95	145	193	144	276	235	253	255	269	234

[a] Insufficient data.

groups which corresponds to that for the 40+ members of group A.

Two comparisons can be made from the data in table 21: the first can be made at the level of the experimental group itself; and the second between group A and group B. Leaving aside for the time being the first level, which is discussed below, the series affecting the groups A and B can be considered.

General fertility rate. Comparing the general fertility rates, all ages combined, in 1966 for the experimental group and the control group gives the following results:

Group B, 1966	*Group A*	*Difference*
269	− 144 (IUD) =	−125
269	− 255 (Pill) =	− 14

The differences between the rates for experimental group A and the results obtained for control group B in 1971 are:

Group B, 1971	*Group A*	*Difference*
234	− 144 (IUD) =	−90
234	− 255 (Pill) =	+21

It is in relation to the 1966 results for group B, therefore, that the fertility of acceptors, regardless of the method of contraception used, has recorded a fairly significant decline. The decline was roughly 46.5 per cent for acceptors of IUD, as against only 5.2 per cent for acceptors of the pill.

On the other hand, the comparison of the observed rates for the experimental group with those of the control group in 1971 still yields a negative difference, though one smaller than that of 1966 for acceptors of IUD, inasmuch as the difference drops from −46.5 per cent to −38.5 per cent; and likewise in the case of pill acceptors, the difference becomes markedly positive (about +9 per cent) precisely because of a higher level of fertility (253) than that recorded for women in general in group B in 1971 (234).

Before commenting on this situation, which may seem anomalous and illogical, some consideration must be given to the distribution of the rates by age.

Figure VI shows the general profile of these rates by age groups. For control group B, there was a fairly substantial decline in fertility between 1966 and 1971, particularly among the youngest age groups. The distance between the two curves is reduced as the age rises, but the downward trend in fertility is systematic regardless of age group. For experimental group A, on the other hand, the curve tends to fluctuate according to the method of contraception. The curve for the IUD acceptors falls steadily as far as the 30–34 age group, rises to reflect the rate for women between the ages of 35 and 39 (130) to a higher level than that of the immediately preceding age group (120), and then resumes its downward slope. Another feature demonstrated by figure VI is that IUD acceptors owe their fall in fertility primarily to those among them in the 25–29 and 30–34 age groups. For acceptors of the pill, the curve follows a zigzag course. The negative difference that it shows in relation to the 1966 curve is, as can be clearly seen from the graph, almost entirely due to the very low fertility rate of the 30–34 age group. In

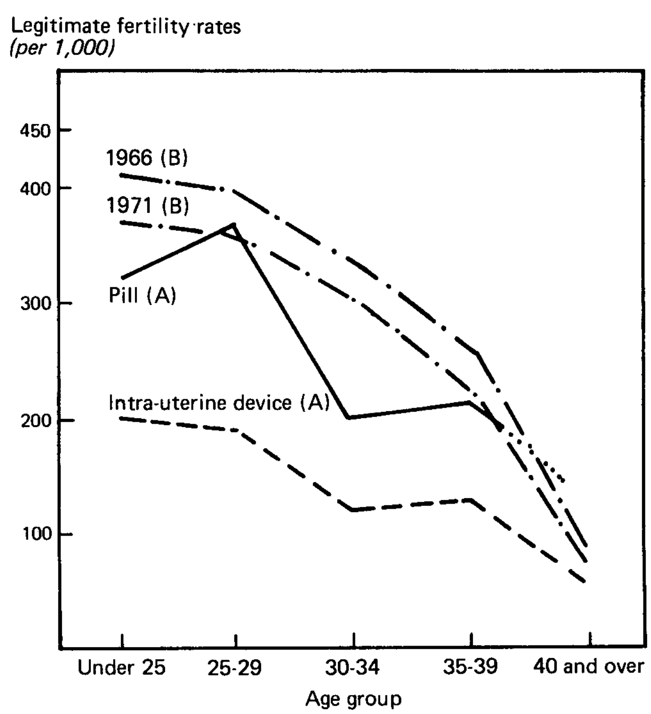


Figure VI. Experimental group A and control group B: legitimate fertility rates by age group

1971, pill acceptors between the ages of 25 and 29 have a slightly higher fertility level than women of the same age in the control group.

Lastly, the evidence shows that it is generally acceptors, whether of the pill or IUD, between the ages of 30 and 34 who contribute most to the general fall in the fertility level of their group, (A), compared with the level of women not practising contraception.

However, as has been shown, the distribution of the rates by age does not explain the situation in 1971, which has been described as anomalous.

The present authors hold that there are two main reasons for this finding. The first concerns the control population B itself. Although it can be regarded as genuinely constituting the total number of women experiencing natural fertility in 1966 (the official family planning programme was not launched until June 1965), it cannot be so regarded in 1971 (after five years of official and national practice of contraception in Tunisia) without underestimating the disparity between it and experimental group A, since it will inevitably contain a fairly large number of women using contraception.

It is therefore believed that the disparity found for 1971 in fact underestimates the true situation, which partially explains why the fertility level of acceptors of the pill is higher than that of group B.

The second reason concerns the experimental group itself and, more precisely, the continuation rate for each of the contraceptive methods used. As is demonstrated below, the continuation rate for IUD is much higher than that for the pill, which after only one year falls from unity to about 40 per cent. Because it is a birth-spacing contraceptive chiefly used by women who are young in the absolute sense and are of a younger average age than IUD acceptors, the fluctuating behaviour of the curve and the high level of fertility are both understandable.

The third reason, which is examined in greater detail in the following section, is associated with one exceptional characteristic of contraceptive users in general and of the experimental group in particular, namely, their very high fertility level.

Differences in fertility between experimental group A and control group C

This method is applied to the same group of women "before and after" acceptance of a method of contraception; its obvious merit is that the socio-economic, cultural and demographic characteristics are identical. Its drawback is, naturally, a chronological one, since it compares two situations, each of which relates to a different time-frame.

In this method, fertility rates before and after acceptance are compared at equal ages, using, for a given age group (e.g., 30–34) the average of "pre-acceptance" rates (see table 22). The result obtained characterizes the fertility of the group at "about age 30" and can be compared with the average of the "post-acceptance" fertility rates of women in the 25–29 age group, which likewise characterizes their fertility at "about age 30".

Estimates computed on the basis of the data in table 22 yield the average rates given in table 23.

Figure VII illustrates these estimates. Average fertility rates before and after acceptance are shown in figure VII (*a*); it reveals the decline in fertility among acceptors over an average interval of five and a half years. Figure VII (*b*) illustrates the difference, at comparable ages, between the pre- and post-acceptance fertility of the women.

A review of the results provided by this method suggests two important observations. First, the decline in fertility is definitely greater for acceptors of IUD than for acceptors of the pill, in the case of whom the

TABLE 22. FERTILITY RATES, BY AGE GROUP, OF WOMEN IN EXPERIMENTAL GROUP A BEFORE AND AFTER ACCEPTANCE OF CONTRACEPTION
(Rates per 1 000 women)

	Before acceptance				After acceptance		
Method and age group	Fifth year	Fourth year	Third year	Second year	Second year	Third year	Fourth year
Intra-uterine device							
Under 25	452	498	438	563	177	203	220
25-29	383	425	380	389	81	226	261
30-34	309	326	365	381	105	61	196
35-39	342	243	305	332	30	168	194
40 and over	210	167	131	292	66	45	52
Pill							
Under 25	445	524	380	552	373	303	272
25-29	326	394	371	434	365	369	362
30-34	369	342	419	370	213	105	281
35-39	307	244	270	303	249	221	171
40 and over	350	254	157	278	41	45	[a]

Source: Tunisia, Office national du planning familial et de la population, *Enquête nationale sur la continuation des methodes contraceeptives, 1973;* vol. II, *Exploitation et résultats* (Tunis, 1975).

[a] Insufficient data.

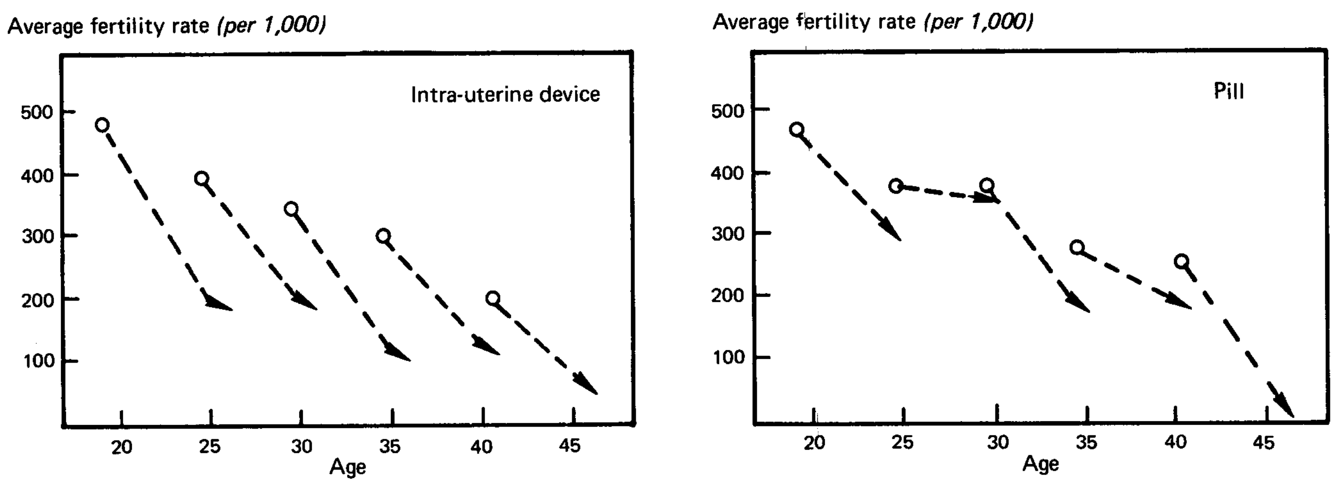
TABLE 23. AVERAGE FERTILITY RATES BEFORE AND AFTER ACCEPTANCE, BY AGE GROUP
(Rates per 1 000 women)

Method and age group	Average rates: Before acceptance	Average rates: After acceptance	Difference: Value	Difference: Corresponding age
Intra-uterine device				About
Under 25	488	200	194	25
25-29	394	189	156	30
30-34	345	121	185	35
35-39	306	131	69	40
45 and over	200	54	—	45
Pill				About
Under 25	475	316	65	25
25-29	381	365	10	30
30-34	375	200	81	35
35-39	281	213	47	40
40 and over	260	43	...	45

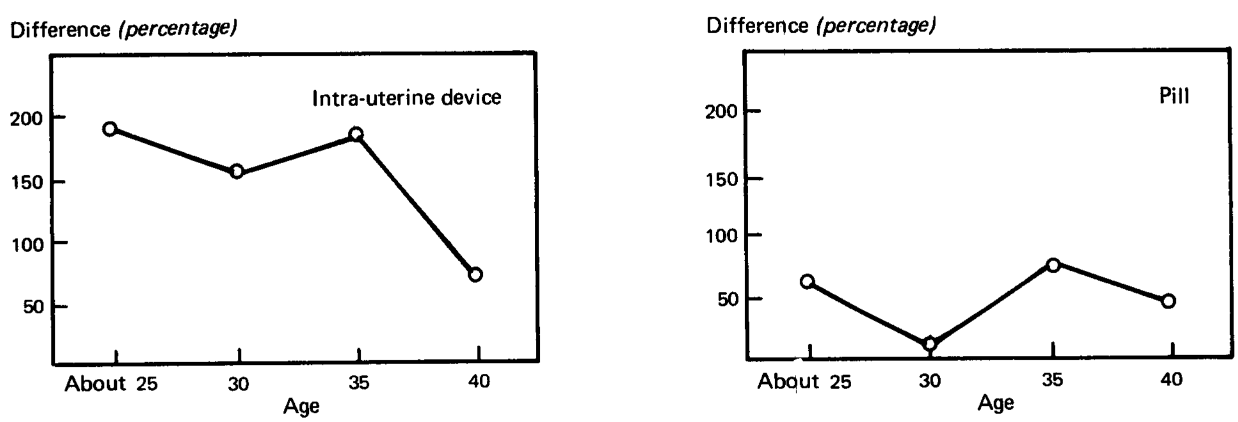
decline is very small at the level of the 25–29 (10) and 35–39 (47) age groups; secondly, the differences in fertility at comparable ages between contraceptive users (after acceptance) and non-users (before acceptance) are much smaller for pill acceptors (from 10 to 80 points) than for IUD acceptors (from 69 to 194 points).

Generally speaking, however—and contrary to the conclusion emerged from the comparison of group A with group B—the two groups of acceptors (pill and IUD) did show a real decline in fertility, compared with their pre-contraception status. Indeed, although acceptors of the pill had a higher general fertility rate than the control group in 1971, that was because the pill acceptors were women who were much more fertile than the average of women in general.

That such is the case provides further confirmation of the slight impact of oral contraception on fertility in Tunisia, as already indicated by the low rate of continuation for this form. Many plausible explanations could be advanced, and the situation could be the subject of valuable research.



(*a*) Average fertility rates before and after acceptance, by age group at time of acceptance



(*b*) Differences in fertility, at comparable ages, between acceptors and those who are not yet acceptors

Figure VII. Pre-acceptance and post-acceptance fertility rate and differences in fertility

Appraisal of the method

The experimental-design method has one indisputable advantage—it is simple to apply. It is, however, very inadequate because it is too crude in the way in which it measures the impact of family planning on fertility. Its inadequacy derives basically from two factors.

Selection of the experimental and control groups. The method is based on the full comparability of the two groups. However, it is extremely difficult, if not impossible, to find two population groups having identical socio-economic, cultural and demographic characteristics and differing only in the practice or non-practice of contraception.

In the example used, it was found that in the first case (groups A and B) there were significant social and economic differences between the populations compared; and although they differed in the fundamental characteristic of being acceptors or non-acceptors, illogical results were produced. On the other hand, the second case (groups A and C) appeared to have solved the problem of identical characteristics. In the authors' view, however, this was only apparently so, and the effect was due to the procedure of displacement over time that had to be applied in order to make the comparison.

In comparing the fertility of women of, say, ages 30–34 before acceptance with that of the 25–29 age group after acceptance, one is no longer comparing the same women even though this may appear to be the case, but two groups of women who are completely different (according to age group) and who consequently have different socio-economic, cultural and demographic characteristics.

Furthermore, in order to be significant, the average rates must be estimated over a long period (an average of five years). The contention here is that the behaviour of a woman at the beginning of this period necessarily differs from that of another woman, even if she has the same characteristics, at the end of the period, because such behaviour is linked to the social and—more importantly—the economic evolution of the country.

Exceptional nature of the experimental group

The second defect of the method is that it ignores the exceptionally fertile nature of women contraceptive users as compared to all non-users. It has been shown, after all, that acceptors of both IUD and the pill are both young (average ages of 30.8 and 30, respectively) and very fertile, especially during the year preceding acceptance (555 per 1,000 for acceptors of IUD and 485 per 1,000 for acceptors of the pill one year before acceptance, and 390 per 1,000 and 386 per 1,000, respectively, on average, for the five years preceding acceptance, as against a general fertility rate at the national level of 269 per 1,000 in 1966 and 234 per 1,000 in 1971). The method makes no adjustment or correction of this extra-high fertility level among acceptors and this factor, obviously, distorts the results and the conclusions.

Couple-years of protection index

The estimates

In applying the couple-years of protection method, use was made of the formula applied to Pakistan; but vasectomies, which are not performed under the Tunisian programme, were excluded, and two other methods, the pill and social abortion, were added.

Secondary methods

Data are available in Tunisia, not for the number of condoms or tubes of cream and jelly distributed, but for the average number of users per month. In the case of condoms and jelly, each acceptor is usually given the quantity considered necessary for one cycle, in other words, 12 condoms and one tube of jelly. This judgement is largely empirical, and there has been no study to verify its validity.

Family planning centres consider these methods to be temporary (either for a client who comes in half the way through her cycle or in the expectation of persuading the woman at her next visit to adopt a more effective method (IUD or pill)).

It is also assumed that, in order to provide protection for one year, each woman must be supplied with contraceptives for 13 cycles.

In 1971, there were, on average, 2,237 users of condoms and 401 users of jelly and cream per month; the corresponding number of couple-years of protection is:

$$\text{CYP}_{1971}\ (\text{condoms} + \text{jelly}) = \frac{(2{,}237 + 401) \times 12}{13}$$
$$= 2{,}435 \text{ couple-years of protection.}$$

Pill

The reasoning given above applies also to pill users. In 1971, there were, on average, 7,612 pill users per month and 96,570 packs were distributed during the year, which gives:

$$\text{CYP}_{1971}\ (\text{pill}) = \frac{96{,}570}{13}$$
$$= 7{,}427 \text{ couple-years of protection.}$$

In some centres, multicycle packs are issued, which accounts for the difference between the total number of packs and the number of users (7,612 × 12 = 91,344 users).

Tubal ligation

In the formula used for Pakistan, the total number of ligations is calculated from the start of the programme to year *n* which, in Tunisia, corresponds to 11,144 couple-years of protection in 1971 (see tables 24 and 25).

TABLE 24. WOMEN REMAINING *n* YEARS FOLLOWING LIGATION PER 1 000 WOMEN HAVING UNDERGONE LIGATION, ALL AGES COMBINED

n	*Women remaining*	*n*	*Women remaining*
1	933	15	101
2	867	16	75
3	793	17	50
4	718	18	34
5	645	19	27
6	572	20	20
7	498	21	14
8	429	22	7
9	372	23	4
10	313	24	3
11	253	25	2
12	196	26	1
13	150	27	0
14	126	28	0

It was deemed preferable to use the results of one demographer's estimate based on Tunisian data[22] and to allow for the following events which might have occurred in the absence of tubal ligation: (*a*) death between the ages of 25 and 50; (*b*) being widowed between the ages of 25 and 50; (*c*) divorce; (*d*) becoming sterile; (*e*) resort to other methods of contraception.

From table 24 and figure VIII, an estimate was made of the number of women having undergone ligation who remained in 1971, on the basis of the number of women having undergone the procedure since 1964 (see table 25).

Table 25 gives the number of women in 1971 who had undergone the procedure and hence the value of CYP:

CYP_{1971} (tubal ligation)
= 8,722 couple-years of protection.

Intra-uterine device

The Pakistan formula was adopted, but the coefficients were replaced with Tunisian estimates of the proportion of IUD acceptors who were still protected after one year, two years etc., following initial acceptance, as estimated by the 1973 continuation survey (table 26).

Rates for periods exceeding 48 months were determined by trend extrapolation (figure VIII). A figure for women remaining who continued to be protected by an IUD in 1971 was thus obtained (see table 27).

The value of CYP, therefore, is:

CYP_{1971} (IUD)
= 28,659 couple-years of protection.

The results

The sum of the number of couple-years of protection obtained by the various methods gives the value of the CYP index for 1971:

[22] See annex II, "Naissances évitées par les ligatures de trompes en Tunisie", prepared by L. Behar.

TABLE 25. RESIDUE IN 1971 OF WOMEN HAVING UNDERGONE LIGATION

Year	*Number of women having undergone ligation*	*Residue in 1971 of women having undergone ligation*	
		Per 1 000	*Total*
1964	293	429	126
1965	384	498	191
1966	766	572	438
1967	742	645	478
1968	1 627	718	1 168
1969	2 513	793	1 993
1970	2 539	867	2 201
1971	2 280	933	2 127
TOTAL	11 144		8 722

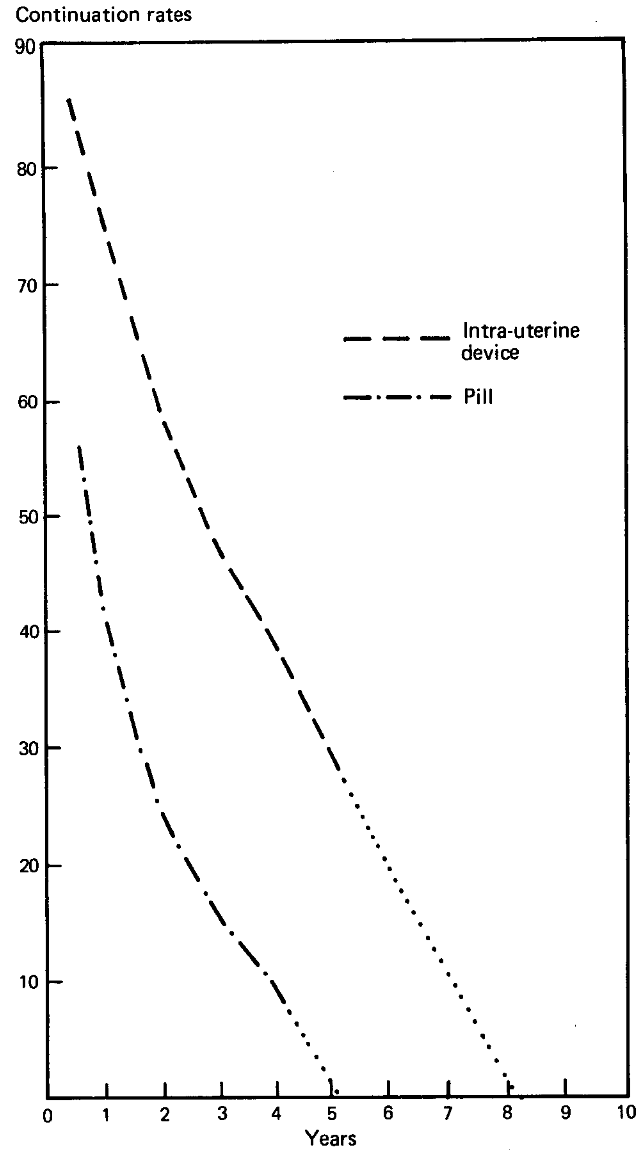


Figure VIII. Trend of continuation rates, first method, for acceptors of the pill and the intra-uterine device

TABLE 26. RATES OF CONTINUATION FOR ACCEPTORS OF AN INTRA-UTERINE DEVICE, AT 6, 12, 24, . . . 96 MONTHS

	Average period (months)								
	6	*12*	*24*	*36*	*48*	*60*	*72*	*84*	*96*
Rate given by the survey	84.9	74.6	58.5	46.7	38.3	28	18	10	0

$$CYP_{1971} = 2,435 + 7,427 + 8,722 + 28,659 = 47,243$$ couple-years of protection.

In order to express this value as births averted, it was assumed that in Tunisia 1 CYP = 0.25 birth averted, which would correspond to a legitimate fertility rate of 250 per 1,000, which is close to but about 10 per cent higher than that for all married women (15–54 years) in 1971 (234 per 1,000) and is the rate used by ONPFP for its evaluations.

On the basis of that formula, the number of births averted would be 11,810, not including births averted by the 3,197 abortions performed in 1971.

Appraisal of the method

The CYP method was found by the authors to be very crude and open to criticism at every level. Although is might be of some use to the authorities when comparing the results of programme activities, it is particularly misleading if it is assumed to represent the number of births genuinely averted and is therefore used to measure the impact of the family planning programme on fertility.

Estimating protection by the various methods

With respect to secondary methods, it was not possible to gather data as detailed as the reference material recommended. The number of units distributed, the number of units needed to protect a couple during one cycle and the percentage of contraceptives efficiently used were all unknown. The number of years of protection secured by these methods are certainly over-estimated. As stated above, clients are advised to use them only as temporary measures and they are only a marginal activity of the family planning centres.

For the pill, the present estimate of protection, which is based on the number of packs distributed, is also optimistic. The survey on the continuation of the IUD and the pill has demonstrated how little impact oral contraceptives have had on fertility in Tunisia: after one year, 60 per cent of the acceptors had abandoned that method. The level of protection is poor, since half of the acceptors become pregnant within two years and their pregnancy rate does not appear to be much lower than the rate for the population as a whole. It seems that for these women contraception provides little motivation or that they use it for spacing purposes, in order to defer further pregnancy for a few months or a year.

TABLE 27. WOMEN STILL PROTECTED BY AN INTRA-UTERINE DEVICE IN 1971, BY YEAR OF INSERTION

		Women remaining in 1971 and still protected by IUD	
Year	*Primary insertions*	*Rate per 1 000*	*Total*
1964	1 154	0	
1965	12 832	10	1 283
1966	12 077	18	2 174
1967	9 657	28	2 704
1968	9 304	38.3	3 563
1969	8 696	46.7	4 061
1970	9 638	58.5	5 638
1971	12 381	74.6	9 236
TOTAL	75 739		28 659

Tubal ligation

The CYP method can be considered satisfactory with respect to tubal ligations, even though protection for 1971 was over-estimated since ligations performed in January or December were assumed to provide the same protection over the year. Moreover, that assumption is equally applicable to preceding years. The method, in fact, is as good as the assumptions with regard to mortality, divorce, widowhood, sterility and use of contraception in the absence of ligation (see annex II).

These comments on ligation are also applicable to IUDs; the method of estimation adopted is similar. Moreover, the protection provided by the IUD also involves other variables—fertility, mortality, dissolution of unions, amenorrhoea—about which no accurate information is available.

The reasons for discontinuation which were taken into account in estimating continuation rates are as follows: (*a*) pregnancy; (*b*) expulsion; (*c*) withdrawal on medical grounds; (*d*) withdrawal for other reasons. The data in table 28 show that withdrawal because of "pain" or "bleeding" (i.e., medical grounds) is the most frequent reason for discontinuation.

Abortions For the 3,197 abortions, it would be better to use the method of births averted described below in the subsection on the ONPFP method, in order to estimate directly the number of births averted in 1971, taking into account the fact that the interruption of a pregnancy causes a woman to become fertile earlier, thereby diminishing the apparent demographic effectiveness of the operation.

This method of evaluating the impact of the family planning programme presupposes that the years of protection are additive and that the contraceptive methods are independent of one another. It ignores the indirect effects of the programme and takes no account of non-programme contraception through private doctors and pharmacies.

TABLE 28. CUMULATIVE RATE OF DISCONTINUATION BY ACCEPTORS OF AN INTRA-UTERINE DEVICE, BY CAUSE, AT 6, 12, 24, 36 AND 48 MONTHS[a]

Reason for discontinuation	*Number of months* 6	12	24	36	48
Pregnancy	2.3	3.2	4.0	4.4	4.8
Expulsion	4.5	6.6	10.5	10.9	12.4
Withdrawal on medical grounds	6.9	14.0	22.5	30.9	34.6
Withdrawal for other reasons	2.3	3.9	6.5	8.2	10.2
Rate of discontinuation	16.0	27.8	43.4	54.3	62.1

Source: See table 22.
[a] Only those acceptors were surveyed.

Abandoning the first method used does not mean that a couple ceases to practise contraception in general and, in the continuation survey, consideration was also given to the continuation of "all contraception" (see figure IX).

Table 29 shows what proportion of the women who have ceased using the IUD or the pill have adopted other methods, indicating which method.

Over all, only one third of the women who stop using the IUD continue to practise contraception, and half of those who continue use the pill; the other methods practised are, in order of use, the "natural methods" and condoms, jelly and cream. Among acceptors of the pill, three quarters of those who discontinue using it give up all contraception (at least until the next pregnancy). Those who continue use natural methods or condoms, jelly and cream in preference to the IUD.

Lastly, the conversion of the CYP index into number of births averted is very questionable since the

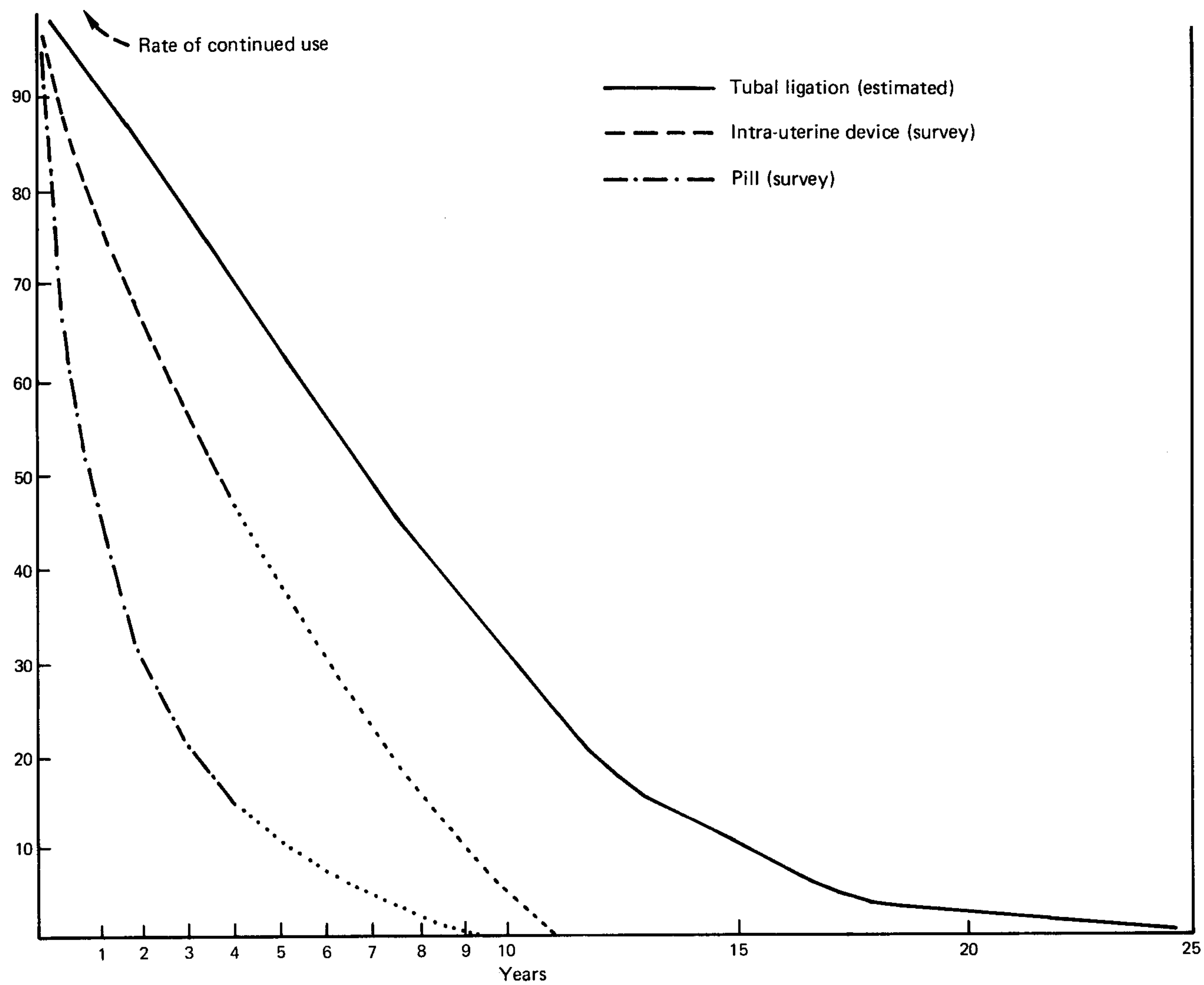


Figure IX. Evolution of rates of continued use of the intra-uterine device and the pill (all contraception) and tubal ligation.

potential fertility of acceptors appears, in fact, to be significantly higher than that of married women in general. This argument is set forth in detail in the next section of this report.

TABLE 29. WOMEN HAVING CEASED TO USE THE FIRST METHOD, BY METHOD ADOPTED, IF ANY, PRIOR TO PREGNANCY, IF ANY
(*Percentage*)

Method adopted	*Intra-uterine device acceptors*	*Pill acceptors*
No method	66.5	76.0
Intra-uterine device	—	5.5
Pill	15.3	—
Condom, jelly, cream	5.1	5.9
Tubal ligation	2.9	1.4
Natural method	6.4	7.3
Traditional method	2.9	2.5
Another method	1.0	1.1
Not stated	—	0.2
TOTAL	100.0	100.0
NUMBER	(313)	(438)

Source: See table 21.

Component projection approach

The estimates

It is assumed for the purposes of this discussion of the component projection approach that in order to avert births in 1971, a couple must have been protected by a contraceptive method from June 1970 to May 1971. The following calculations therefore relate to each age group efficiently practising contraception during that period.

Intra-uterine device

An attempt is made to estimate the number of couples still living in the same union and remaining fertile in which the wife is using an IUD. This estimate was made by wife's age group at the end of 1970, which was taken to be the mid-point of the period (June 1970-May 1971).

For each earlier year (from 1964, when the programme began, to 1970), data are available on primary insertions of IUDs, broken down by wife's age group at the time of acceptance. This distribution, by percentage and in absolute figures, is given in table 30.

In addition, the 1973 continuation survey provides data on the rate of continuation by age group (table 31). It was assumed in the survey that continuation is uniform within age groups; that IUDs inserted in 1970 had been in use an average of six months as of the end of 1970; and that the same applies to earlier years.

Overall, continuation improves with age, as might be expected. Although the rise is steady up to age 30–34, the pattern is less clear thereafter: women in the 35–39 age group show a lower rate of continuation than those in the 30–34 age group, and women aged 40 years and over show a higher rate of continuation only after 18 months. By applying these rates to all women who have had an IUD inserted between 1964 and 1970, it is possible to determine by age group the number of women who at the time of insertion will continue to use the IUD in the absence of any other disruptive factor in their married life.

No data were available for continuation after 42 months and as extrapolation proved to be a very risky process, it appeared to be appropriate to make the following hypotheses.

The continuation rate for all ages combined is applied to all women who have had an IUD insertion (i.e., 32.5 per cent after 54 months for 1966, 25 per cent after 66 months for 1965 and 14.5 per cent after 78 months for 1964). See table 24. It was assumed that for 1965 and 1966, women continuing to use an IUD were distributed as follows: 25–29 years, one fifth; 30–34 years, two fifths; and 35–39 years, two fifths. For 1964, the following distribution was assumed: 30–34 years, three quarters; and 35–39 years, one quarter. This produces the results given in table 32.

In order to take account of the risks to which the women in question may have been exposed between 1964 and 1970, data on tubal ligations were obtained from the study by L. Behar (see annex II), in which the following events are considered: deaths occurring between the ages of 25 and 50; widowhood between the ages of 25 and 50; divorce; becoming sterile; and resorting to another method of contraception. It will be noted, however, that the probability of resorting to another method of contraception would certainly be lower for women who have accepted IUDs than for women who have undergone tubal ligation. That hypothesis would therefore lead to slightly underestimating the number of women remaining. Each age group is represented by its mean age, and it is assumed that the average age at time of insertion is 23 for the age group under 24 and 42.5 for the age group 40 and over.

With the assistance of table 37, which gives the data for remaining women n months after a tubal ligation, expressed as a percentage, the number of women remaining who used an IUD at the end of 1970, distributed by age group at the time of insertion, was estimated (see table 33).

In determining the distribution of these persons by age group at the end of 1970, it was assumed that the distribution was uniform within age groups and that the total number of IUDs actually in use at the end of 1970 among women who were assumed to be in the same age group in 1971 would therefore be given by:

$$
\begin{aligned}
Q_i\text{, } 1971 = {} & 0.90\, q_i\text{, } 1970 + 0.10\, q_i - 1\text{, } 1970 \\
& + 0.70\, q_i\text{, } 1969 + 0.30\, q_i - 1\text{, } 1969 \\
& + 0.50\, q_i\text{, } 1968 + 0.50\, q_i - 1\text{, } 1968 \\
& + 0.30\, q_i\text{, } 1967 + 0.70\, q_i - 1\text{, } 1967 \\
& + 0.10\, q_i\text{, } 1966 + 0.90\, q_i - 1\text{, } 1966 \\
& + 0.90\, q_i - 1\text{, } 1965 + 0.10\, q_i - 2\text{, } 1965 \\
& + 0.70\, q_i - 1\text{, } 1964 + 0.30\, q_i - 2\text{, } 1964
\end{aligned}
$$

This method was used to derive the number of

TABLE 30. DISTRIBUTION OF WOMEN WHO HAVE HAD A PRIMARY INTRA-UTERINE DEVICE INSERTION, BY AGE GROUP AND DATE OF INSERTION

Age group	*1964*	*1965*	*1966*	*1967*	*1968*	*1969*	*1970*
			(a) *Rate per 1 000 women*				
Under 24	120.6	142.0	96.4	95.7	150.3	163.0	176.3
25-29	223.1	272.9	204.0	195.5	243.6	246.5	239.9
30-34	305.8	281.2	296.6	294.4	284.6	280.2	275.4
35-39	218.2	204.9	290.6	314.1	240.0	242.8	233.4
40 and over	132.3	99.0	112.4	100.3	81.5	67.5	75.0
TOTAL	1 000.0	1 000.0	1 000.0	1 000.0	1 000.0	1 000.0	1 000.0
			(b) *Absolute number*				
Under 24	139	1 822	1 164	924	1 398	1 417	1 699
25-29	257	3 502	2 464	1 888	2 267	2 144	2 312
30-34	353	3 608	3 582	2 843	2 648	2 437	2 654
35-39	252	2 630	3 510	3 033	2 233	2 111	2 250
40 and over	153	1 270	1 357	969	758	587	723
TOTAL	1 154	12 832	12 077	9 657	9 304	8 696	9 638

Source: Tunisia, Office du planning familial et de la population, *Caracteristiques des acceptrices de DIU de 1964 à 1969* (Tunis, 1972).

women who were efficiently practising contraception at the end of 1970, by age group at the end of 1970, as an average over the period June 1970-May 1971, and who would therefore prevent births in 1971. Table 34 gives the final results of those calculations.

TABLE 31. RATES OF CONTINUATION, FIRST METHOD, OF WOMEN ACCEPTING AN INTRA-UTERINE DEVICE, BY AGE GROUP

Age group	*Six months*	*Eighteen months*	*Thirty months*	*Forty-two months*
15-24	79.7	54.3	38.5	29.2
25-29	82.2	64.3	50.1	39.7
30-34	90.2	74.4	61.4	50.0
35-39	87.0	69.7	54.1	44.8
40 +	84.8	72.4	64.6	56.9
TOTAL	84.9	66.5	52.6	42.5

Tubal ligation

The same estimating principles were applied to tubal ligations. However, since the distributions of women by age at the time of ligation was not known for all years, the 1974 distribution (see table 35) was applied to all women who underwent tubal ligation between 1964 and 1970. The distribution of those women by age group would therefore be as shown in table 36.

Assuming that the necessary conditions obtained for applying the method of estimating the number of women remaining in the sample *n* months after a ligation (see annex II), an estimate was made, on the basis of the data in table 37, of the number of women still protected by ligation at the end of 1970; women in the group aged 45 years and over were presumed to have an average age of 47 at the time of ligation (see table 38).

Using the same assumption for distribution by age group as in the case of IUDs, it is possible to estimate, by age group at the end of 1970, the number of women who were protected by tubal ligation at the end of 1970 and who thus prevented births in 1971 (table 39).

Condoms and jelly

An estimate was made of the average numbers of users per month of condoms and jelly, during the period June 1970-May 1971 (table 40). It was assumed that 80 per cent of them were practising contraception efficiently. These contraceptives are, as a rule, supplied for one cycle at a time. For the distribution of users (table 41), the distribution IUD insertions by age group in 1970 was used (see table 30(*a*)).

These couples were assumed to be living in the same union and to be fertile throughout the period.

The pill

The method used to estimate the number of users efficiently practising this form of contraception between June 1970 and May 1971 is very similar to the

TABLE 32. WOMEN CONTINUING TO USE AN INTRA-UTERINE DEVICE AT END-1970 (IN THE ABSENCE OF OTHER DISRUPTIVE FACTORS), BY AGE GROUP AT TIME OF INSERTION

Age group	*1964*	*1965*	*1966*	*1967*	*1968*	*1969*	*1970*
24 and under	–	–	–	270	538	769	1 354
25-29	–	642	785	749	1 136	1 379	1 900
30-34	125	1 283	1 570	1 421	1 626	1 813	2 394
35-39	42	1 283	1 570	1 359	1 208	1 471	1 957
40 and over	–	–	–	551	490	425	613
TOTAL	167	3 208	3 925	4 350	4 998	5 857	8 218

TABLE 33. WOMEN REMAINING IN SAMPLE AND STILL USING AN INTRA-UTERINE DEVICE AS OF END-1970, BY AGE GROUP

Age group	*1964*	*1965*	*1966*	*1967*	*1968*	*1969*	*1970*
24 and under	–	–	–	245	505	742	1 338
25-29	–	530	669	658	1 033	1 302	1 864
30-34	92	1 002	1 297	1 236	1 483	1 719	2 353
35-39	23	797	1 100	1 061	1 033	1 347	1 902
40 and over	–	–	–	307	339	348	576
TOTAL	115	2 329	3 066	3 507	4 393	5 458	8 033

previous method, but a number of remarks need to be made since it takes into account the continued use of the pill.

A recent study[23] of the period 1970–1973 shows that data obtained by applying an adjusted continuation table to the number of acceptors correspond, with very minor disparities, to the number of former registered users. Statistics are available on the average number of users per month, the number of new acceptors per month and the number of packs distributed. Because, in some centres, users are given a supply for several cycles (usually three), a distinction was made between users and packs; then, for each month, women protected by a one-cycle supply of pills were regarded as users for that month, and to them were added users for the previous two months who still had one cycle's supply but had not visited the centre. For December 1970, for example, the figures would be determined as follows:

Users for December 1970 7,342
Difference between number of packs
and number of users:
November 1970 199
October 1970 201
Women protected by a one-cycle supply of pills in
December 1970:

$$7{,}342 + \left(\frac{199 + 201}{2}\right) = 7{,}542$$

This calculation was made for each month from June 1970 to May 1971 in order to estimate the average number of women who were protected by the pill during that period (see table 42).

As was done with respect to condoms and jelly, it was assumed that 80 per cent of those women actually used their supply for the cycle (no information was available on the subject, but it does appear from some surveys that many women do not use their pack of pills or use them incorrectly): 7,003 × 80 per cent = 6,443. On the basis of this calculation, 6,443 women were actually protected by the pill during the period June 1970-May 1971.

It was assumed that 50 per cent of the women were one-year users, that 40 per cent has been users for from one to two years and that 10 per cent had been users for over two years. The age distribution at the time of acceptance (table 43) is that given by the continuation survey sample.

On the basis of these assumptions as to the distribution of women actually protected by the pill over time and by age group at time of acceptance, the method used for IUDs was again utilized to estimate by age group the number of women protected by the pill between June 1970 and May 1971. The results are given in table 44.

Since the age group distribution of women actually protected by the pill was derived *a posteriori,* women leaving the sample by reason of age were replaced, in the estimation process used, in the group aged 40 years and over (see table 45).

For social abortions, the number of births averted was estimated by a specific method described in the next section of this paper.

The results

Data are thus available for women effectively protected by IUD, tubal ligation, condoms and jelly, and pills by average age groups for the period from June

[23] A. Marcoux, "Continuation de la pillule d'après les statistiques de service," Tunis, Population Council, 1974 (mimeographed).

TABLE 34. DISTRIBUTION OF WOMEN REMAINING IN SAMPLE AND USING AN INTRA-UTERINE DEVICE BY AGE GROUP AT END-1970

Age group	*1964*	*1965*	*1966*	*1967*	*1968*	*1969*	*1970*	*Total number of women remaining, end-1970*
Under 25	–	–	–	73	252	519	1 205	2 049
25-29	–	–	66	369	769	1 134	1 811	4 149
30-34	–	477	733	832	1 258	1 594	2 304	7 198
35-39	64	955	1 277	1 183	1 258	1 459	1 947	8 143
40 and over	44	817	990	835	686	648	709	4 729
TOTAL	108	2 249	3 066	3 292	4 223	5 354	7 976	26 268
Women leaving the sample as a result of age	7	80	–	215	170	104	57	
TOTAL	115	2 329	3 066	3 507	4 393	5 458	8 033	

TABLE 35. NUMBER OF WOMEN WHO UNDERWENT TUBAL LIGATION IN 1974, BY AGE GROUP

Age group	Women who underwent tubal ligation in 1974	
	Number	*Percentage*
Under 25	108	1.5
25-29	701	10.0
30-34	1 931	27.5
35-39	2 733	38.9
40-44	1 361	19.3
45 and over	197	2.8
Not stated	403	
TOTAL	7 434	100.0
Average age	35.13	

Source: Tunisia, Office du planning familial et de la population.

1970 to May 1971, who would therefore avert births during 1971 (table 46).

In order to estimate the number of births averted by these women, it was decided to apply as potential fertility the fertility of IUD and pill acceptors in the five years preceding acceptance as assessed on the basis of the 1973 continuation survey.[24] The year preceding acceptance was excluded, as it was clear that acceptance of contraception often occurred in the year following a birth (or an abortion), so that fertility rates were particularly high in the year preceding acceptance. Therefore it was assumed that on average, acceptors in the 30–34 age group accepted at the mean age of 32.5 and that the average of the four pre-acceptance fertility rates would represent their fertility at "about age 30" (see table 47).

In those estimates, that fertility rate was applied to women aged 30–35 at the end of 1970, thereby avoiding an over-estimation of births (see table 48).

For women protected by tubal ligation, the potential fertility of pill acceptors, who are assumed to be very fertile, was used. For women protected by condoms and jelly, the potential fertility rates of IUD acceptors were used.

To the total given in table 48 must be added the number of births (1,817) averted by social abortion, as estimated by the method described in the following section. The total number of births averted in 1971 is thus: 13,876 + 1,817 = 15,693.

[24] These rates are given above in the subsection on experimental designs.

Social abortions

Marcoux argues that although an abortion prevents a birth, the woman concerned becomes fertile again at an earlier date, thereby reducing the demographic effectiveness of the operation. The only information available on the distribution by duration of terminated pregnancies yields an estimated average duration of two months.[25]

Some of these pregnancies would not have run their full term because of spontaneous abortion. The number involved is small, but it could be said that an intentional abortion averts approximately 0.94 birth (90 of 96 pregnancies which reached the two-month point would have run their full term).[26]

The additional fertile period averages seven months but since the average duration of pregnancy is 8.5 months, the additional fertile period is in fact 6.5 months.

The interval until ovulation resumes is taken to be one month after an abortion, and two months after childbirth in the absence of breast-feeding. In order to allow for this factor and for the protection afforded by post-abortum and post-partum IUD insertion, ovulation is regarded as recommencing, on average, six months after confinement.

Induced abortions, compared with the normal course of pregnancy, therefore give rise to an additional 6.5 + 5 = 11.5 fertile months. For every 100 women, this represents 95.8 fertile years.

Given the distribution by marital status (96.6 per cent married women) and by age (29) of women for whom information is available, Marcoux estimates that a population of the same structure has a fertility rate of approximately 290 per 1,000.

At this rate, 27.7 births will occur for every 95.8 woman-years, i.e., 0.277 birth per woman. In short, therefore, an abortion averts: 0.940 − 0.277 = 0.66 birth.

To estimate births averted in 1971, therefore, the number of social abortions carried out under the family planning programme between June 1970 and May 1971 (see table 49) was used, with the following result: 2,753 (social abortions) × 0.66 = 1,817 births averted.

[25] A. Marcoux, "Naissances évitées par les avortements", Tunis, Population Council, 1973 (mimeographed). Also T. B. Ben Cheikh, "L'experience tunisienne de l'avortement provoqué".

[26] *Foetal, Infant and Early Childhood Mortality: vol. I. The Statistics* (United Nations publication, Sales No. 54.IV.7).

TABLE 36. DISTRIBUTION OF WOMEN WHO UNDERWENT TUBAL LIGATION DURING 1964-1970, BY AGE-GROUP DISTRIBUTION IN 1974

Age group	1964	1965	1966	1967	1968	1969	1970
Under 25	4	6	12	11	24	38	38
25-29	29	38	77	74	163	251	254
30-34	81	106	210	204	447	691	698
35-39	114	149	298	289	633	978	988
40-44	57	74	148	143	314	485	490
45 and over	8	11	21	21	46	70	71
TOTAL	293	384	766	742	1 627	2 513	2 539

TABLE 37. WOMEN REMAINING IN SAMPLE *n* MONTHS AFTER TUBAL LIGATION
(Rates per 1 000 women)

Average age at the time of ligation	Number of months: 6	18	30	42	54	66	78
23	988	965	939	909	879	849	820
27.5	981	944	909	878	852	825	797
32.5	983	948	912	870	826	781	736
37.5	972	916	855	781	701	621	541
42.5	940	819	693	557	417	279	139
47	833	500	166	–	–	–	–
All ages combined	966	900	830	755	681	608	500

Source: Annex II, "Naissances évitées par les ligatures de trompes en Tunisie", prepared by L. Behar.

TABLE 38. NUMBER OF WOMEN REMAINING IN SAMPLE AT END-1970 WHO HAD UNDERGONE TUBAL LIGATION, BY AGE GROUP

Age group	1964	1965	1966	1967	1968	1969	1970
Under 25	3	5	10	10	22	37	37
25-29	23	31	66	65	148	237	249
30-34	60	83	173	177	408	655	686
35-39	62	92	209	226	541	896	960
40-44	8	21	62	80	218	397	461
45 and over	–	–	–	–	8	35	59
TOTAL	156	232	520	558	1 345	2 257	2 452

TABLE 39. DISTRIBUTION OF WOMEN REMAINING IN SAMPLE WHO HAD UNDERGONE TUBAL LIGATION, BY AGE GROUP AT END-1970

Age group	1964	1965	1966	1967	1968	1969	1970	Total remaining, end-1970
Under 25	–	–	1	3	11	26	33	74
25-29	2	4	16	26	85	177	228	538
30-34	17	29	76	99	278	529	642	1 670
35-39	49	78	177	192	474	824	933	2 727
40-44	61	91	194	182	380	547	511	1 966
45 and over	25	28	56	56	113	143	99	520
TOTAL	154	230	520	558	1 341	2 246	2 446	7 495
Women leaving the sample as a result of age	2	2	0	0	4	11	6	
TOTAL	156	232	520	558	1 345	2 257	2 452	

TABLE 40. AVERAGE NUMBER OF USERS OF CONDOMS AND JELLY PER MONTH
JUNE 1970-MAY 1971

1970	Condom	Jelly	1971	Condom	Jelly
June	2 393	374	January	2 339	326
July	2 291	432	February	1 880	218
August	2 044	276	March	2 891	315
September	2 233	284	April	2 526	400
October	2 233	342	May	2 531	465
November	2 219	220	TOTAL:		
December	2 520	353	June 1970-May 1971	28 100	4 005

Note: Average number of users (June 1970-May 1971): 2 342 + 334 = 2 676.
Number of users efficiently practising contraception by means of condoms and jelly: 2 676 × 80 per cent = 2 141 users.

TABLE 41. DISTRIBUTION OF USERS PROTECTED BY CONDOMS AND JELLY, BY AGE GROUP, JUNE 1970-MAY 1971

Age group	*Number of users protected*
Under 25	377
25-29	514
30-34	590
35-39	500
40 and over	160
TOTAL	2 141

TABLE 42. WOMEN PROTECTED BY THE PILL, PER MONTH, JUNE 1970-MAY 1971

Year and month	*Users who came to family planning centre*	*Additional packs distributed*	*Women protected by a pill cycle for the month of:*
1970			
June	6 528	245	6 625[a]
July	6 354	100	6 540
August	5 898	171	6 070
September	6 024	119	6 159
October	6 621	199	6 766
November	6 429	201	6 588
December	7 342	187	7 542
1971			
January	6 945	149	7 139
February	6 435	365	6 603
March	8 141	478	8 398
April	7 308	405	7 729
May	7 441	482	7 882
Monthly average	6 789		7 003

[a] For June 1970, half of the additional packs distributed in April 1970 and May 1970 (i.e., 97) must be added.

TABLE 43. PERCENTAGE DISTRIBUTION OF PILL ACCEPTORS BY AGE GROUP AND AVERAGE AGE AT TIME OF ACCEPTANCE, ACCORDING TO CONTINUATION SURVEY, 1973

Age group	*Distribution of acceptors*
Under 25	20.4
25-29	22.0
30-34	25.7
35-39	22.1
40 and over	9.8
All groups	100.0
Average age	30.0

TABLE 44. WOMEN ACTUALLY PROTECTED BY THE PILL BETWEEN JUNE 1970 AND MAY 1971, BY AGE GROUP AT TIME OF ACCEPTANCE

Age group	*June 1968-May 1969*	*June 1969-May 1970*	*June 1970-May 1971*
Under 25	131	526	657
25-29	142	567	709
30-34	166	662	828
35-39	143	570	711
40 and over	63	252	316
All ages	645	2 577	3 221

Appraisal of the method

The component projection approach has great merit since it attempts far more than the methods previously discussed, to incorporate accurate indexes of the effectiveness of contraception. However, to expect to obtain a much more scientific result appears to be over-ambitious, as despite its refinements and perhaps precisely because of them, the method demands a set of data which is impossible to collect and numerous assumptions which are difficult to verify.

Data on the characteristics of acceptors, which are needed to assess the volume of contraceptives really being used, are not readily available; even if they were available on age distribution at time of acceptance per year or on age distribution of users at a given date, the efficiency of contraceptive practice at the level of the couple would still have to be assessed.

These difficulties, which arise from the need for a high degree of accuracy in applying the method, are heightened still further in developing countries—where, generally speaking, official family planning programmes are in existence—because the collection and processing of statistical data are rudimentary. Tunisia is a very favourable case, inasmuch as it has a relatively elaborate statistical organization, particularly for family planning activities.

By conducting surveys of specific problems, more detailed data can, of course, be obtained on certain aspects of the use of contraceptives, but the question arises whether the required investment of effort and financial resources would really be warranted by the importance of the results. Generally speaking, the authorities concerned are interested in action rather than theory.

At all levels of application of the method, assump-

TABLE 45. WOMEN ACTUALLY PROTECTED BY THE PILL BETWEEN JUNE 1970 AND MAY 1971, BY AVERAGE AGE GROUP AT END-1970

Age group	*Year of acceptance*			*Women protected by pill at end-1970*
	June 1968-May 1969	*June 1969-May 1970*	*June 1970-May 1971*	
Under 25	65	368	591	1 024
25-29	137	555	704	1 396
30-34	154	633	816	1 603
35-39	154	598	723	1 475
40 and over	135	423	387	945
All ages	645	2 577	3 221	6 443

TABLE 46. WOMEN PROTECTED BY ALL METHODS BETWEEN JUNE 1970 AND MAY 1971

Age group	Contraceptive method				
	Intra-uterine device	*Tubal ligation*	*Condoms and jelly*	*Pill*	*All methods*
24 and under	2 049	74	377	1 024	3 524
25-29	4 149	538	514	1 396	6 597
30-34	7 198	1 670	590	1 603	11 061
35-39	8 143	2 727	500	1 475	12 845
40 and over	4 729	2 486	160	945	8 320
All ages	26 268	7 495	2 141	6 443	42 347

TABLE 47. POTENTIAL FERTILITY RATES OF CONTRACEPTIVE USERS

Age group	Potential fertility	
	Acceptors of intra-uterine devices, condoms and jelly	*Acceptors of pills and tubal ligation*
About 20	488	475
About 25	394	381
About 30	345	375
About 35	305	281
About 40	200	260
All ages	349	361

tions had to be made which influenced assessment of the programme impact quite considerably. To change their content would alter, for instance, the relative weight of the individual contraceptive methods. Thus, for condoms, jelly and pills, the theoretical base used here is very precarious and questionable. The continuation rates applied for IUD insertions were drawn from a survey which covered a four-year period; beyond that period, extrapolations were made which would have to be verified.

Lastly, as regards both the potential fertility of users and the probability that they used other contraceptive methods, the deductions made here are based on an unreal situation about which the authors will never have full knowledge: the rates adopted for each contraceptive method can be justified only by a subjective judgement. If a potential fertility rate approximately 20 per cent higher than the legitimate fertility rate for 1966 (base year where the behaviour of couples was natural) had been used, a total of 15,453 births averted (at a rate of 322 per 1,000) would have resulted. A result very close to that of the authors was obtained by a very crude calculation; there are therefore grounds for asking whether it is really useful and profitable to devote a great deal of time and thought to producing such results.

Estimating births averted by the method used by the National Family Planning and Population Office, Tunis

Before discussing the main methodological issues raised by these methods, it should be mentioned that ONPFP has adopted a method of estimating births averted which is quite similar to the component projection approach, but more general in that it does not introduce the age distribution of the women protected.

The continuation rates adopted (see figure X) were established before the survey on continuation of the IUD and the pill was undertaken. For IUDs in particular, the degree of underestimation is sizable, but no allowance is made for various hazards (mortality, widowhood etc.).

For tubal ligation, on the other hand, the number of births averted was over-estimated in relation to the estimates. The rates for proportions of women remaining were drawn from a study made in another country and should also be re-examined.

Lastly, potential fertility was estimated at an over-all rate of 250 per 1,000, i.e., at a rate close to the over-all legitimate fertility rate in the absence of contraception.

For abortions, the estimate of births prevented considerably underestimates lactation amenorrhoea, i.e., four abortions prevent three births.

Estimating process

The number of women protected at the end of 1970 are estimated as follows:

(*a*) Protection by IUD, by applying the continuation rates shown in figure X, 23,710 women;

(*b*) Protection by tubal ligation, by applying the survivorship rates of figure X, 8,520 women;

TABLE 48. BIRTHS AVERTED IN 1971 BY WOMEN PROTECTED BY AN INTRA-UTERINE DEVICE, TUBAL LIGATION, CONDOMS AND JELLY, AND PILL, BY AGE GROUP OF THE WOMEN AT END-1970

Age group	Births averted by				
	Intra-uterine device	*Tubal ligation*	*Condoms and jelly*	*Pill*	*All methods*
24 and under	999	35	183	486	1 703
25-29	1 635	205	202	532	2 574
30-34	2 483	626	203	601	3 913
35-39	2 484	766	152	414	3 816
40 and over	946	646	32	246	1 870
All ages	8 547	2 278	772	2 279	13 876

TABLE 49. DISTRIBUTION OF SOCIAL ABORTIONS CARRIED OUT BETWEEN JUNE 1970 AND MAY 1971, BY MONTH

Year and month	*Number of social abortions*	*Year and month*	*Number of social abortions*
1970		1971	
June	251	January	241
July	237	February	165
August	211	March	281
September ...	242	April	279
October	216	May	250
November ...	163	TOTAL	2 753
December ...	217		

(*c*) Protection by the pill, by applying the method used, except for the month of December 1970; 7,542 women (users in December plus half the additional packs for October and November);

(*d*) Protection by condoms and jelly, by taking the November and December users namely, 5,312 women with a potential fertility rate of 250 per 1,000: 11,266 births averted by IUD, ligation, pill, condom and jelly, to which must be added 2,064 births averted by social abortion (2,753 abortions between June 1970 and May 1971).

This calculation results in the total: 13,330 births averted in 1971.

METHODOLOGICAL ISSUES

The aim of the study was to test various techniques for measuring the impact of family planning programmes on fertility by applying them to the particular case of Tunisia; it therefore appears useful to gather all the criticisms and observations which emerged in the preceding section and to generalize from them. The discussion covers the following aspects:

(*a*) Potential fertility of acceptors;

(*b*) Availability of data;

(*c*) Interaction of factors;

(*d*) Uncontrolled variables;

(*e*) Independence of method;

(*f*) Cost-precision analysis.

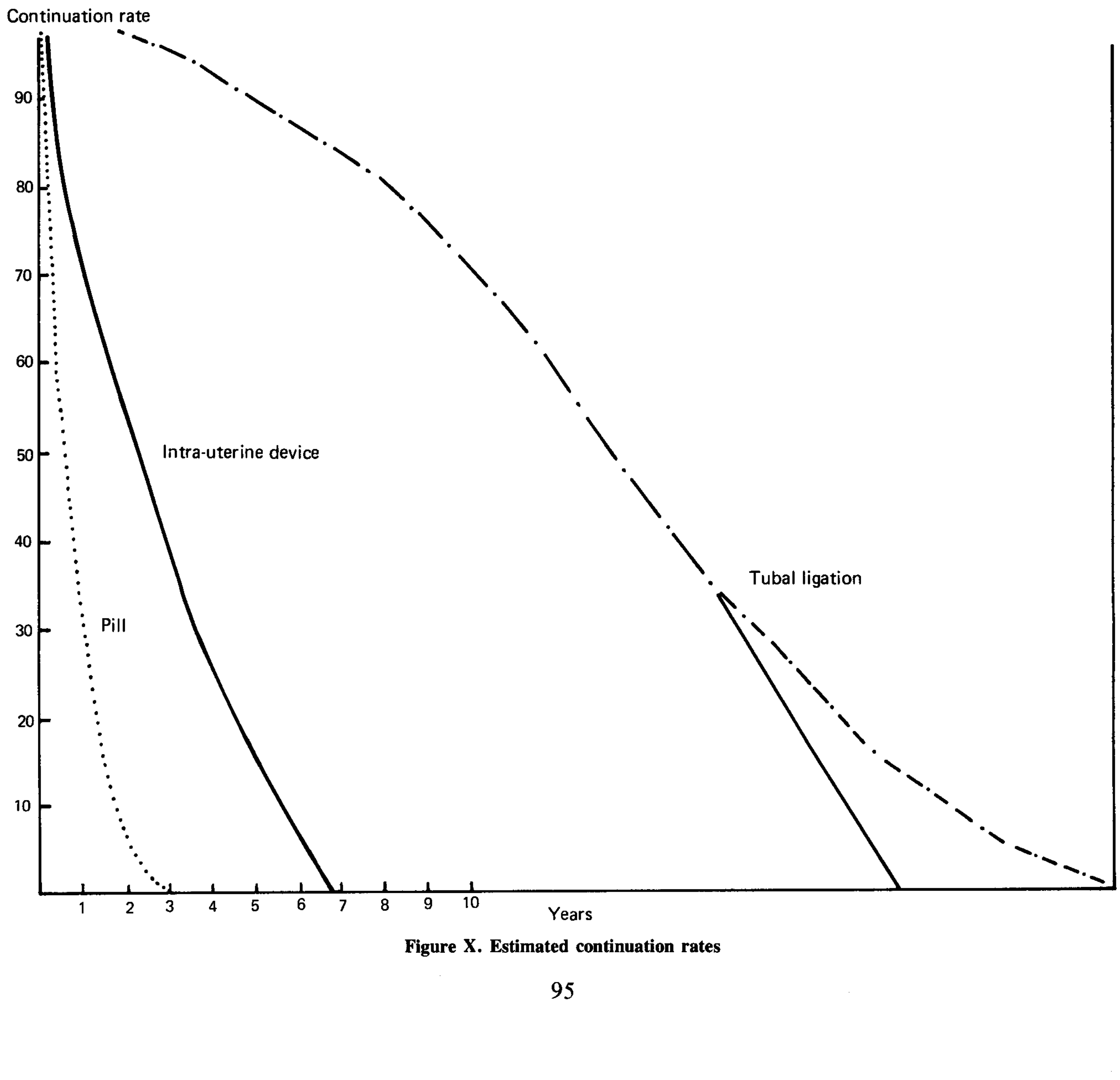


Figure X. Estimated continuation rates

Potential fertility of acceptors

One of the major assumptions on which several measurement methods are based is the estimate of the potential fertility of participants. This process involves assessing an imaginary situation by using reasonable estimates and an analysis of local demographic and social characteristics.

Proposed solutions

In the preceding section, the problem of estimating potential fertility was dealt with in a number of ways. For the method using an index of couple-years of protection, it was assumed that 1 CYP = 0.25 births; that assumption presupposes that the potential fertility of acceptors is 250 per 1,000, which is roughly equal to the legitimate fertility rate of the female population in 1971 increased by 10 per cent. This assumption would mean that contraceptive users were slightly more fertile than married women in general, and the same assumption is used in the estimates of UNPFP, which goes so far as to claim that there is no difference between the two universes if the age distribution is taken into account because acceptors are in the age brackets in which reproduction reaches its highest level.

In fact, this position seems untenable, because the survey on continuation of IUD and the pill showed clearly that fertility is higher among contraceptive users than in the general population (see table 17 and figure V, which show acceptors by number of children living at time of acceptance). Furthermore, cases of sterility are much less frequent among users; given the prevailing social conditions and outlook, the likelihood is that it is chiefly women with more than the average number of children who use contraception to prevent further pregnancies. The idea of using contraception to space births is still unfamiliar to the general population; social norms encourage couples to have children immediately after marriage and regularly thereafter, although permitting contraception after four or five offspring have been produced (even this is a great advance on earlier attitudes, which wanted everything left to nature).

For the component projection method, the potential fertility of acceptors was taken to be the fertility rate of acceptors of IUD and the pill before acceptance, as estimated on the basis of the continuation survey. This estimating process also calls for a number of criticisms:

(*a*) In the absence of any information, these rates had to be applied uniformly to acceptors of tubal ligation, condoms and jelly, whose fertility levels before acceptance were unknown. A survey of sterilization currently in progress ought to clarify this point shortly, but for condoms and jelly no reasonable indexes are available;

(*b*) It was assumed that the fertility rate "at about 30 years of age" of acceptors whose average age was 32.5 could be taken as the potential fertility rate of women aged 30–34, which implies a not insignificant degree of over-estimation of the births which would have been prevented among them;

(*c*) In the case of social abortions, a fertility rate of 290 per 1,000, i.e., the 1971 legitimate fertility rate in a population of the same structure. That estimate appears reasonable, as abortion is not exclusively a method of contraception and the motivation is different.

Although the method involving analysis of the reproductive process was not tested, the following formula was proposed to the present authors for estimating the average interval between births:

$$I = \frac{12{,}000}{\text{fertility rate by age or mean age group}}$$

This formula is an approximation which is subject to the same criticisms as before, since it uses the fertility rates of acceptors before their participation in the family planning programme.

As well as the type of data adopted by us, one could conceivably use as potential fertility rates the legitimate fertility rates for the base year 1966, with or without an increment for the higher fertility rate of contraceptive users. This approach was not adopted because it is almost certain that in 1966, fertility was particularly high (owing to, among other things, the minimum age for marriage having been raised in 1964); moreover, economic and social conditions have changed and it is inconceivable that the acceptors alone were not affected by the change.

Non-contraception programme

It is difficult to evaluate the extent to which contraceptive users might have resorted to a non-contraception programme. In Tunisia, as stated earlier, the family planning programme is part of a State-directed national population policy and all measures carried out under that policy are interdependent. The unrestricted sale and distribution of contraceptives, which dates from 1961, and the complete liberalization of social abortion in 1973 are legislative reforms which could have taken place only within the context of co-ordinated family planning. A private sector for the distribution of contraceptives does exist; but it is supervised by the Central Pharmacy of Tunisia and, although it was not included in statistics until recent years, activities within the private sector have been evaluated regularly since the establishment of the National Family Planning and Population Office. In the third quarter of 1975, almost the same number of pill cycles were distributed by the private sector as under the programme (about 32,362 for the programme and 39,976 for the private sector).

The programme authorities place considerable emphasis on educating and informing the public by direct contact and through the mass media.[27] The real effects of the programme are therefore difficult to measure and are closely linked with other aspects of economic and social development.

The ideal estimate and the most reasonable solution

It might be thought, in the absence of any kind of verification, that the ideal would be to ascertain the fertility rate of a control group having the same characteristics as the population of contraceptive users. But that is exactly where the problem arises, as the authors were unable to find a group of women which satisfied these conditions and their opinion is that, in reality, it is impossible to isolate such a group. It is therefore suggested that an intermediate assumption will serve the purpose, namely, to take as the potential fertility of acceptors the legitimate fertility rate, by method of contraception, of contraceptive users before they joined the programme, subject to adjustment for suitable age groups.

Availability of data

Certain statistical data vital for the application of these methods are non-existent or not easily accessible or would be too costly to obtain; in some cases, such data as do exist are of dubious reliability. An account follows of the difficulties encountered in applying the methods described in the preceding section.

Quality of available data

Tunisia has an adequate statistical organization, the activities of which are described below.

Censuses

The first census count was made in 1926 under the French protectorate. Similar very rudimentary counts were made every five years. They provide information mainly on the geographical distribution of the population and distinguish between different nationalities and ethnic groups. Since the obvious objectives of these surveys were of a fiscal or military nature, three age groups are distinguished (under 15 years, 15–20 years and over 21 years).

Beginning in 1936, censuses took place every 10 years and began to serve their true purpose, although in fact only the censuses of 1966 and 1975 deserve the name. Those censuses are considered satisfactory and have been corrected to allow for any possible distortions. The results are, of course, still open to criticism if a very high standard is demanded, but it is generally recognized that the quality of the data obtained is comparable to that of the most developed countries.

Vital records

Vital records were instituted in 1908, but the first very fragmentary data are available only from 1926. After independence was attained in 1956, there was a marked improvement as a result of legislation in 1958 which made the registration of births, deaths and marriages obligatory (the Personal Status Code of 13 August 1956 and the Acts of 1 August 1957 and 4 July 1958), the reorganization of regional administration and of the expansion of school attendance.

As previously stated (see foot-note 9), the estimated rate of coverage of births is 95 per cent, and the statistics are very reliable. However, it has been suggested that in 1971 the omission rate was 6.9 per cent, which would mean that INS underestimated the true situation (see discussion of trend analysis in preceding section).

The distribution of births by age of the mother is known only for 1960 and from 1965 to 1973.

Detailed statistics on births are still very scanty: data are available on births by sex, length of marriage, birth order and governorate from 1966 to 1970; and by socio-professional category of the father and type of confinement (multiple, place, type of attendance) for 1970.

Surveys

Demographic surveys directed to obtaining a better knowledge of the conditions of the Tunisian population began in 1964. These surveys are described briefly below:

(*a*) In 1964, as part of the preparations for the introduction of family planning to Tunisia, J. Morsa organized a KAP survey of 2,175 married women in 12 maternal and child welfare centres. Only one very short preliminary report was published; it served chiefly to demonstrate that the women were favourably disposed towards contraception;

(*b*) In 1967–1968, another survey of the same type was conducted by the Centre for Economic and Social Studies and Research, Tunis. The sample consisted of 1,440 people (820 women and 620 men) from Tunis and its suburbs. The analysis of the results was divided into three parts: the concept of the family size: family planning and the motivations for it; and the characteristics of the legitimate fertility rate;

(*c*) In 1968, INS undertook a national demographic survey[28] the main aim of which was to assess the quality of vital data registration. As pointed out previously, the results were very encouraging. The opera-

[27] Hedi Jemai, *Attitudes des responsables tunisiens vis-à-vis de la politique de planning familial à travers la presse tunisienne*, study for the project "Droit et population" (Tunis, Office national du planning familial et de la population, 1975); and S. Sahli, *Attitudes des responsables tunisiens vis-à-vis de la politique de planning familial à travers les émissions télévisées et les discours*, study for the project "Droit et population" (Tunis, Office national du planning familial et de la population, 1975).

[28] Tunisia, Institut national de la statistique, *Enquête nationale demographique, 1968–1969*, Demographic Series, No. 6 (Tunis, 1974).

tion also provided more detailed material for fertility measurement (in particular, measurement of the gap between urban and rural areas). The survey covered 27,000 households interviewed on three separate occasions;

(*d*) In 1972–1973, INS undertook a large-scale survey of migration and employment at Tunis.[29] Part of the questionnaire dealt with fertility: description of all the wife's offspring and birth control behaviour. The analysis covers 1,790 women and is especially concerned with distinguishing areas of origin (Tunis, urban area, semi-urban area or rural area).

(*e*) In 1973, as previously stated, ONPFP carried out a national survey on the continuation of contraceptive methods (IUD and pill) by 2,060 acceptors (between January 1969 and August 1972). The results have been issued,[30] and continuation rates for those two methods and an analysis of the fertility of contraceptive users before and after acceptance are now available. As a follow-up, a field survey is being carried out in order to provide similar information on women who have had an abortion or tubal ligation.

The results of the surveys described in (*a*)—(*d*) in the field of fertility and family planning are relatively detailed.

Statistics on family planning activities

Since 1964, all statistics compiled for use by the programme itself have been published. The general statistics (on the number of acceptors by method and by governorate) have been published in two volumes covering the period 1964–1973.[31]

Since January 1974, the statistics department of ONPFP has been publishing a quarterly statistical bulletin covering: (*a*) the activities of the centres over time; (*b*) the activities of the centres in space; (*c*) trends in the centres; (*d*) results in relation to number of medical and paramedical personnel.

For the post-partum and post-abortum programme, statistics were published from 1969 to 1971. The possibility of reviving the programme is currently being studied. The ONPFP periodically publishes reports on the demographic characteristics of acceptors, births averted (targets and results) and number of women protected (1974–1975).

Estimates regarded as valid

Some of the data used in preparing this report were not collected or drawn from surveys, but were the product of a valid estimate made by INS or a competent demographer. One example is the INS population perspectives for 1971–2001, the main assumptions for which are described above in footnote 15. It will be recalled that estimates were also used for the proportion of married women in 1971, for the residue of women x years after a tubal ligation and for births prevented by abortion. Reference to these studies will confirm the validity of the estimates.

Assumptions from imperfect data

Apart from the data just mentioned, which may not be unimpeachable but are, despite their faults, quite reliable, it was necessary to formulate for other factors assumptions whose soundness must still be assessed. However, it would be futile to expect to obtain information to verify such data since, despite its very advanced statistical organization, Tunisia is still a developing country; and the establishment of some indicators requires time, material resources and specialized personnel.

Furthermore, the application of the measurement methods demands at many points very detailed data on matters that are difficult to quantify; such data would be difficult to obtain even for a universe that was ideal from the point of view of collecting statistical data. This last observation applies to the use of contraceptives in a couple's private life, to the impact of economic, social and psychological factors on fertility, to the age distribution of women making effective use of contraception and to all the socio-economic indicators required to apply the regression method.

Listed below are only the main assumptions made by the authors in the preceding section on the application of the various methods:

(*a*) Factors affecting fertility are independent and additive;

(*b*) Illegitimate births are negligible;

(*c*) The same fertility rates are attributed to unmarried and unborn women as to women actually married in 1971;

(*d*) The natality trend is taken to be satisfactorily measured by determining the gradient of a line estimated by the least-squares method;

(*e*) The control and experimental groups have the same social, demographic and economic characteristics (see the subsection on this method for details);

(*f*) The acceptance of contraceptive methods is spread evenly over the year;

(*g*) Four abortions protect one woman for one year;

(*h*) Continuation rates can be extrapolated beyond four years;

(*i*) Users of condoms and jelly receive the quantity necessary to provide protection for one cycle;

(*j*) The effects of all the contraceptive methods are additive;

[29] Tunisia, Institut national de la statistique, *Enquête migration et emploi à Tunis, 1972–1973*.

[30] Tunisia, Office national du planning familial et de la population, *Enquête nationale sur la continuation des méthodes contraceptives en 1973*.

[31] Tunisia, Office national du planning familial et de la population, *Statistiques des activités du programme de planning familial de 1964 à 1970* (Tunis, 1971); and *idem*, *Statistiques des activités du programme de planning familial de 1971 à 1973*.

(*k*) Age distribution within age groups is uniform;

(*l*) Continuation within age groups is the same for each age;

(*m*) Eighty per cent of the condoms, jelly and pills distributed are efficiently used by couples.

Given the number of assumptions required for the application of these methods, scepticism as to the real scope and scientific value of the results is justified. Nevertheless, it was found that in many cases the errors cancel each other out and the conclusions are not inconsistent with reality.

In fact, the doubts raised by the reasoning underlying each of the methods and the evaluation process which they imply are more fundamental.

Interaction among factors

On the basis of the foregoing comments, it appears that interaction, among the factors which might affect fertility rates, is one of the basic problems to be considered in any discussion on measuring the impact of family planning programmes. It is hard to imagine family planning programmes being imposed in a social context that would reject them, and this factor is well understood by the Tunisian authorities.

Where the traditional structures of the Moslem family and its relationship with the economic environment prevail, it is certainly difficult to imagine modern contraceptive methods being accepted or even considered by the couple. (This practice would run counter to all the psycho-sociological and religious characteristics of these societies.) It has been possible for the activities of the Tunisian family planning programme to expand precisely because the demographic policy was an integral part of the over-all economic and social development model, as advocated by the authorities.

Interactions among non-programme and programme factors undoubtedly exist, as witnessed by, among other examples, the decline in family planning activities after a speech by President Bourguiba advocating population growth.[32]

The authors are inclined to think that the interdependence of all the political, economic, social and demographic factors is a fundamental principle in the understanding of national trends. However, it is realized why, in order to facilitate the measurement of social phenomena, these factors should be assumed to be independent, provided that the *a priori* bias thereby introduced is not overlooked.

Uncontrolled variables

Some measurement techniques—in particular, the standardization approach and regression analysis—try to account for all factors that may affect the level of fertility. In the present analysis, prominence is given to those factors which can be measured, namely, structure by age and by marital status, and family planning activities.

Attention has been drawn to the influence of universal schooling and of internal and external migration, although it was not possible to evaluate them in quantitative terms. This point was discussed in the preceding section (see subsection on the standardization approach) and is also dealt with in a recent study[33] prepared for the United Nations Educational, Scientific and Cultural Organization (UNESCO), which contains a detailed analysis of trends in a number of economic and social factors which may have affected fertility in Tunisia for the past 10 years: (*a*) economic growth, level of living and income distribution; (*b*) employment, migration and rural development; (*c*) education, health and well-being.

Some of these factors (emancipation of women, universal schooling, decline in mortality, migration etc.) have probably had a significant impact, but it is difficult to measure.

On the other hand, the majority of the population have not experienced decisive improvements in income, employment and housing.

However, it may well be that even though there has been only slight progress in any single one of these factors, they have together caused the change in outlook and the widespread aspiration to attain the social model exemplified by the élite, who are seen to profit from it.

Independence of methods

The main value of this report derives in large part from the juxtaposition of several methods of measuring the impact of family planning programmes on fertility. The simultaneous application of these methods to the same country and the same family planning programme in the same period makes it possible to compare their validity in relation to the same set of facts and the consistency of the reasoning on which they are based.

Two basic types of approach can be distinguished. Under the first type, the methods require an analysis of fertility trends and thereafter an estimate of the relative influence of the determining factors, which ought to demonstrate the relative impact of the family planning programme: such is the case with the standardization, projection, experimental design and regression analysis methods.

The application of the first two methods to Tunisia showed that the declining trend in fertility, which began in 1956, accelerated significantly after the beginning of family planning activities in 1964. It is still difficult, however, to identify the specific share at-

[32] Yolande Jemai, "Droit et population in Tunisie", paper prepared for the International Symposium on Law and Population; Tunis, Office national du planning familial et de la population, 1974 (mimeographed); M. Ayad and Yolande Jemai, *Tunisian Fertility Models* (Paris, UNESCO, forthcoming); and Hedi Jemai and Yolande Jemai, *La politique de la planification familiale en Tunisie*, study for the project "Droit et population" (Tunis, Office national du planning familial et de la population, 1975).

[33] M. Ayad and Y. Jemai, *op. cit.*

tributable to the family planning programme; not only have variations in the female age structure and changes in nuptiality been conducive to a decrease in the average number of children but all economic and social development factors are creating a new society, whose structures and objectives are incompatible with maintaining the traditional concept of the family and therefore imply a fairly rapid spread of family planning among couples.

Despite the imperfections of the control group in the present example the experimental method represents a novel approach to the problem; it shows clearly the fall in fertility among programme acceptors compared with the particularly high fertility rates which they displayed before using contraception and even compared with the female population as a whole. This finding confirms that even if all the decline in fertility cannot be attributed to the programme, the programme has indisputably had direct and indirect effects through information and education, and it has helped to spread new attitudes to reproduction in all sections of society.

A review of these methods indicates that, although their assumptions can be disputed and the measurement of economic and social factors is a difficult process, the standardization method achieves results which are clearer than those of other methods and which are sufficiently accurate. They should, obviously, be supplemented by a reasonable estimate of births averted by the programme—the main objective of the second series of methods tested.

These methods seek to measure the impact of family planning programmes in another way, by evaluating the result of family planning activities in demographic terms. This process involves estimating the degree of protection afforded by contraception or the number of births which have been averted by its use: that is the aim of the methods involving couple-years of protection, component projection or analysis of the reproductive process.

As applied to Tunisia, the first two methods yield almost identical results, which are also similar to the estimates made by ONPFP. According to the assumptions used for the potential fertility of acceptors, some 13,000–15,000 births are believed to have been averted in 1971. Although differing in detail, the estimates lead to the same conclusions: the most effective methods are IUD and tubal ligation. Abortion should be treated separately, since it may or may not be considered a method of contraception.

Cost-precision analysis

It does appear that refining the analysis adds nothing of moment to an assessment of the results. All the conclusions reached in the present study are consistent with one another, and it seems unnecessary to spend a great deal of time and effort to show in greater detail the magnitude of the impact of a family planning programme. It would be enough to improve the standardization method and to estimate the number of births averted by family planning by a homogeneous procedure in order to permit comparisons between countries.

On the other hand, it would certainly be useful to improve techniques for measuring the impact of family planning programmes on fertility from the scientific point of view, in order to make possible an in-depth assessment of the consequences of a State-directed population policy at the national level.

In most countries, the authorities are more concerned with action than with abstract thinking; they should be given the tools for measuring the result of family planning activities by methods that are simple and easy to apply systematically. On the other hand, the specialists must seek methods of monitoring population policies in order to evaluate their long-term consequences and to warn against any possible dangers to society.

Conclusion

The application of the methods discussed above to Tunisia for the period 1966–1971 has served to confirm what was already known from other studies. It would have been much more instructive to have done this study on a more recent period (1971–1975), but lack of statistical data made that impossible.

Family planning programmes are still very new and organizational changes occur frequently. Until 1973, structural changes were being made—a good indication that the responsible authorities were seeking an ideal operational model. Since 1973, when the National Family Planning and Population Office was established, family planning activities have been given a definite, permanent place in national economic and social development planning.

Family planning is now seen as a fundamental option of national policy, knowledge well known to and accepted by the population.

Currently, however, despite the unquestionable success experienced in recent years, the fact that since the liberalization of abortion in 1973, abortions have been increasingly resorted to, to the detriment of preventive contraception, gives cause for concern.

The authorities want to see couples make greater use of birth-spacing contraception rather than absolute contraception, a development which would provide tangible evidence of the success of their policies.

The function of evaluation experts is to help the authorities to assess the measures taken and, more importantly, to try to become more familiar with the workings of a policy which has a direct impact on personal, family and social well-being.

Annex I

TUNISIA: DEMOGRAPHIC AND FAMILY PLANNING INDICATORS

TABLE 50. DEMOGRAPHIC INDICATORS, 1966 AND 1971-1975

Indicator	*1966*	*1971*	*1972*	*1973*	*1974*	*1975* [a]
Population						
Population at mid-year [b]	4 717 500	5 228 400	5 331 800	5 444 200	5 616 300	5 572 229
Male	2 323 700	2 557 200	2 612 600	2 652 600	2 758 900	2 808 622
Female	2 393 800	2 671 200	2 719 200	2 791 600	2 857 400	2 763 607
Women of reproductive age (15-54) [c]	1 071 300	1 229 300	1 268 200	1 314 200	1 350 000	...
Population of communes (*percentage*) [d]	40.0	...	...	...	...	49.1
Urban population (*percentage*) [d]	...	50.4	...	...	...	...
Rural population (*percentage*) [d]	...	49.6	...	...	...	...
Structure						
Age distribution (*percentage*) [e]						
0-14 years	46.5	45.5	44.8	44.3	43.7	43.2
15-64 years	49.9	50.5	51.1	51.6	52.1	52.6
65 years and over	3.6	4.0	4.1	4.1	4.2	4.2
Vital statistics [b]						
Births registered [f]	206 730	183 311	198 785	194 764	191 049 [g]	...
Deaths registered [f]	48 307	48 625	40 053	43 716	39 062 [g]	...
Marriages registered	27 037	37 750	45 043	43 183	...	...
Divorces registered	4 616	4 584	4 930	5 099	...	...
Corrected birth rate (*per 1,000 population*)	43.8	36.9	39.3	37.7	35.8	...
Corrected death rate (*per 1,000 population*)	14.0	12.7	10.3	11.0	9.5	...
Corrected general fertility rate	193.0	157.0	165.0	156.0	149.0	...
Marriage rate (*per 1,000 population*)	5.7	7.2	8.4	7.8	...	...
Divorces (*per 1,000 population*)	0.98	0.88	0.93	0.93	...	...
Increase						
Balance of migration [h]	−12 637	−32 281	−24 552	−12 768	+2 352	...
(Estimated) births minus deaths [b]	140 556	126 350	154 830	145 128	137 540	...
Crude rate of natural increase (*percentage*) [b]	29.8	24.2	29.0	26.7	26.3	...
Net rate of increase (*percentage*) [i]	27.1	18.0	24.4	24.4	26.7	...

[a] Numbers taken from provisional returns from census of 8 May 1975.

[b] Tunisia, Institut national de la statistique, *Statistiques de l'INS*, Serie démographie No. 5 (Tunis, December 1975).

[c] Tunisia, Institut national de la statistique, *Niveau et tendances de la fécondité en Tunisie*, Serie démographie No. 5 (Tunis, May 1974).

[d] Tunisia, [Ministère du Plan], *Aménagement du territorie: l'armature urbaine en Tunisie, 1973.*

[e] Tunisia, Institut national de la statistique, *Perspectives d'évolution de la population, 1971-2000*, fasc. II (Tunis, 1972).

[f] The numbers of births and deaths registered have been increased by 5 per cent and 27 per cent, respectively. The figure for births in 1966 has not been corrected; INS believes the rate to be 100 per cent for that year (figure becomes 46.1 per cent for births if increased by 5 per cent).

[g] Unpublished data provided by Institut national de la statistique.

[h] Tunisia, Ministère de l'Intérieur, cited in Institut national de la statistique, *Economie de la Tunisie en chiffres de 1967 et 1971* (for 1966 and 1971); and in *idem, Bulletin mensuel de la statistique* (November-December 1974 and January 1975) (for 1972-1975).

[i] In estimating net rates of increase, the balance of migration was included.

TABLE 51. FAMILY PLANNING INDICATORS, 1966 AND 1971-1975

Indicator	*1966*	*1971*	*1972*	*1973*	*1974*	*1975*
Contraceptive users						
Total number of female contraceptive users...	41 517	239 916	246 675	241 335	256 984	51 714[a]
New female contraceptive users	16 176	40 360	43 665	43 840	50 901	30 927[a]
New acceptors of:						
Intra-uterine device	12 077	12 381	13 250	16 790	19 084	9 917[a]
Pill	350	11 778	12 026	11 194	10 795	7 709[a]
Tubal ligation	766	2 280	2 453	4 964	10 757	6 503[a]
Social abortions	1 326	3 197	4 621	6 547	12 427	7 833[a]
Women protected by family planning programme at end of year	...	49 900	54 868	64 109	77 959	...
Rate of protection per 100 married women, aged 15-49, at end of year	...	...	...	8.47	10.06	...
Rate of protection per 100 married women, aged 15-54, at end of year	...	...	...	7.75	9.21	...
Target and achievement						
Births to be averted	...	12 000	15 500	19 000	22 500	26 250
Births averted by family planning programme..	...	13 330	15 515	17 288	23 117	29 720
Target achievement (*percentage*)	...	111	100	91	103	113
Infrastructure						
Operational family planning centres	...	360	330	309	392	...
Married women of reproductive age (15-49) per centre	...	...	...	2 432	1 976	...
Medical specialists	...	35	...	...	69	83[a]
Midwives	...	40	...	...	92	81[a]
Medical aides	...	60	...	...	265	447[a]
Female contraceptive users per doctor	...	6 855	...	...	3 724	...
Female contraceptive users per midwife	...	5 998	...	...	2 793	...
Female contraceptive users per medical aide ...	...	3 999	...	...	970	...

[a] Data for first half-year.

Annex II

BIRTHS AVERTED BY TUBAL LIGATIONS IN TUNISIA*

Tubal ligations have been performed in Tunisia since the beginning of the experimental programme in 1964, and the number performed has risen steadily (see table 52). Since the demographic effects of this radical method of birth control are felt over a long period, a way of accurately quantifying the number of births averted by tubal ligations is required.

TABLE 52. TUBAL LIGATIONS PERFORMED FROM 1964 TO MID-1974

Year	*Number of ligations*	*Year*	*Number of ligations*
1964	293	1970	2 539
1965	384	1971	2 280
1966	766	1972	2 459
1967	742	1973	4 964
1968	1 627	1974 (first quarter) ..	3 480
1969	2 513	1974 (second quarter).	3 096

GENERAL PRINCIPLES OF THE ESTIMATING PROCESS

Ostensibly, a tubal ligation performed on a woman aged 25 years would cause all possible births to her between that age and menopause at, say, 50 years of age, to be averted. This reasoning implicitly assumes that this 25-year-old woman, if she had not undergone ligation, would have been exposed to all the risks of fertility during the remainder of her reproductive span. However, even if she had not undergone tubal ligation, she might have: (*a*) died between the ages of 25 and 50; (*b*) been widowed between the ages of 25 and 50; (*c*) had a divorce; (*d*) become sterile; or (*e*) resorted to another method of contraception.

In any event, the average reproductive span of the couple would have lasted less than 25 years. These factors must therefore be taken into account in producing a breakdown of births actually averted by tubal ligation. All of the following estimating processes therefore rely upon a comparison between the absence of births following a tubal ligation and what might have happened in the absence of tubal ligation. It is assumed, with no great risk of error, that fertility is nil after age 50. Illegitimate births are ignored.[a]

ESTIMATING BIRTHS AVERTED BY TUBAL LIGATION

Assuming that the probabilities of death, widowhood, divorce or becoming sterile from the time of ligation to menopause (at 50 years) are independent of one another, changes in the number of fecund couples surviving and cohabiting after tubal ligation are first estimated.[b] When all these factors are taken into account, one will obtain, for each age, the number of surviving, fecund, cohabiting couples there would have been in the absence of tubal ligation per 1,000 women having undergone tubal ligation at a given age (see table 53). It is to this residual number at each age that the age-

*Prepared by L. Behar for the Tunisian case study.

[a]Births classified as illegitimate constituted 0.37 per cent of all births in 1968, according to Institut national de la statistique, *Naissances 1966 à 1968, statistiques détailleés,* Demographic Series, No. 4 (Tunis, February 1974), pp. 57–59.

[b]For mortality, use was made of the life tables for both sexes prepared by Institut national de la statistique, on the basis of the National Demographic Survey, 1968–1969. It was assumed that, on average, husbands are five years older than their wives. The distribution of divorces by length of marriage was estimated from divorces in the 1964 marriage cohort and the distribution by length of marriage of divorces in 1969. See "Naissances, décès, mort-nés, mariages, divorces, 1969", Demographic Series, No. 3 (February 1973), pp. 229 and 243. For sterility, the age-specific data are those estimated by L. Henry, "Fécondité des mariages, nouvelle méthode de mesure", *Cahiers de l'I.N.E.D.*, No. 16 (1953). An attempt was made to allow for the possibility that a woman who has not undergone a ligation may resort to traditional means of contraception by relying on the age distribution of new acceptors of intra-uterine devices and the age-specific rates of exposure to family planning.

TABLE 53. RESIDUE OF FECUND COHABITING COUPLES BY AGE AT TIME OF LIGATION

Age at time of ligation	Age of couples					
	25	30	35	40	45	50
25	1 000	829	703	532	285	0
30		1 000	845	639	342	0
35			1 000	744	398	0
40				1 000	535	0
45					1 000	0
50						0

specific legitimate fertility rates will be applied in order to estimate the number of births averted by tubal ligations.

On the basis of table 53, the average period[c] of protection by tubal ligation (see table 54) is estimated. It will be found that the average periods of protection are much lower than the difference between age at menopause (50 years) and age at time of ligation.

However, in order to eliminate as far as possible distortion arising from incorrect statements of age, five-year age groups should be taken as the departure point.

TABLE 54. AVERAGE PERIOD OF PROTECTION BY AGE AT TIME OF TUBAL LIGATION

(*Years*)

	Age at time of ligation			
	25	30	35	40
Average period of protection	14.3	11.6	8.2	5.2

For each five-year age group between ages 20 and 50, 1,000 women who had undergone ligation were taken and that initial arbitrary number was subjected to the risks of death, widowhood, divorce, sterility and resort to another method of contraception up to the age of 50 years, year by year (see table 58 in appendix). Each age group at time of ligation is represented by its mid-point, except the age group 20–24, in which the average age at time of ligation is 23, and the 45–49 age group, in which the average age is 47. These average ages were estimated on the basis of the ages of women who underwent ligation during the first half of 1974. The residual numbers of women are reduced to zero after a number of years equal to the difference between the average age and the upper age limit for fertility, i.e., 50 years.

In order to determine for all ages combined the residual number of women 1, 2, . . . *n* years after ligation (table 55), each residual number was weighted by a quantity representing the proportion of women in that age group at the time of ligation to the total number of women having undergone ligation. The weighting coefficients are those relating to the age distribution of women having undergone ligation during the second quarter of 1974.[d]

It is clear, therefore, that 35 per cent of the women are outside the field of observation at the end of 5 years and over 68 per cent at the end of 10 years. After 15 years, the residue is 10 per cent of the women and after 20 years only 2 per cent. It must be borne in mind that these figures are for surviving, cohabiting, fecund couples in which the age distribution of the wives is the same as that of women having undergone ligation, but in the absence of ligation.

On the basis of table 58, and proceeding in the same way as in estimating life expectancy from a life table, the average period of protection was estimated (table 56).

It is to the residue at each age, as shown in table 1, that the legitimate fertility rates are applied in order to derive the number of births averted by tubal ligation. Use was made of the legitimate fertility rates taken from the Tunisian national population survey.[e] As the women grow older, the legitimate fertility rates corresponding to the age groups into which they move are applied (see table 59). In the summation (table 57), for each age group at time of ligation the number of births averted per 1,000 women having undergone ligations in the age group in question is determined.

Thus, for each woman who has undergone ligation between the ages of 20 and 25 (the average age being 23 years), 4.06 births are averted; for each woman who has undergone ligation between the ages of 25 and 30, 2.8 births, on average, are averted etc.

The weighted average of births averted by age at time of ligation, using the same weighting coefficients as before, will yield the actual number of births averted by tubal ligation. This number is 1,095 births averted per 1,000 women who have undergone ligation. To simplify the computation, the figure may be rounded off to 1.1 births averted by ligation. The distribution over time of these 1,095 births averted shows a heavy concentration in the years immediately following ligation. In fact, 48 per cent of the averted births are averted in the 3 years following ligation, 66 per cent within 5 years of ligation and 92 per cent in the first 10 years after ligation.

[c] In this connexion, it was assumed that the "departure" for the universe used here was a linear process between age x and age $x + 5$. That hypothesis is probably invalid after age 40, but the relatively small numbers remaining after that age greatly reduces the risk of error in the results.

[d] The age distribution of women who have undergone tubal ligation is available for only a few years, the most reliable data—and the only data covering the whole of Tunisia—being those for the second quarter of 1974. Moreover, the age distribution of women who have undergone tubal ligations and the average age do not appear to have changed significantly, except perhaps for greater concentration between the ages of 30 and 40, a corollary of the slightly smaller relative weight of the extreme age groups, 20–24 and 45–49 (see table 60 in appendix).

[e] See Tunisia, Institut national de la statistique, *Niveau et tendances de la fécondité en Tunisie,* Demographie Series, No. 5 (Tunis, May 1974), p. 16.

TABLE 55. RESIDUE OF WOMEN *n* YEARS AFTER LIGATION PER 1 000 WOMEN HAVING UNDERGONE LIGATION, ALL AGES COMBINED

n	Residue of women	*n*	Residue of women	*n*	Residue of women	*n*	Residue of women
1	933	8	429	15	101	22	7
2	867	9	372	16	75	23	4
3	793	10	313	17	50	24	3
4	718	11	253	18	34	25	2
5	645	12	196	19	27	26	1
6	572	13	150	20	20	27	0
7	498	14	126	21	14	28	0

TABLE 56. AVERAGE PERIOD OF PROTECTION BY AGE GROUP AT TIME OF LIGATION

(Years)

Age group at time of ligation	*Average period of protection*
20-24	15.88
25-29	12.98
30-34	9.99
35-39	6.79
40-44	3.87
45-49	1.50
All ages	7.70

TABLE 57. BIRTHS AVERTED, BY AGE AT TIME OF LIGATION

Age group at time of ligation	*Births averted per 1,000 women having undergone ligation in the age group*
20-24	4 057
25-29	2 810
30-34	1 630
35-39	716
40-44	191
45-49	3
All ages	1 095

A PRACTICAL EXAMPLE: WOMEN WHO UNDERWENT A LIGATION IN 1973

Since the age distribution of women who underwent a tubal ligation operation in 1973 was not available, it was assumed that the distribution was the same as that for the second quarter of 1974. The residue of women and births averted each year from 1973 to 2000, when this group of women will be completely outside the field of observation, are given in table 61. All told, 5,440 births will be averted by the 4,964 ligations performed in 1973. Most of these births will be averted in the years immediately following the ligation and 94 per cent before 1985.

CONCLUSION

Each tubal ligation operation performed in Tunisia averts, on the average, 1.1 births. The births averted among a group of women who undergo the operation during a given year are spread over 30 years, but over 90 per cent of them are averted in the 10 years following the operation. It is very unlikely that there will be any great change in the age distribution and average age of the women undergoing the operation which would significantly alter the data.

Births averted by tubal ligations have been underestimated owing to the fact that the life table for the years 1968–1969 was used. The death rate in Tunisia is falling steadily, a fact which increases the proportion of surviving couples allowed for in these estimates. Nevertheless, the fact that use was made of fertility data for the years 1968–1969 may offset this underestimation in view of the decline in fertility in Tunisia.

Appendix

Table 58. Residue of women *n* years after ligation

Average age at time of ligation	Number of years 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28
23	977	954	924	894	864	835	865	781	758	734	740	687	655	623	592	560	528	480	432	384	336	288	230	173	115	57	0	
27.5	963	926	892	865	839	811	784	751	714	677	640	603	556	502	448	394	339	280	218	156	94	32	0					
32.5	966	931	893	848	804	758	715	655	595	530	466	408	328	260	185	111	37	0										
37.5	944	889	821	742	661	582	501	414	322	231	137	47	0															
42.5	880	758	628	486	349	210	68	0																				
47	667	333	0																									
All ages combined	933	867	793	718	645	572	498	429	372	313	253	196	150	126	101	75	50	34	27	20	14	7	4	3	2	1	0	

Table 59. Births averted each year
(Per 1 000 women)

Average age at time of ligation	Number of years 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	Births averted per 1,000 women who have undergone ligation in the age group
23	379	370	328	317	307	296	286	217	211	204	198	190	130	123	117	111	105	37	34	30	27	23	5	4	3	2	1	0	4 057
27.5	342	329	284	241	231	226	218	179	141	134	127	120	78	39	35	31	27	15	5	4	3	1	0						2 810
32.5	269	260	200	168	159	151	142	92	46	42	37	32	17	6	5	3	1	0											1 630
37.5	188	177	115	58	52	46	40	21	7	6	4	2	0																716
42.5	69	61	33	11	8	6	3	0																					191
47	2	1	0																										3
All ages combined	201	191	139	199	92	86	80	54	32	29	27	24	14	7	6	5	4	2	1	1	1	04							1 095

TABLE 60. AGE DISTRIBUTION AND AVERAGE AGE OF WOMEN WHO HAVE UNDERGONE TUBAL LIGATION

	Age distribution of women who have undergone tubal ligation (Per 1 000)						
	Under 25	*25-29*	*30-34*	*35-39*	*40-44*	*45+*	*Average age*
1969 (seven hospitals)	23	153	280	338	151	45	34.9
1970 (Bizerta and M. Bourguiba)	30	152	255	336	186	41	34.6
1974 (first quarter)	10	99	307	357	204	23	35.2
1974 (second quarter)	15	94	269	413	184	25	35.17

TABLE 61. WOMEN WHO UNDERWENT TUBAL LIGATION IN 1973; RESIDUE OF WOMEN REMAINING AND BIRTHS AVERTED, 1974-2001

	1974	*1975*	*1976*	*1977*	*1978*	*1979*	*1980*	*1981*	*1982*	*1983*	*1984*	*1985*	*1986*	*1987*	*1988*	*1989*	*1990*	*1991*	*1992*	*1993*	*1994*	*1995*	*1996*	*1997*	*1998*	*1999*	*2000*	*2001*
Women remaining at mid-year ..	4 631	4 304	3 936	3 564	3 202	2 839	2 472	2 130	1 847	1 554	1 256	973	747	625	501	372	248	169	134	99	69	34	20	15	10	5	0	0
Births averted ..	998	948	690	492	457	427	317	268	159	144	134	119	70	35	30	25	20	10	5	5	5	2						

APPLICATION OF METHODS OF MEASURING THE IMPACT OF FAMILY PLANNING PROGRAMMES ON FERTILITY: THE CASE OF CHILE*

*Erica Taucher*** and Albino Bocaz****

INTRODUCTION

Chile, which is situated on the south-western coast of South America, had a population of 8.9 million at the time of the 1970 census. Of that number, 219 million lived in the capital. Of the total population, 75.1 per cent live in urban areas.

At the same date, the female population of childbearing age, from 15 to 49 years, accounted for 24.1 per cent of the population. The illiteracy rate was 10.6 per cent, and 8.9 per cent for women of childbearing age.

The crude birth rate, which was stable at 35 per 1,000 or more for several decades, has been steadily declining since 1964; by 1974 it had reached a rate of almost 25 per 1,000. Over the same period, the infant mortality rate dropped to approximately 40 per cent and the general mortality rate fell from 11 to 7.4 per 1,000.

The annual growth rate, which was 2.51 per cent between 1959 and 1964, dropped to 1.95 per cent during the period 1969–1974.

Family planning activities were officially begun in 1964. They do not constitute a programme as such, but form part of the mother and child care programme. Most of the activities are carried out under the National Health Service, which covers approximately 70 per cent of the country's population.

The objectives of family planning activities are defined in terms of health indexes and are directed mainly to lowering the maternal mortality rate, and particularly mortality due to abortion. It is therefore not the purpose of these activities to lower the fertility rate, although that is one effect of making available to women the means of spacing births and limiting families.

This framework explains why there is no separate statistical system for family planning and why the information on it is gathered as if it were just one other activity within the health statistics system.

Data sources

In order to apply the various methods of evaluating the impact of family planning activities on fertility, data were needed which can be classified as follows: measurements of fertility; non-programme factors; and programme factors.

Measurements of fertility

To obtain the data concerning measurements of fertility, two sources were used: vital statistics and censuses.

Vital statistics in Chile are compiled and published by two bodies, the National Institute of Statistics and the National Health Service. Although the former organization is the official agency for the entire country, its most recent publication is dated 1970; the National Health Service, on the other hand, has issued publications up to and including 1974. For this reason, both sources of information had to be used. Furthermore, both institutions have different methods of tabulating births which made it necessary to use one or the other, depending upon the needs of the method. It should be explained that births are published taking into account the date of occurrence.

For one of the methods, information on the total number of live births as recorded in the census was used.

Non-programme factors

From the vital statistics, data were obtained on nuptiality by age and on the age of mothers of live-born children.

As an indicator of the existence of medical resources, the percentage of confinements without professional care, published in the annual birth record of the National Health Service, was used for the regression method.

The remainder of the non-programme factors were obtained from census data, for which use was made of the publications of the 1960 and 1970 censuses, and the samples of these censuses in the data bank of Centro Latinoamericano de Demografia (CELADE).

Programme factors

As previously mentioned, in Chile there is no separate statistical system to provide data on family planning activities. Various sources therefore had to be used to obtain data.

---

* The original version of this paper appeared as document ESA/P/AC.7/4.

**Chief, Health and Population Sector, Centro Latinoamericano de Demografia, Santiago de Chile.

***Centro Latinoamericano de Demografia, Santiago de Chile.

Information on new participants in the programme during the period 1964–1966 was obtained from a publication by Requena and Monreal;[1] in 1967–1973, from a CELADE publication; and in 1974, from the Chilean Family Protection Association (APROFA). This association also made available data covering the family planning programme by province.

To break down the new participants by age group, use was made of data from publications by CELADE concerning the application of the Service Statistics System (SIDES) in the National Employees' Medical Service (SERMENA). Continuation rates for the use of contraceptive methods were obtained from the same source.

Methods applied

Of the eight evaluation methods suggested, for the study of cases, only five could be applied. With respect to the experimental design method, there was no study especially designed for the purpose and the accessible data were not suited for this type of analysis. Nor was it possible to obtain data on intervals between births, required for analysis of the reproductive process. Lastly, there was neither enough information nor facilities for producing a simulation model.

The present study therefore covers the following methods: standardization approach; trend analysis; couple-years of protection (CYP); component projection approach and regression analysis.

Discussion

The methods whose application is described below are discussed in the light of the problems encountered during application and the results obtained.

Application of methods

Availability and quality of data

It was only in the standardization approach that problems were encountered with the measurement of the fertility selected, which was the number of live births per woman as reported in the census. In 1960, there was a major omission in the datum; and in 1970, it was discovered that a coding error had produced an inflated figure. However, these errors appear to have had a uniform effect on different groups, as the ratios between groups for each census were reasonable. This led the authors to correct the information, which made it possible to determine the significance of the changes given by different variables. No estimate was made of the amount of the differences between the two years, which would only have reflected the correction procedure used.

In the other methods, vital statistical publications provided the necessary data on births. Omissions in the register of births were studied both for the country as a whole and for individual regions, and the data could be satisfactorily corrected.

Potential fertility for the couple-years of protection and the component projection methods was estimated on the basis of marital fertility rates. That procedure entailed correcting the census data on married women by age group, making a nuptiality cohort analysis, with the data obtained from vital statistics.

There was no difficulty in obtaining data on legitimate children since they are published in the vital statistics.

With regard to programme factors, they were not included directly either in the trend analysis or in the standardization approach.

In the regression analysis, use was made of the datum of coverage of women of childbearing age by province.

Therefore, the two most demanding methods as concerned programme data were the couple-years of protection and the component projection approach. It was necessary to assume that the age distribution of new participants and the rates of continuation in the use of contraceptive methods in the SERMENA experiment could be applied to new participants throughout the country. The validity of this assumption was tested by comparing this information with that contained in a number of Chilean studies; and it appeared to be reasonable, in the absence of more comprehensive new and up-to-date information.

The data on non-programme factors, necessary for the regression analysis, did not present any major problems. With regard to the standardization approach, it is unfortunate that the wealth of information contained in the statistical report on live births (age, level of education, occupation, occupational category and place of residence of each parent and the number of live-born children still alive, live-born children now dead or stillborn children of the mother), has not been published in full. It was therefore necessary to resort to data from census samples, with the drawbacks already mentioned.

Problems of application

Once the problem of obtaining data was solved, the couple-years of protection method and the component projection approach presented no obstacles, since they are quite clear-cut.

An important application problem appeared in the regression analysis, concerning the choice of the analytical method. Several computer programs were available, but all of them are based on distribution assumptions which might not be met by the available data. On the other hand, the small number of observation units (25 provinces) did not allow multivariate analysis with distribution free methods.

[1] M. Requena and T. Monreal, "Evaluación del programa de control del aborto inducido y planificación familiar en Chile", *Milbank Memorial Fund Quarterly*, vol. XLVI, No. 3 (July 1968), part 2, pp. 213–246.

The fertility projections also posed a problem of selection with regard both to the functions to be used and to the base data for the projections.

In the present case, the standardization approach required the use of a computer because the data were recorded on magnetic tape, and it was further complicated by the need to make a special computer program.

Results of application

Validity

With the background information already provided on the quality of the data and on the assumptions adopted, it is difficult to be categorical about the validity of the methods. An effort is made to analyse this aspect of the question by comparing the results obtained by independent methods designed to achieve the same estimates.

Before proceeding further with this analysis, mention should be made of two problems that affect the interpretation of the results obtained: abortions and sterilizations.

It should be noted that family planning activities in Chile were mainly designed to reduce abortion, which was a serious problem as concerned the health of mothers and hospitalization costs (see annexed Table 40). According to data of the National Health Service, of the total number of obstetrical hospitalizations, the percentage for abortion which had amounted to 21 per cent in 1964, had declined to about 16 per cent by 1973. Mortality due to abortions dropped between 1964 and 1974, from 10 per 10,000 live births to 4 per 10,000 live births. This decline could be an indication that the use of contraceptives is partially replacing abortion as a means of avoiding childbirth.

As for sterilization, although there are no specific figures, the growing proportion of births by Caesarian section implies sterilization after the third or fourth child delivered by this method. Furthermore, in a number of clinics—just how many it is hard to determine—female sterilization is carried out in accordance with certain socio-economic and health criteria, although this activity is not quantifiable.

If one analyses the results obtained by means of both methods used in an attempt to identify the factors having the greatest impact on fertility, one finds that both in the standardization approach and in regression analysis, the level of education proves to be an important variable.

The same is not true of economic activity, which becomes important in regression analysis but not in the standardization approach.

The urban or rural area of residence has been found significant in the standardization approach but is negligible in regression analysis.

Lastly, the effect of the coverage of family planning activities in 1970 could be studied only in the regression analysis.

The fertility figures estimated using the variables of illiteracy, female economic activity and family planning coverage are very close to the figures observed. However, the positive sign of the regression coefficient of the coverage variable makes this method useless, in this case, as a means of measuring the effects of the programme on fertility.

It is possible that the variables selected and the types of regression used may be responsible for this situation. It would appear that the indirect effects of illiteracy and economic activity on fertility cancel and exceed the negative correlation between coverage and fertility ($r = -0.5156$), so that the direct effect of coverage on fertility becomes positive. There was not enough time to undertake a more exhaustive analysis of these problems.

The three methods directed towards estimating live births averted gave more satisfactory results which may be intuitively judged sound.

If the fertility projection is taken as a reference point and it is borne in mind that the decline in the number of live births is due not only to activities in connexion with the family planning programme it is expected that the methods for quantifying the number of live births averted (couple-years of protection and component projection approach) will give lower results. In fact, one finds the following situation:

Difference between live births averted according to the projection and live births observed

Period	*Difference*
1965–1970	283 709
1965–1974	639 623

Live births averted

Period	*Intra-uterine device*	*Oral gestagens*[a]	*Total*
	Couple-years of protection		
1965–1970	138 379	34 720[a]	173 099
1965–1974	328 704	84 832[a]	413 536
	Component projection approach		
1965–1970	140 577	35 271[b]	175 848
1965–1974	391 135	100 944[b]	492 079

[a]Estimate made on new participants using oral gestagens, with distribution and continuation rates by the National Employees' Medical Service (SERMENA), marital fertility adjusted by age and on the basis of a two-year period of participation in the programme.

[b]Estimate, applying proportions of live births averted through oral gestagens out of the total in the couple-years of protection method.

It will be noted that the two methods for estimating live births averted give lower figures than the projection figures.

While comparing the number of births averted resulting from the application of the CYP and component projection methods, a relatively greater difference in results for the period 1965–1970 is found, always with higher values for the component projection method. An analysis of the assumptions underlying each method explains this differential behaviour.

In applying the component projection method, the potential fertility of contraceptive users was obtained by increasing the marital fertility rates by 20 per cent above the level adopted in applying CYP, on the assumptions that acceptors have higher than average fertility.

The other factor responsible for an increase in the component projection estimate is the assumption regarding continuation rates. In this method, acceptors are corrected only once by a hypothetical continuation rate modified for death and widowhood. In the present case, this correction coefficient was 0.714085. On the other hand, CYP assumes a geometrical decrease of continuation rates. Comparing figures for six years, the following ratios show the difference in results obtained with the component projection and CYP methods:

	Continuation rates		
Year of the programme	*Couple-years of protection*	*Component projection*	*Couple-years of protection component projection*
1.	0.8223	0.714085	0.8684
2.	0.7307	0.714085	0.9773
3.	0.5890	0.714085	1.2124
4.	0.4750	0.714085	1.5033
5.	0.3832	0.714085	1.8635
6.	0.3090	0.714085	2.3109

Another differential factor is that in the component projection method, it is assumed that in a specific year t only one half of the acceptors of year (t-1) avoid births. Although it is correct that, on average, these acceptors remain only six months of the year (t-1), all of them reach one year of permanence at some moment of year t. Therefore, the $Qi_{(t)}$ underestimate the population of year t.

For these reasons, the underestimation features of the component projection method for the first year have more weight over short periods of analysis, whereas for long periods the factors tending to lead to an over-estimate have a predominant influence on the results.

Assuming that CYP would be preferred for calculating births averted by the programme, there still remains the problem that they might over-estimate its impact, since an unknown number of women might have resorted to abortion if they had not had access to the programme. Therefore, this measurement is somewhere in between births and pregnancies averted.

Interpretation

There are no interpretation problems in the couple-years of protection and component projection methods once the assumptions on which they are based are accepted. Interpretation of the results of projections requires study of the factors affecting fertility and is open to subjective influences.

Accepting that the trend projection contains some effects of non-programme variables, it was assumed that in the present case, this estimate would show higher figures than the births averted, calculated directly with CYP and component projection. This assumption was based on the fact that neither sterilization nor use of contraceptives outside the programme are considered in the latter methods. Although the over-all results were in accordance with this assumption, the interpretation still presents problems. Programme effects are probably not additive and they might interact with some of the non-programme variables affecting trends, especially with abortion practice for which the programme might be partially substitutive.

Therefore, trend projections might be considered to be a useful descriptive technique but not a method for measuring impacts of the programme.

Regression analysis should, in theory, lead to objective interpretations. None the less, in the authors' experience, it was necessary to find reasons to explain the results that did not turn out as expected.

In the path analysis, the subjective assumptions of the research worker acquire even greater importance.

Lastly, the interpretation of the results of the standardization approach is most complex, since it is difficult to separate direct effects from interactions.

Standardization approach

In order to standardize according to different variables, it would have been necessary to have cross tabulations of births on several criteria. Such tables could not be found in the vital statistics publications, nor could access be had to original birth data. Therefore, the 1960 and 1970 Chilean census samples in the CELADE data bank were used for the application of this method.

To measure fertility, the average number of children born alive to married women of 15 years of age and over was taken, classified by:

(a) *Age group:* 15–19, 20–24, 25–29, 30–34, 35–39, 40–44, 45–49, 50 and over;

(b) *Area of residence:* city, capital, urban, rural;

(c) *Economic activity:* economically active, not economically active;

(d) *Level of education:* no education, 1–3 years, 4–6 years, 7–9 years, 10 years or more.

Procedure

The 1970 population was taken as the standard population. The average number of children born alive per woman for each age group was standardized separately by: level of education; economic activity; and area of residence. In addition, it was standardized simultaneously by level of education and economic activity, and by level of education, economic activity and area of residence. The average number of children born alive for the total number of women in 1960 was standardized by age and, simultaneously, by age and the same variables and associations of variables described for the standardizations in each age group.

Application of the method

When the average number of children born alive in the master tables had been obtained, it was found that for 1960 there was an omission in the registration of children born alive. In 1970, owing to a coding error, the figures were higher than the probable real amounts. However, the relative position of the averages between the various categories was consistent, as were the directions in which the changes occurred in the course of standardization.

Since the object of this method is to determine the influence of the different variables on the change in the average number of children born alive rather than the actual change in this index, the averages obtained were corrected. The correction procedure consisted of comparing the average number of children born alive by age group of mothers from the census sample with an estimate of children born alive, based on age-specific fertility rates obtained from vital statistics.

Since the census sample data referred to married women, assuming legitimacy of their children, whereas the fertility rates were based on total births per total women in each age group, the first step had to be the conversion of the averages from the census sample into total averages per total women:

$$\frac{HL}{MC} \times \frac{MC/MT}{NL/NT} = \frac{HT}{MT}$$

Where HL = legitimate children born alive (census sample);
HT = total children born alive;
MC = married women (census data corrected with nuptiality data from vital statistics);
MT = total women (census);
NL = legitimate births (vital statistics);
NT = total births (vital statistics).

The estimate of parity from age-specific fertility rates was based on the methodology developed by Brass.[2]

Lastly, the ratios between estimated parity and average number of children born alive from the census sample gave the correction factors for the various age groups in 1960 and 1970.

In order to obtain the corrected averages by level of education, economic activity and area of residence, regardless of age, which appear in table 2, an average correction factor was needed. Therefore, the correction factors per age group were weighted by the overall age structure of the population that appears in table 1.

A more refined technique would have been to weigh the coefficients by age, by the specific age structure of

TABLE 1. RELATIVE DISTRIBUTION OF MARRIED FEMALE POPULATION AGED 20-49 YEARS ACCORDING TO THE DIFFERENT VARIABLES INVESTIGATED, IN CENSUS SAMPLES OF 1960 AND 1970 *(Percentages)*

Age group and variables investigated	*Distribution in:* 1960	1970
Age group		
20-24	14.4	13.4
25-29	19.2	19.6
30-34	20.9	18.7
35-39	18.2	19.9
40-44	15.2	16.4
45-49	12.1	12.0
TOTAL	100.0	100.0
Level of education		
No education	15.3	9.6
1-3 years	20.0	18.6
4-6 years	38.9	39.0
7-9 years	13.8	16.0
10 years or more	12.0	16.8
TOTAL	100.0	100.0
Economic activity		
Active	10.4	11.9
Non-active	89.6	88.1
TOTAL	100.0	100.0
Area of residence		
Capital	29.7	31.4
Urban	41.8	47.8
Rural	28.5	20.8
TOTAL	100.0	100.0

TABLE 2. AVERAGE NUMBER OF CHILDREN BORN ALIVE TO MARRIED WOMEN AGED 20-49 YEARS IN SUBGROUPS OF VARIABLES INVESTIGATED IN CENSUS SAMPLES OF 1960 AND 1970

Age group and variables investigated	*Average number of live-born children* 1960	1970
Age group		
20-24	2.524	2.163
25-29	3.646	2.665
30-34	4.747	3.261
35-39	5.711	3.735
40-44	6.464	4.092
45-49	7.049	5.051
Level of education		
No education	6.560	5.428
1-3 years	5.792	4.505
4-6 years	4.715	3.547
7-9 years	3.636	2.605
10 years or more	3.252	2.262
Economic activity		
Active	3.292	2.572
Non-active	5.071	3.668
Area of residence		
Capital	3.869	2.998
Urban	4.659	3.421
Rural	6.275	4.619

[2] See *Manual IV. Methods of Estimating Basic Demographic Measures from Incomplete Data* (United Nations publication, Sales No. 67.XIII.2).

each subgroup of the variables. The procedure actually employed is based on the assumption that the different structure by age of these groups might have a minor effect on the average correction factors. The validity of this assumption was not tested.

The percentage distribution of the 1960 and 1970 populations is given below according to different variables, as well as the average numbers of children born alive for the same groups and the results of standardization. In these tables and in subsequent calculations, the 15–19 age group was omitted because the data obtained for it did not seem reliable. Moreover, it was found that in this group fertility is low and relatively constant over the years; and, therefore, it was not suited for study in terms of changing factors. One of the more striking changes in structure between 1960 and 1970 was that of the level of education, where a definite increase in level can be discerned for the last year. The increase in urban residence is also noteworthy. In general, all the changes in structure were statistically significant:

$$p < 0.01 \text{ for the proof of } X^2$$

The great difference in the average number of live children born to women in the different categories within the variables investigated will be noted.

The standardized averages, the absolute differences with 1970 averages and the percentage reductions with respect to 1960 are given in table 3.

The averages resulting from standardizing separately by level of education (*I*), economic activity (*AE*) and area of residence (*Z*), as well as from standardizing simultaneously by the three variables, are shown in part (*a*) of table 3. The standardized averages are always greater than the 1970 figures for the same age groups, which shows that the differences between 1970 and 1960 averages cannot be attributed solely to structural changes in these variables but might be due also to others not investigated or to real changes in parity.

In part (*b*) of table 3 it is easy to see that absolute differences between standardized averages and those observed in 1970 increase with age, up to age group 40–44. This means that in the older groups the structural changes play a less important role than in the younger ones.

A similar fact is shown in part (*c*) of table 3, where the absolute differences registered were related to values of averages in 1960 in order to show the percentage reductions of observed and standardized averages.

Under the assumption of independence among the variables considered for standardization, the percentage of reduction attributable to each variable and to the three variables simultaneously, were calculated. For example, in the first age group, 20–24, the percentage of reduction attributable to changes in level of education (*I*) came from: 14.30 – 8.52 / 14.30 – 0.4042 = 40.42 per cent.

This result means that if the structures by level of

TABLE 3. RESULTS OF STANDARDIZATION BY LEVEL OF EDUCATION (*I*), ECONOMIC ACTIVITY (*AE*) AND AREA OF RESIDENCE (*Z*)

(a) *Average figures*

Age group	Averages observed 1970	Averages observed 1960	Averages standardized by *I*	*AE*	*Z*	*I×AE×Z*
20-24	2.163	2.524	2.378	2.524	2.458	2.390
25-29	2.665	3.646	3.456	3.579	3.542	3.428
30-34	3.261	4.747	4.565	4.730	4.654	4.566
35-39	3.735	5.711	5.508	5.664	5.495	5.457
40-44	4.092	6.464	6.248	6.418	6.276	6.210
45-49	5.051	7.049	6.861	7.030	6.846	6.811

(b) *Absolute differences compared with 1970*

Age group	From averages observed in 1960	From averages standardized by *I*	*AE*	*Z*	*I×AE×Z*
20-24	0.361	0.215	0.361	0.295	0.227
25-29	0.981	0.791	0.914	0.877	0.763
30-34	1.486	1.304	1.469	1.393	1.305
35-39	1.976	1.773	1.929	1.760	1.722
40-44	2.372	2.156	2.326	2.184	2.118
45-49	1.998	1.810	1.979	1.795	1.760

(c) *Percentage reductions with respect to 1960*

Age group	From averages observed in 1970	From averages standardized by *I*	*AE*	*Z*	*I×AE×Z*
20-24	14.30	8.52	14.30	11.69	8.99
25-29	26.91	21.70	25.07	24.05	20.93
30-34	31.30	27.47	30.95	29.34	27.49
35-39	34.60	31.04	33.78	30.82	30.15
40-44	36.70	33.35	35.98	33.79	32.77
45-49	28.34	25.68	28.08	25.46	24.97

TABLE 4. CONTRIBUTION OF EACH VARIABLE INVESTIGATED

Age group	Percentage reduction in crude average	Percentage reduction attributable to			
		I	AE	Z	$I \times AE \times Z$
20-24	14.30	40.42	0.00	18.25	37.13
25-29	26.91	19.36	6.84	10.63	22.22
30-34	31.30	12.24	1.12	6.26	12.17
35-39	34.60	10.29	2.37	10.92	12.86
40-44	36.70	9.13	1.96	7.93	10.71
45-49	28.34	9.39	0.92	10.16	11.89

Note: I = level of education; AE = economic activity; Z = area of residence.

education had not changed within the 20–24 age group, the average standardized by this variable would have been the same as that for 1960 (as is the case for AE in this same group). The difference between the percentage of reduction of the crude average and the standardized average, related to the crude average, shows the effect of this change (table 4).

In order to interpret the percentages of reduction attributable to the different variables and to their combinations, an analysis of the components of effects of standardization must be made.[3] Age groups 20–24 and 40–44 were chosen for this analysis, because they represent extreme cases of differences between 1960 and 1970 averages and the contributions of the variables under study. Standardizing separately for each of the three variables and for all their possible combinations, the following standardized averages were obtained:

Variables standardized for	Standardized averages for age group	
	20–24	40–44
I	2.378	6.248
AE	2.524	6.418
Z	2.458	6.276
$I \times AE$	2.399	6.247
$I \times Z$	2.436	6.195
$AE \times Z$	2.464	6.251
$I \times AE \times Z$	2.390	6.210

If the effect E of standardizing is: E = average observed in 1960 minus standardized average, there are in this case three orders of effects:

First order: $E^{(1)}$ = average observed in 1960 minus average standardized by one variable i:

$$\text{Second order: } E^{(2)} = \sum_{1}^{2} E^{(1)} - I^{(2)}$$

where $I^{(2)}$ = the interaction between two variables:

$$\text{Third order: } E^{(3)} = \sum_{1}^{3} E^{(1)} - \sum_{1}^{3} I^{(2)} + I^{(3)}$$

where $I^{(3)}$ = the interaction among three variables.

The upper figures on the summation signs represent the number of effects or interactions that have to be added according to the number of variables considered in each specific case. The variables corresponding to these effects and interaction are put as subindices.

As an example, $E^{(1)}$ is the effect of standardizing by level of education.

For age group 20–24, the 1960 average was 2.524, whereas the average standardized by level of education was 2.378; therefore:

$$E^{(1)}_{I} = 2.524 - 2.378 = 0.146$$

All the effects can be calculated in the same way.

To provide a clear image of the different components of the effects in age groups 20–24 and 40–44, the values for effects and interactions are given below in table 5.

TABLE 5. DIRECT AND INTERACTION EFFECTS

	Age group 20-24		Age group 40-44	
	Absolute value	Percentage[a]	Absolute value	Percentage[a]
$E^{(1)}_{I}$	0.146	40.4	0.216	9.1
$E^{(1)}_{AE}$	0.000	0.0	0.046	2.0
$E^{(1)}_{Z}$	0.066	18.3	0.188	7.9
$I^{(2)}_{I \times AE}$	0.021	5.8	0.045	1.9
$I^{(2)}_{I \times Z}$	0.124	34.3	0.135	5.7
$I^{(2)}_{AE \times Z}$	0.006	1.7	0.021	0.9
$I^{(3)}_{I \times AE \times Z}$	0.073	20.2	0.005	0.2

[a] Proportion of absolute value with respect to difference: observed crude average in 1960 minus observed crude average in 1970.

Note: I = level of education; AE = economic activity; Z = area of residence.

Therefore, the percentage reduction attributable to first-order, second-order and third-order effects are:

First-order effects: $E^{(1)}$

Effect	Age group 20-24	Age group 40-44
$E^{(1)}_{I}$	40.4	9.1
$E^{(1)}_{AE}$	0.0	2.0
$E^{(1)}_{Z}$	18.3	7.9

[3] Analysis based on an unpublished paper by Albino Bocaz, Centro Latinoamericano de Demografia, Santiago, Chile.

Second-order effects: $E^{(2)} = \Sigma E^{(1)} - I^{(2)}$

Effect	Age group 20-24	Age group 40-44
$E^{(1)}_{I}$	40.4	9.1
$E^{(1)}_{AE}$	0.0	2.0
$I^{(2)}_{I \times AE}$	-5.8	-1.9
$E^{(2)}_{I \times AE}$	34.6	9.2

Effect	Age group 20-24	Age group 40-44
$E^{(1)}_{I}$	40.4	9.1
$E^{(1)}_{Z}$	18.3	7.9
$I^{(2)}_{I \times Z}$	-34.3	-5.7
$E^{(2)}_{I \times Z}$	24.4	11.3

Effect	Age group 20-24	Age group 40-44
$E^{(1)}_{AE}$	0.0	2.0
$E^{(1)}_{Z}$	18.3	7.9
$I^{(2)}_{AE \times Z}$	-1.7	-0.9
$E^{(2)}_{I \times Z}$	16.6	9.0

Third-order effects: $E^{(3)} = \Sigma E^{(1)} - \Sigma I^{(2)} + I^{(3)}$

Effect	Age group 20-24	Age group 40-44
$E^{(1)}_{I}$	40.4	9.1
$E^{(1)}_{AE}$	0.0	2.0
$E^{(1)}_{Z}$	18.3	7.9
$I^{(2)}_{E \times AE}$	-5.8	-1.9
$I^{(2)}_{I \times Z}$	-34.3	-5.7
$I^{(2)}_{AE \times Z}$	-1.7	-0.9
$I^{(3)}_{I \times AE \times Z}$	20.2	0.2
$E^{(3)}_{I \times AE \times Z}$	37.1	10.7

In summary, the percentage reductions for these two age groups are:

Effect	Age group 20-24	Age group 40-44
$E^{(1)}_{I}$	40.4	9.1
$E^{(1)}_{AE}$	0.0	2.0
$E^{(1)}_{Z}$	18.3	7.9
$E^{(2)}_{E \times AE}$	34.6	9.2
$E^{(2)}_{I \times Z}$	24.4	11.3
$E^{(2)}_{AE \times Z}$	16.6	9.0
$E^{(3)}_{I \times AE \times Z}$	37.1	10.7

This table clearly shows the greater influence of the structure by the considered variables on the 20–24 age group than on the 40–44 age group. Neither of the standardizations is satisfactory from the point of view of explaining the decrease of averages. By standardizing simultaneously for the three variables, there remains unexplained changes of 62.9 per cent in the 20–24 age group and 89.3 per cent in the 40–44 age group.

The standardization by age leads to an average of 4.934, compared to an average of 3.538 for 1970 and an average of 4.886 for 1960.

The percentage reduction attributable to age is −3.55 per cent. The interpretation of this negative figure is that the age distribution of women in 1970 changed by comparison with 1960 in the sense that there was a slight relative increase of older age groups which showed a greater reduction in the average number of children born alive. This change explains the relative higher decline in the average adjusted by age (1.396) in comparison with the crude figure (1.348). The standardization by age, level of education, economic activity and area of residence yields an average of 4.719. The percentage contribution of the four variables is 12.40 per cent, which leaves 83.6 per cent of reduction unexplained by structural changes with respect to these variables.

Trend analysis

Figure I shows the development of the crude birth rate for the period 1935–1974; the sharp drop in this rate after 1964 is striking. Although the period marked by this drop is that in which development of family planning activities began, changes in fertility obviously cannot be attributed solely to those activities. It was decided to discontinue work with the crude birth rate because it was too approximate an indicator and better sources of information were available; it was therefore decided to make projections of specific fertility rates according to age.

In order to estimate fertility rates according to age group, use was made of information on registered births published by the National Institute of Statistics (INS) and the National Planning Office (ODEPLAN) population projection for the period 1950–2000. Analysis of information on the female population of childbearing age in the ODEPLAN projection showed that population growth rates between 1950 and 1960 for the different age groups did not tally with growth rates for 1960–1970. In order to even out the popula-

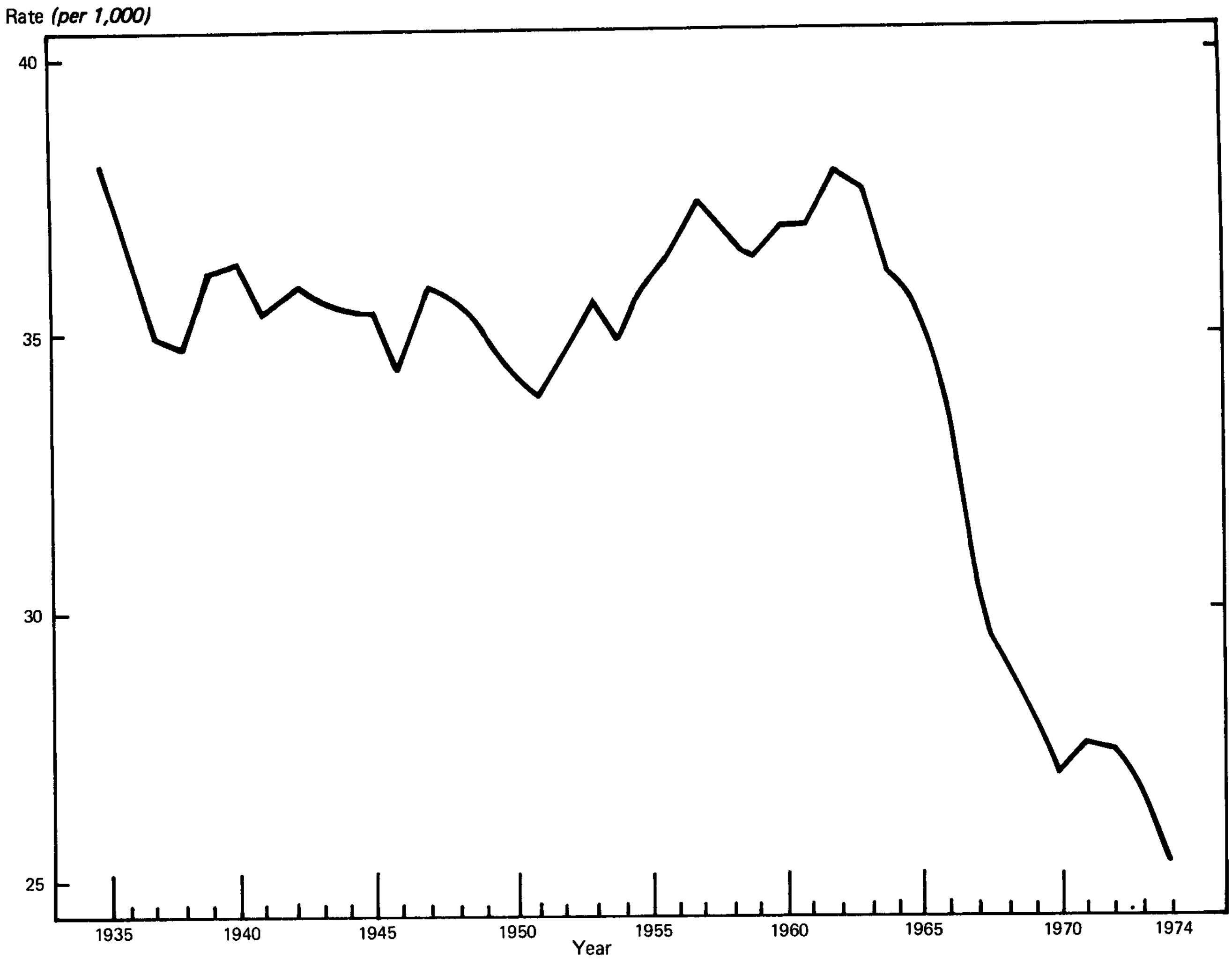


Figure I. Chile: crude birth rate, 1935–1974

Source: Chile, Servicio Nacional de Salud, *Anuario 1974, defunciones y causes de muerte* (Santiago, 1975).

tion series, the rule of a single but specific geometrical variation per age group was adopted for the period 1950–1970. Registered births were also evened out by a geometrical function because they showed sharp annual fluctuations. The values thus obtained were corrected by official estimates of omissions.

Once the two series had been evened out, specific fertility rates were determined for the period 1956–1964, and extrapolated values for the period 1965–1974 were calculated. The resulting fertility rates show a slight variation in time, as can be seen from table 6 and from figure II. Fertility levels in age groups under 20 years were relatively constant in time.

The values recorded for age group 20–24 were less than those projected (see table 7); however, the differences were not as striking as those recorded in the four age groups between 25 and 44 years. Lastly, the projection for women between 45 and 49 years of age was almost identical with the figures recorded. This variation in behaviour could be interpreted in terms of the impact of family planning, that impact being the greatest between the ages of 25 and 44 years, less in women between 20 and 24 years who are still setting up families and practically nil in those under 20 years. Family planning could be regarded as a substitute for abortion in women between 44 and 49 years.

It must be borne in mind that, although probably correct, these interpretations are subjective. Moreover, one must consider the influence on fertility of other factors, such as age of marriage, the use of contraceptives outside the family planning programme, induced abortion and stability of marriage.

Couple-years of protection

The information used for the CYP method refers to the entire country and covers the period 1964–1974. The results presented here refer only to new participants using the intra-uterine device (IUD).

In order to estimate potential fertility, use was made

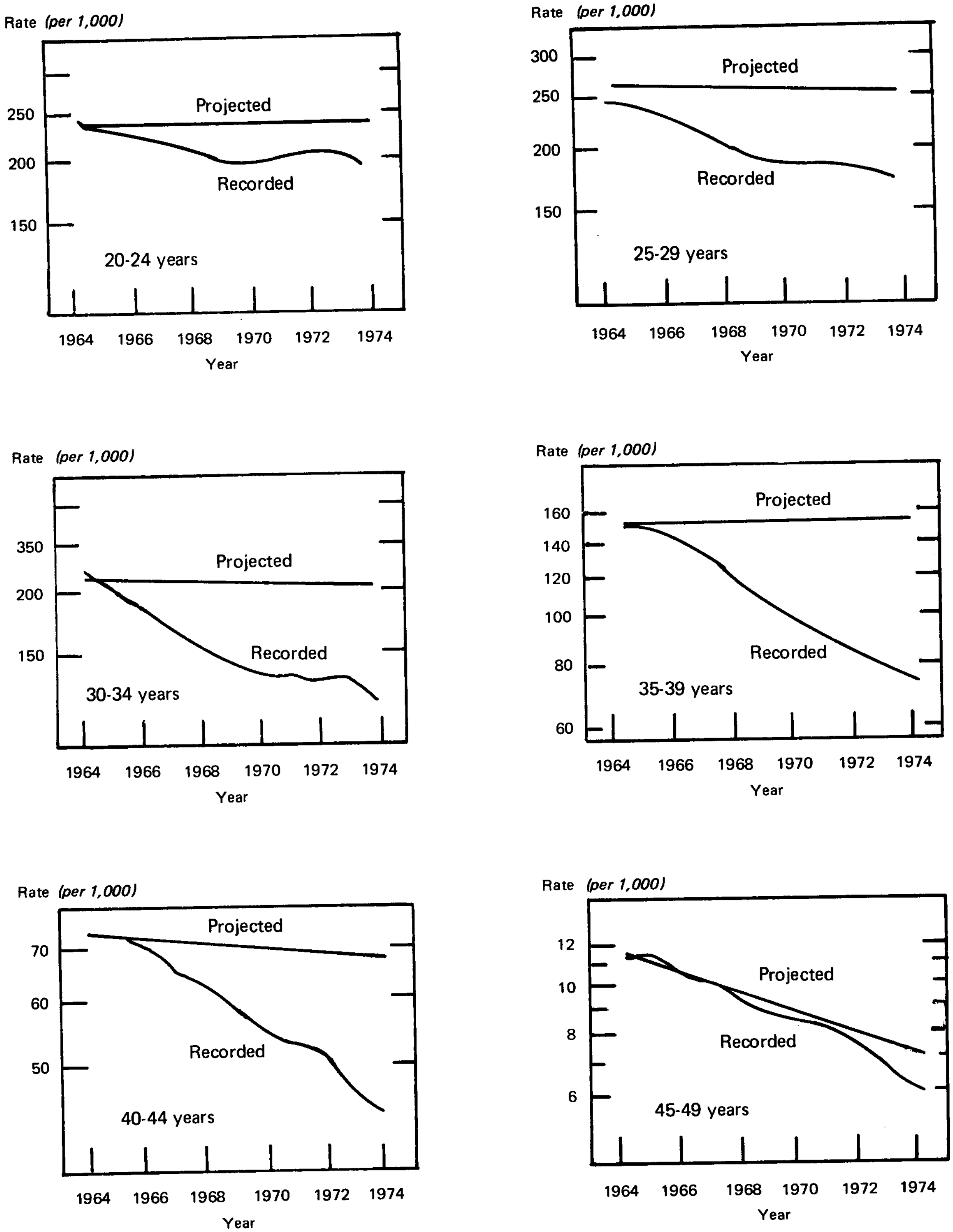


Figure II. Chile: trends in age-specific projected and recorded fertility rates, 1964–1974
(Semi-logarithmic scale)

TABLE 6. RECORDED AND ESTIMATED[a] AGE-SPECIFIC FERTILITY RATES, 1964-1974

Age group	*1964*	*1965*	*1966*	*1967*	*1968*	*1969*	*1970*	*1971*	*1972*	*1973*	*1974*
					Estimated on basis of trends						
12-14	2.7	2.8	2.8	2.8	2.8	2.8	2.9	2.9	2.9	3.0	3.0
15-19	80.8	80.8	80.8	80.8	80.8	80.8	80.7	80.7	80.7	80.7	80.7
20-24	231.6	231.9	232.1	232.4	232.6	232.9	233.1	233.4	233.6	233.9	234.2
25-29	264.7	263.5	262.2	261.0	259.8	258.5	257.3	256.1	254.9	253.7	252.4
30-34	214.2	213.2	212.2	211.2	210.2	209.2	208.2	207.2	206.2	205.9	204.3
35-39	152.8	153.0	153.1	153.3	153.5	153.7	153.9	154.1	154.3	154.5	154.7
40-44	65.1	64.5	63.9	63.3	62.7	62.1	61.5	60.9	60.4	59.8	59.2
45-49	11.6	11.0	10.5	10.0	9.6	9.1	8.7	8.3	7.9	7.5	7.2
					Recorded						
12-14	2.9	2.9	2.8	2.9	2.5	2.7	2.3	2.8	3.1	2.8	3.1
15-19	81.9	82.8	82.1	82.5	79.2	78.9	80.7	84.3	86.4	83.5	79.3
20-24	235.4	227.9	220.2	215.0	206.2	194.8	192.7	200.1	203.0	202.5	191.3
25-29	247.5	242.7	230.5	213.2	200.0	186.1	182.5	186.0	183.6	175.5	169.4
30-34	220.1	207.4	186.3	168.8	152.1	141.6	135.0	136.2	131.0	123.9	116.9
35-39	150.4	149.5	142.4	120.4	116.0	105.2	94.4	88.7	82.3	76.4	73.4
40-44	65.5	64.5	60.3	55.3	52.9	48.0	44.7	42.4	40.0	35.4	33.5
45-49	11.1	11.4	10.4	9.9	8.9	8.4	8.4	7.9	7.3	6.3	6.0

[a] Estimated by geometrical adjustment of the trend for 1956-1964.

of the average rates for 1961–1963, the period immediately preceding the official beginning of family planning activities. The over-all rate of marital fertility was adjusted to the age structure of new participants in order to obtain a more realistic estimate of births avoided by IUD users. Information on new participants, according to method, was obtained from various sources for the period 1964–1974 (see annexed table 37). The continuation rates for IUDs were taken from an application of the Service Statistics System of CELADE in the National Employees' Medical Service (see annexed table 39). The age structure of new participants was also drawn from that source (see annexed table 38).

TABLE 7. ESTIMATED AND RECORDED BIRTHS AND DIFFERENCES BETWEEN ESTIMATES AND TOTALS RECORDED, BY AGE GROUP, 1964-1975

Age group	*1964*	*1965*	*1966*	*1967*	*1968*	*1969*	*1970*	*1971*	*1972*	*1973*	*1974*	*1975*
					Projected birth estimates							
12-14	723	745	783	801	819	838	860	895	911	938	965	994
15-19	33 737	34 725	35 742	36 789	37 867	38 976	40 117	41 292	42 502	43 746	45 028	46 346
20-24	82 401	84 922	87 521	90 199	92 958	95 802	98 734	101 754	104 868	108 076	111 383	114 791
25-29	79 712	80 614	81 527	82 450	83 384	84 328	85 283	86 249	87 226	88 213	89 212	90 222
30-34	56 215	56 851	57 495	58 146	58 804	59 470	60 144	60 825	61 513	62 210	62 915	63 627
35-39	37 061	38 059	39 085	40 138	41 219	42 330	43 470	44 641	45 844	47 079	48 348	49 650
40-44	13 206	13 266	13 326	13 387	13 448	13 509	13 571	13 633	13 695	13 757	13 820	13 883
45-49	2 113	2 046	1 981	1 919	1 838	1 799	1 743	1 688	1 634	1 583	1 533	1 484
					Recorded births							
12-14	771	763	761	846	755	831	751	912	1 054	973	1 121	...
15-19	33 960	35 057	35 356	36 060	35 157	35 561	36 973	40 254	42 931	43 189	42 538	...
20-24	84 253	84 251	83 479	83 547	82 086	79 381	80 323	84 808	87 397	88 572	85 004	...
25-29	73 774	73 998	72 993	70 013	67 993	65 450	66 309	69 368	70 202	68 781	65 993	...
30-34	58 103	54 925	50 591	47 014	43 435	41 404	40 385	42 345	42 265	41 475	40 414	...
35-39	36 995	38 002	36 364	33 408	29 857	27 166	24 508	23 625	22 477	21 384	21 054	...
40-44	13 323	13 311	12 959	12 349	12 227	11 500	11 072	10 580	9 992	8 900	8 451	...
45-49	2 029	2 098	1 946	1 877	1 736	1 649	1 662	1 645	1 578	1 440	1 400	...
					Differences between estimates and totals recorded							
12-14	–48	–18	22	–45	64	7	109	–27	–143	–36	–156	...
15-19	–223	–332	386	729	2 710	3 415	3 144	1 038	–429	557	2 490	...
20-24	–1 852	671	4 042	6 652	10 872	16 421	18 411	16 946	17 471	19 504	26 379	...
25-29	5 938	6 616	8 534	12 437	15 391	18 878	18 974	16 891	17 024	19 432	23 219	...
30-34	–1 888	1 926	6 904	11 132	15 369	18 066	19 759	18 480	19 248	20 795	22 501	...
35-39	66	57	2 721	6 730	11 362	15 164	18 962	21 016	23 367	25 695	27 294	...
40-44	–117	–45	367	1 038	1 221	2 009	2 499	3 053	3 703	4 857	5 369	...
45-49	84	–52	35	42	122	150	81	43	56	143	133	...
TOTAL	1 960	8 823	23 011	38 715	57 111	74 110	81 939	77 440	80 297	90 948	107 229	...

TABLE 8. RELATIVE REDUCTION IN BIRTHS DUE TO USE OF THE INTRA-UTERINE DEVICE IN THE FAMILY PLANNING PROGRAMME, BY AGE GROUP, 1965-1970, 1971-1974 AND 1965-1974

Age group	*Live births*		*Reduction in births (All causes)*	*Births averted (component projection)*	*Relative reduction due to use of intra-uterine device (percentage)*
	Recorded value	*Trend value*			
		1965-1970			
12-14	4 707	4 846	139	...	...
15-19	214 164	224 216	10 052	...	...
20-24	493 067	550 136	57 069	41 429	72.6
25-29	416 756	497 586	80 830	47 230	58.4
30-34	277 754	350 910	73 156	32 066	43.8
35-39	189 305	244 301	54 996	13 305	24.2
40-44	73 418	80 507	7 089	2 752	38.1
45-49	10 968	11 346	378	...	...
		1971-1974			
12-14	4 060	3 699	−361	...	...
15-19	168 912	172 568	3 656	...	...
20-24	345 781	426 081	80 300	58 436	72.8
25-29	274 334	350 900	76 566	87 762	114.6
30-34	166 439	247 463	81 024	64 731	79.9
35-39	88 540	185 912	97 372	28 136	28.9
40-44	37 923	54 905	16 982	6 435	37.9
45-49	6 063	6 438	375	...	...
		1965-1974			
12-14	8 767	8 545	−222	...	...
15-19	383 076	396 784	13 708	...	...
20-24	838 848	976 217	137 369	99 865	72.7
25-29	691 090	848 486	157 396	134 992	85.8
30-34	444 193	598 373	154 180	96 797	62.8
35-39	277 845	430 213	152 368	41 441	27.2
40-44	111 341	135 412	24 071	9 187	38.2
45-49	17 031	17 784	753	...	...

Methodology and results

The estimates of annual number of births averted were based on the CYP prevalence index,[4] taking continuation rates from the experience in SERMENA:

CYP prevalence index for year

$$T = \frac{a}{r}\sum_{i=0}^{j} N_{T-i}\ \{e^{-ri} - e^{-r(i+1)}\}$$

where $a = 0.9884$ = continuation rate at the end of the first month;

$r = 0.2150$ = rate of attrition during one year;

N_{T-i} = number of insertions in year $T-i$;

i = years of participation in the programme—i varies from 0 to j; $j = 5$.

The addenda of this formula: $N_{T-i}\frac{a}{r}\{e^{-ri} - e^{-r(i+1)}\}$ are shown in the columns of table 9, and the CYP prevalence indexes for each year appear in the total row at the bottom of this table.

[4] Samuel M. Wishik and K. H. Chen, *The Couple-Year of Protection: A Measure of Family Planning Program Output,* Manuals for Evaluation of Family Planning and Population Programs, No. 7 (New York, Columbia University, International Institute for the Study of Human Reproduction, 1973).

The coefficients that multiply N_{T-i} in each of the six years of permanence considered in the programme (i = 0,....,5) are the same for every cohort:

i	Coefficient: $\frac{a}{r}\{e^{-ri} - e^{-r(i+1)}\}$
0	0.8223
1	0.7307
2	0.5890
3	0.4750
4	0.3832
5	0.3090

Therefore, each row of table 9 shows the successive products of new participants and these coefficients.

For example, in the row of the year 1964, the 11,264 new participants multiplied by 0.8223 yield a prevalence of 9,262 for 1964, multiplied by 0.7307 a prevalence of 8,231 for 1965 and so on.

Average marital fertility rates for 1961–1963 were used to estimate the number of births avoided. The information available was specific fertility rates, which were converted into marital fertility rates using a conversion factor corresponding to the ratio between the percentage of legitimate births and the percentage of married women (see table 10).

TABLE 9. CYP PREVALENCE: $\frac{a}{r}\sum_{i=0}^{5} N_{T-i}\{e^{-ri}-e^{-r(i+1)}\}$

($a = 0.9884$; $r = 0.2150$)

Year	New participants	1964	1965	1966	1967	1968	1969	1970	1971	1972	1973	1974
1974	121 879											100 221
1973	80 155										65 911	58 569
1972	35 167									28 918	25 697	20 713
1971	36 142								29 720	26 409	21 288	17 167
1970	43 602							35 854	31 860	25 682	20 711	16 708
1969	40 674						33 446	29 720	23 957	19 320	15 586	12 568
1968	46 422					38 173	33 921	27 343	22 050	17 789	14 344	
1967	45 375				37 312	33 156	26 726	21 553	17 388	14 021		
1966	33 686			27 700	24 614	19 841	16 001	12 908	10 409			
1965	20 467		16 830	14 955	12 055	9 722	7 843	6 324				
1964	11 264	9 262	8 231	6 634	5 350	4 316	3 481					
TOTAL		9 262	25 061	49 289	79 331	105 208	121 418	133 702	135 384	132 139	163 537	225 946

On the assumption that the percentage of married women according to age group does not vary significantly in the short term, information available for the 1960 census was used. On the same assumption, the percentage of legitimate births in turn corresponds to the average of the percentages for 1959–1961:

$$\text{Marital fertility rates} = \frac{N_i}{M_i} \times C_i$$

$$C_i = \frac{NL_i}{N_i} \Big/ \frac{MC_i}{M_i}$$

where N_i = births to mothers in age group i;
M_i = total number of women in age group i;
NL_i = legitimate births to mothers in age group i;
MC_i = married women in age group i;
C_i = conversion factor for group i.

If the rates for women of 20 years and over are correct, those for the women between 15 and 19 years must be underestimated. No adjustment was made in view of the small proportion which this group represents in the total, although the rate for the Hutterites could, for example, have been used as a reference to make an adjustment. One of the methods of estimating the number of live births averted consisted in applying the over-all marital fertility rate to couple-years of protection, which were obtained using the achievement and prevalence indexes. In order to make a more precise estimate, the age distribution of new participants and its evolution during their years of participation in the programme were taken into consideration.

TABLE 10. AGE-SPECIFIC FERTILITY RATES
(Rates per 1 000)

Age group	Fertility rate, 1961-1963 average	C_i	Marital fertility rate
15-19	82.47	6.41	528.63
20-24	235.13	2.35	552.36
25-29	265.90	1.52	404.17
30-34	240.53	1.30	312.69
35-39	156.10	1.24	193.56
40-44	67.10	1.26	84.55
45-49	12.83	1.33	17.06
TOTAL	141.20	1.70	240.04

The five-year age groups were divided up into one-year age groups, adopting a logarithmic variation of numbers, and women who had reached the highest age in a given group were transferred to the following group. This procedure resulted in the percentage distribution given in table 11.

By weighting the specific rates of marital fertility for the age structure in successive years, over-all potential fertility rates for the different age groups were obtained for each year of participation in the programme (see table 12).

The need to consider the age structure of participants becomes obvious if one examines the weighted rates obtained and compares them with the over-all marital fertility rate of 240.04 in the population as a whole. For the particular distribution assumed in the present case, a comparison was made between live births avoided per woman in the programme, estimated on the basis of the over-all marital fertility rates, with and without adjustment (table 13).

It should be pointed out that, for the purposes of the estimate, it was assumed that births were averted, as from the year following that in which the women concerned joined the programme, when they were already one year older; for this reason, the rates corresponding to the following year were applied to couples who had joined the programme the previous year.

Thus, for the six-year period of participation considered, there would in this case be a 28.3 per cent underestimation if the age of new participants was not considered. The need for information on rates of continued participation is reflected in the following comparison (table 14) between cumulative couple-years of protection obtained with the prevalence rates used in this report and those used in Pakistan: 0.75 for the first year; 0.50 for the second; and 0.35 for the third.

If one assumes that the over-all fertility rate is 240.04 per 1,000, one obtains the estimates of live births averted given in table 15. Thus, for the period 1965–1970, live births averted totalled 93,512.38; and for 1965–1974, they reached a total of 229,077.60.

Using marital fertility rates adjusted by age of new participants in successive years, one obtains the fig-

TABLE 11. RELATIVE DISTRIBUTION BY AGE GROUP IN SUCCESSIVE YEARS OF PARTICIPATION IN THE PROGRAMME

Age group	Number of years of participation in programme						
	Under one	One	Two	Three	Four	Five	Six
15-19	4.0	2.47	1.45	0.77	0.31	0.00	0.00
20-24	30.0	25.61	20.67	15.35	9.77	4.00	2.47
25-29	29.0	29.84	30.39	30.61	30.49	30.00	25.61
30-34	21.0	22.77	24.47	26.10	27.62	29.00	29.84
35-39	12.0	13.84	15.71	17.57	19.36	21.00	22.77
40-44	4.0	4.98	6.21	7.73	9.63	12.00	13.84
45-49	– [a]	0.49	1.10	1.87	2.82	4.00	4.98
50 and over	–	–	–	–	–	–	0.49

[a] It was assumed that there were no new participants aged over 45 but that women who joined the programme between the ages of 40 and 44 continued to participate.

ures for births averted per couple and over-all. On this basis, for the period 1965–1970, live births averted totalled 138,379.1; for 1965–1974, the total was 328,703.8 (see table 16).

The live births averted per couple up to 1970 and up to 1974 are based on the figures given in the last column of part (*b*) of table 13.

As can be seen, the live births averted by a couple entering the programme in 1964 are the same up to 1970 and up to 1974 since in both cases it was assumed that a six-year duration constituted permanence in the programme.

Since a woman entering in 1965 can only be counted for five years up to 1970, assuming that she begins averting births in the year following that of her entrance, the live births averted reduce to 1.02593 in this group, whether up to 1974 she still will be able to avoid 1.10845 births in a six-year span. Therefore, the columns of live births averted per couple differ according to the time left from the time of entrance up to the year of the evaluation.

COMPONENT PROJECTION APPROACH

The data used for the component projection approach relate to the country as a whole and, like the couple-years of protection method, cover the period 1964–1974. The method was applied only to data on IUD users.

Measurement of fertility and estimation of other parameters

Vital statistics were used to obtain the data needed to calculate age-specific fertility rates, using as denominators the official population projections of the National Planning Office. The average rates for 1961–1963 were taken as the base fertility, as this was the period immediately preceding the official commencement of family planning activities.

To determine the actual user population, use was made of information on available acceptors (see annexed table 35). The age distribution of participants and the continuation rates for the purpose of estimating drop-outs were calculated on the basis of data from the SERMENA experiment (see tables 36 and 37). The survivorship of men and women was obtained from life tables for Chile compiled by Tacla and Pujol.[5] In applying the method, no account was taken of women in age group 45–49, since the number of participants in this age group is negligible in comparison with the other groups. For purposes of comparison with the CYP figures calculated for six years, it was assumed that all the women remained in the programme for six years, regardless of age.

[5] Odette Tacla and José M. Pujol, *Tablas abreviadas de mortalidad 1952–1953 y 1960–1961*, CELADE Series C, No. 11 (Santiago, 1965).

TABLE 12. MARITAL FERTILITY RATES WEIGHTED BY RELATIVE AGE DISTRIBUTION

Age group	Marital fertility rate	Number of years of participation in programme						
		Under one	One	Two	Three	Four	Five	Six
15-19	528.63	21.14	13.06	7.66	4.07	1.64	–	–
20-24	552.56	165.77	141.51	114.21	84.82	53.98	22.10	13.65
25-29	404.17	117.21	120.60	122.83	123.72	123.23	121.25	103.51
30-34	312.69	65.66	71.20	76.51	81.61	86.36	90.68	93.31
35-39	193.56	23.23	26.79	30.41	34.01	37.47	40.65	44.07
40-44	84.55	3.38	4.21	5.25	6.54	8.14	10.15	11.70
45-49	17.06	–	0.08	0.19	0.32	0.48	0.68	0.85
Over-all marital fertility rate		396.39	377.45	357.06	335.09	311.30	285.51	267.06

TABLE 13. LIVE BIRTHS AVOIDED PER COUPLE, CALCULATED BY VARIOUS METHODS

Year	*Proportion of protected couples*	*Marital fertility rate*	*Live births avoided per couple*	*Cumulative live births avoided*
		(a) *Unadjusted*		
Under one	0.8223	0.24004	0.19738	0.19738
One	0.7307	0.24004	0.17540	0.37278
Two	0.5890	0.24004	0.14138	0.51416
Three	0.4750	0.24004	0.11402	0.62818
Four	0.3832	0.24004	0.09198	0.72016
Five	0.3090	0.24004	0.07417	0.79433
		(b) *Adjusted*		
Under one	0.8223	0.37745	0.31038	0.31038
One	0.7307	0.35706	0.26090	0.57128
Two	0.5890	0.33509	0.19737	0.76865
Three	0.4750	0.31130	0.14787	0.91652
Four	0.3832	0.28551	0.10941	1.02593
Five	0.3090	0.26706	0.08252	1.10845

(c) *Comparison*

	Cumulative live births avoided per couple		*Percentage underestimation due to lack of adjustment*
	Unadjusted	*Adjusted*	
Under one	0.19738	0.31038	36.4
One	0.37278	0.57128	34.7
Two	0.51416	0.76865	33.1
Three	0.62818	0.91652	31.5
Four	0.72016	1.02593	29.8
Five	0.79433	1.10845	28.3

TABLE 14. CUMULATIVE PREVALENCE FOR NEW WOMEN PARTICIPANTS

Year	*Pakistan*	*Chile*	*Percentage underestimation under Pakistan assumptions*
Under one	0.75	0.8223	8.79
One	1.25	1.5530	19.51
Two	1.60	2.1420	25.30

TABLE 15. LIVE BIRTHS AVERTED PER PROGRAMME YEAR, ESTIMATED FROM PREVALENCE INDEX

Year	*Protected couples, prevalence*	*Live births averted per couple protected the previous year*	
		Annual	*Cumulative*
1964	9 263		
1965	25 061	2 223.49	2 223.49
1966	49 289	6 015.64	8 239.13
1967	79 331	11 831.33	20 070.46
1968	105 208	19 042.61	39 113.07
1969	121 418	25 254.13	64 367.20
1970	133 702	29 145.18	93 512.38
1971	135 384	32 093.83	125 606.21
1972	132 139	32 497.57	158 103.78
1973	163 536	31 718.64	189 822.42
1974	225 946	39 255.18	229 077.60
1975	...	54 236.08	283 313.68

TABLE 16. LIVE BIRTHS AVERTED PER COUPLE AND OVER ALL, 1964-1970 AND 1964-1974

Year	*New participants*	*Live births averted per couple up to 1970*	*Live births averted up to 1970*	*Live births averted per couple up to 1974*	*Live births averted up to 1974*
1964	11 264	1.10845	12 485.6	1.10845	12 485.6
1965	20 467	1.02593	20 997.7	1.10845	22 686.6
1966	33 686	0.91652	30 873.9	1.10845	37 339.2
1967	45 375	0.76865	34 877.5	1.10845	50 295.9
1968	46 422	0.57128	26 250.0	1.10845	51 456.5
1969	40 674	0.31038	12 624.4	1.02593	41 728.7
1970	43 602			0.91652	39 962.1
1971	36 142			0.76865	27 780.5
1972	35 167			0.57128	20 090.2
1973	80 155			0.31038	24 878.5
1974	121 879				
TOTAL	514 833		138 379.1		328 703.8

Procedures and results

It was necessary to obtain:

$$f_{i,t} = \frac{F_{i,t}\, f_{i,o} - Q_{i,t}\, G_i}{F_{i,t}}$$

where $f_{i,t}$ = fertility of women in age-group i in year t;
$f_{i,o}$ = fertility rate of women in age-group i, before beginning the programme;
$F_{i,t}$ = total number of women in age group i in year t;
G_i = potential fertility of users in group i;
$Q_{i,t}$ = women in age group i actually using IUD in year $t-1$;
i = 1, ... 6: age groups 15-19, ... 40-44.

The average fertility rates for 1961–1963 were taken for $f_{i,o}$ and the corresponding marital fertility rates calculated for the CYP figures, increased by 20 per cent to adapt them to the present method, were taken for G_i (table 17).

TABLE 17. FERTILITY RATES, BY AGE GROUP, 1961-1963

Age group	*Average fertility rates, 1961-1963 ($f_{i,o}$)*	*Marital fertility rates*	*Potential fertility of users (G_i)*
15-19	82.47	528.63	634.36
20-24	235.13	552.56	663.07
25-29	265.90	404.17	485.00
30-34	240.53	312.69	375.23
35-39	156.10	193.56	232.27
40-44	67.10	84.55	101.46

On the basis of the age distribution of new participants by conventional groups, a breakdown by individual ages was made by means of geometrical interpolation, and transference of the figures for women at the age-group limits was used to obtain the distribution for women one year younger (table 18).

The percentage of new participants actually using IUD, which was estimated at 70 per cent by component projection, was estimated in this case at 74.825 per cent. This figure was obtained by averaging the first- and second-year continuation rates observed in SERMENA: first year, 81.16 per cent; second year, 68.49 per cent; average, 74.825 per cent.

TABLE 18. AGE DISTRIBUTION OF NEW PARTICIPANTS USING AN INTRA-UTERINE DEVICE

Conventional age group	*Percentage*	*Age group one year younger*	*Percentage*
15-19	4.00	14-18	2.67
20-24	30.00	19-23	25.61
25-29	29.00	24-28	29.84
30-34	21.00	29-33	22.77
35-39	12.00	34-38	13.84
40-44	4.00	39-43	4.98

The correction for female mortality and widowhood was approximated by calculating survivorship over five years (until reaching the following age group) of women of average age on admission to the programme (28.5 years). For widowhood, it was estimated that men are, on average, three years older than women:

Women	*Men*
$l_{28.5} = 0.82754$	$l_{31.5} = 0.78750$
$l_{33.5} = 0.81317$	$l_{36.5} = 0.76482$

$$\underset{\text{Women}}{\frac{l_{33.5}}{l_{28.5}}} \times \underset{\text{Men}}{\frac{l_{36.5}}{l_{31.5}}} = 0.95434$$

Consequently, the admissions to the programme were corrected by 0.74825 for continuation and by 0.95434 for mortality and widowhood, giving a combined correction factor of 0.714085.

When $Q_{i,t}$ had been obtained for each year, the next step was to calculate the anticipated fertility rates in each age group in subsequent years of the programme, as a result of the use of IUD contraception ($f_{i,t}$) (see table 19).

A comparison of the fertility rates calculated for 1970–1974 with those of the base period 1961–1963 gave the results shown in table 20.

For both years it can be seen that the decline is greatest in the 25–29 age group, followed closely by the 30–34 age group.

TABLE 19. FERTILITY RATES EXPECTED AS A RESULT OF USE OF INTRA-UTERINE DEVICES, 1965-1974

Year	Fertility rate in age group					
	15-19	20-24	25-29	30-34	35-39	40-44
1965	82.33	233.39	264.01	239.19	155.56	66.98
1966	81.96	228.99	259.00	235.47	153.98	66.64
1967	81.41	222.42	250.97	229.19	151.15	66.02
1968	80.72	214.14	239.93	220.11	146.86	65.06
1969	80.11	206.64	228.03	209.57	141.55	63.85
1970	79.82	202.61	217.95	199.43	135.97	62.52
1971	79.81	200.97	211.41	193.31	132.37	61.42
1972	79.96	201.39	208.35	190.11	130.19	60.62
1973	80.16	203.60	209.58	191.02	130.20	60.40
1974	79.96	199.93	207.68	190.67	129.98	60.32

TABLE 20. DECLINE IN FERTILITY RATES IN 1970 AND 1974, COMPARED WITH THE BASE PERIOD, 1961-1963

Age group	Base rate	Estimated rate		Percentage decline with respect to the base year	
		1970	1974	1970	1974
15-19	82.47	79.82	79.96	3.2	3.0
20-24	235.13	202.61	199.93	13.8	15.0
25-29	265.90	217.95	207.68	18.0	21.9
30-34	240.53	199.43	190.67	17.1	20.7
35-39	156.10	135.97	129.98	12.9	16.7
40-44	67.10	62.52	60.32	6.8	10.1

Births avoided were estimated for 1965–1970 and 1965–1974 by adding, in each age group, the Q_iG_c values calculated for the corresponding periods (see table 21).

TABLE 21. LIVE BIRTHS AVERTED, 1965-1970 AND 1965-1974

Age group	1965-1970	1965-1974
15-19	3 795	8 853
20-24	41 429	99 865
25-29	47 230	134 992
30-34	32 066	96 797
35-39	13 305	41 441
40-44	2 752	9 187
TOTAL	140 577	391 135

REGRESSION ANALYSIS

The 25 provinces of Chile were used as observation units and regressions were calculated for the years 1960 and 1970 (see tables 22 and 23).

Measurement of fertility and estimation of other parameters

The fertility measurement used was the gross reproduction rate (variable 1), derived from birth data from the National Health Service and population data from the census. Since data on births by age groups of mothers for each province were only published from 1963 onward, the age structure for 1960 was obtained by linear extrapolation based on the years 1963–1965 and 1968–1970.

The following variables were used:

(1) Gross reproduction rate;

(2) Average of age-specific percentages of urban residence of women between the ages of 10 and 49;

(3) Average of age-specific percentages of illiteracy of women between the ages of 15 and 49;

(4) Average of age-specific percentages of economically active women between the ages of 15 and 49;

(5) Percentage of economically active population engaged in agriculture, aged 12 and over, both sexes;

(6) Percentage of births without professional attention;

(7) Percentage of women of childbearing age covered by the family planning programme.

Variables 2–5 were obtained from the 1960 and 1970 censuses.

In order to reproduce the uniform age distribution implicit in the gross reproduction rate, the average of the percentages by age group was taken for variables 2, 3 and 4. Since variable 5 was not broken down by sex, use was made of the over-all percentage of economically active population aged 12 and over engaged in agriculture. Variable 6, obtained from the information on births published by the National Health Service, was selected as an indicator of the existence of general medical facilities, as no data were available on personnel engaged in family planning activities and it was assumed that there would be a direct relationship between the number of such personnel and general facilities. The coverage datum (variable 7) was kindly provided by APROFA. It exists only for 1970 because in 1960 family planning activities had not yet been

TABLE 22. CHILE: VARIABLES FOR THE STUDY OF REGRESSION, BY PROVINCE, 1960

Province	Variable No. 1	2	3	4	5	6
Tarapacá	2.34	90.20	7.79	11.56	14.01	17.1
Antofagasta	2.29	96.27	5.93	10.57	3.02	9.8
Atacama	2.95	79.16	11.64	17.11	11.95	44.8
Coquimbo	2.89	56.82	20.94	19.27	36.53	54.8
Aconcagua	2.59	61.16	16.41	20.01	45.38	34.7
Valparaíso	2.05	91.17	6.04	25.86	12.63	14.8
Santiago	2.10	92.33	7.83	33.38	8.45	13.0
O'Higgins	2.97	58.78	17.29	16.23	48.49	35.3
Colchagua	3.23	37.88	27.28	15.05	63.14	56.6
Curicó	3.29	48.21	25.11	17.51	58.11	54.3
Talca	2.89	50.84	23.15	20.70	50.93	46.8
Maule	2.59	46.19	23.28	16.23	60.29	58.7
Linares	3.05	41.31	23.78	16.42	58.44	50.9
Ñuble	2.91	45.67	26.35	17.28	59.91	62.2
Concepción	2.80	84.85	15.17	23.13	16.68	36.0
Arauco	3.38	39.60	32.79	14.05	46.78	68.9
Bío-Bío	2.94	43.20	29.01	16.78	55.70	61.5
Malleco	3.39	50.61	31.66	16.14	53.63	55.1
Cautín	2.48	43.94	29.17	16.30	55.76	69.3
Valdivia	3.22	49.26	24.41	17.20	45.59	55.8
Osorno	2.80	52.17	22.48	21.61	46.29	58.9
Llanquihue	3.01	46.46	21.99	18.60	47.82	61.0
Chiloé	2.69	23.04	18.72	26.32	67.38	81.0
Aysén	3.25	58.16	21.10	17.64	42.33	43.0
Magallanes	1.49	93.14	6.39	23.28	17.00	3.0

Sources: For variables 1 and 6, Chile. Servicio Nacional de Salud, *Anuario de nacimientos, 1960;* for variables 2-5, data of 1960 census.

Notes: Variable 1: gross reproduction rate;
Variable 2: average percentage of women residing in urban areas;
Variable 3: average percentage of illiterate women;
Variable 4: average percentage of economically active women;
Variable 5: percentage of economically active population engaged in agriculture;
Variable 6: percentage of births without professional attention.

officially introduced. Since the coverage corresponds to the balance between participants joining and leaving the programme between 1964 and 1970, for each province, and therefore depends upon developments during that period, it appeared appropriate to relate it to fertility in 1970. The coverage relates only to institutional data, mainly those of the National Health Service and SERMENA, which cover almost 90 per cent of the population.

Procedures and results

After constructing the zero-order correlation matrices with all the variables for 1960 and 1970 (tables 24 and 29), a forward stepwise inclusion procedure was applied to arange the variables in order of importance (tables 25 and 30). The following order was obtained:

1960: 3, 4, 6, 5, 2 $R^2 = 0.67093$
1970: 3, 5, 4, 7, 6, 2 $R^2 = 0.84957$

If variable 5 (percentage of economically active population engaged in agriculture) (tables 26 and 31) is excluded, one finds that the order is as follows:

1960: 3, 4, 6, 2
1970: 3, 4, 7, 6, 2

The regression with variables 3 and 4 for 1960 and variables 3, 4 and 7 for 1970 was then calculated (tables 27 and 32):

1960: 3, 4 $R^2 = 0.62713$
1970: 3, 4, 7 $R^2 = 0.80785$

Using the component regression method[6] in which the measurement of the total variation is the sum of all elements of the zero-order correlation matrix, the following order was found:

1960: 3, 4, 1, 6, 5, 2
1970: 3, 7, 4, 1, 6, 5, 2

This means that dependent variable 1 is explained by 3 and 4 in 1960 and by 3, 7 and 4 in 1970.

In this way, both procedures gave a similar order, according greatest importance, among the selected variables, to the indicators of level of education and economic activity and, also, coverage in 1970.

In the authors' opinion, if information had been available for measuring the level of education by years

[6] Per Ottestad, "Component analysis: an alternative system". *International Statistical Review*, vol. 43, No. 1 (April 1975).

TABLE 23. CHILE: VARIABLES FOR THE STUDY OF REGRESSION, BY PROVINCE, 1970

Province	Variable No. 1	2	3	4	5	6	7
Tarapacá	2.02	92.91	5.39	21.76	11.78	7.1	14.53
Antofagasta	1.94	97.51	4.06	17.30	2.78	5.2	10.23
Atacama	2.13	88.54	8.80	16.54	9.03	18.2	5.40
Coquimbo	2.21	65.19	13.99	17.89	26.59	32.3	5.06
Acomcagua	1.86	65.29	12.07	20.64	31.16	17.2	8.19
Valparaíso	1.55	92.93	4.49	24.25	9.54	5.8	12.27
Santiago	1.58	94.72	5.82	30.84	5.79	6.6	15.18
O'Higgins	2.12	61.41	12.22	17.71	35.72	20.9	8.41
Colchagua	2.40	41.72	20.05	15.65	53.80	31.1	1.16
Curicó	2.19	52.58	18.25	18.70	48.15	29.0	3.55
Talca	2.12	59.27	17.10	20.70	44.06	34.6	4.50
Maule	2.08	51.55	18.95	16.81	51.70	39.9	3.68
Linares	2.44	48.43	17.60	17.60	53.58	30.4	0.57
Ñuble	2.58	53.02	19.16	16.82	49.63	42.5	10.41
Concepción	2.06	88.33	10.09	22.96	13.27	20.7	4.62
Arauco	2.86	54.21	23.36	13.81	43.17	47.1	0.26
Bío-Bío	2.59	54.58	21.66	16.48	48.11	36.2	0.58
Malleco	2.88	56.65	25.16	15.59	49.08	35.8	5.75
Cautín	2.36	54.82	18.84	17.67	48.22	38.2	3.89
Valdivia	2.34	58.25	16.87	17.24	41.66	32.6	11.91
Osorno	2.15	59.58	15.48	18.88	41.24	31.3	2.36
Llanquihue	2.40	55.30	16.05	19.20	43.92	34.8	6.16
Chiloé	2.00	34.33	13.64	24.78	60.28	49.1	5.27
Aysén	2.62	69.18	15.66	20.96	36.11	28.2	10.02
Magallanes	1.45	91.06	4.17	22.78	17.42	2.1	9.95

Sources: For variables 1 and 6, Chile, Servicio Nacional de Salud, *Anuario de nacimientos, 1970;* for variables 2-5, data of 1970 census; for variable 7, Asociación Chilena de Protección de la Familia (APROFA).

Notes: Variable 1: gross reproduction rate;
Variable 2: average percentage of women residing in urban areas;
Variable 3: average percentage of illiterate women;
Variable 4: average percentage of economically active women;
Variable 5: percentage of economically active population engaged in agriculture;
Variable 6: percentage of births without professional attention;
Variable 7: percentage of coverage of family planning programmes.

of school attendance instead of illiteracy, the present estimates might have been better.

The regressions of variable 1 on variables 3 and 4 for 1960 and of variable 1 on variables 3, 7 and 4 for 1970 gave the following F figures:

1960: F: $18.50, F_{2.22;\ 0.99} = 5.72$
1970: F: $29.43, F_{3.21;\ 0.99} = 4.87$

	Sum of squares	Additional reduction/ mean square error for all variables
1960		
Regression with all the variables	3.4423	0.2248/0.088859 = 2.53
Regression with variables 3 and 4	3.2175	
1970		
Sum of squares for regression all variables	2.7466	0.1349/0.027018 = 4.99
Sum of squares for regression 3, 4, 7	2.61169	$F_{1.18;\ 0.95} = 4.41$ $F_{1.18;\ 0.99} = 8.29$

	Sum of squares	Additional reduction/ mean square error for all variables
Sum of squares for regression 3, 4, 7	2.61169	0.07189/0.0295817 = 2.43
Sum of squares for regression 3, 4	2.5398	$F_{1.21;\ 0.95} = 4.32$ $F_{1.21;\ 0.99} = 8.02$

It will be seen that, in formulating the regression on variables 3, 4 and 7, the additional increment including the remaining variables is not very significant.

The odd thing about the results is that variable 7 (coverage) appears with positive regression coefficients, although statistically it does not differ from 0 in all the regressions calculated, despite the fact that the zero-order coefficient of correlation between variable 1 (gross reproduction rate) and variable 7 is −0.5156. One possible explanation of this phenomenon is that because coverage is highly dependent upon the other variables, its direct effect, once those variables are fixed, is of no significance.

[*Tables 24–34 on pages 126–129; text continued on page 129.*]

TABLE 24. CORRELATION MATRIX, 1960

Variable No.	Variable No. 1	2	3	4	5	6
1	1	–0.6796	0.77464	–0.42247	0.60444	0.69952
2	–0.6796	1	–0.87005	0.20308	–0.9599	–0.93152
3	0.77464	–0.87005	1	–0.34611	0.84601	0.85841
4	–0.42247	0.20308	–0.34611	1	–0.20451	–0.19252
5	0.60444	–0.9599	0.84601	–0.20451	1	0.83634
6	0.69952	–0.93152	0.85841	–0.19252	0.83634	1

TABLE 25. ARRANGEMENT OF VARIABLES IN ORDER OF IMPORTANCE USING METHOD 1, 1960
(Forward stepwise inclusion)

Variable No.	Coefficient: Partial correlation		(Multiple correlation)	
3	r_{13}	0.77464	R^2_{13}	0.60006
4	$r_{14.3}$	–0.26015	$R^2_{1.34}$	0.62713
6	$r_{16.34}$	0.17306	$R^2_{1.346}$	0.63829
5	$r_{15.346}$	–0.19329	$R^2_{1.3465}$	0.65181
2	$r_{12.3465}$	–0.23434	$R^2_{1.34652}$	0.67093

TABLE 26. MULTIPLE REGRESSION OF VARIABLE 1 ON VARIABLES 3, 4, 6, 5 AND 2, AND ANALYSIS OF VARIANCE, 1960

Variable No.	Coefficient	Standard error	T value
3	0.038062	0.017379	2.1901
4	–0.017455	0.013649	–1.2788
6	–0.0021915	0.010711	–0.20461
5	–0.018568	0.013885	–1.3373
2	–0.019998	0.019032	–1.0507
Constant	4.4028		

Multiple correlation coefficient, $R = 0.8191$ ($R^2 = 0.67093$)

Analysis of variance

	Regression	Error	Total
Degrees of freedom	5	19	24
Sum of squares	3.4423	1.6883	5.1306
Mean square	0.68845	0.088859	
Standard error of estimate	0.29809		
F value	7.7477		

TABLE 27. MULTIPLE REGRESSION OF VARIABLE 1 ON VARIABLES 3 AND 4, AND ANALYSIS OF VARIANCE, 1960

Variable No.	Coefficient	Standard error	T value
3	0.039926	0.0077602	5.145
4	–0.016569	0.013111	–1.2637
Constant	2.3038		

Multiple correlation coefficient, R: 0.79191 ($R^2 = 0.62713$)

Analysis of variance

	Regression	Error	Total
Degrees of freedom	2	22	24
Sum of squares	3.2175	1.913	5.1306
Mean square	1.6088	0.086957	
Standard error of estimate	0.29488		
F value	18.501		

TABLE 28. OBSERVED AND ESTIMATED VALUES, 1960

Observed value	*Estimated value: 1 on 2, 3, 4, 5 and 6*	*Estimated value: 1 on 3 and 4*
2.34	2.40	2.42
2.29	2.44	2.36
2.95	2.64	2.48
2.89	2.93	2.82
2.59	2.54	2.63
2.05	2.09	2.11
2.10	2.09	2.06
2.97	2.62	2.72
3.23	3.12	3.14
3.29	2.89	3.01
2.89	2.86	2.88
2.59	2.83	2.96
3.05	3.00	2.98
2.91	2.94	3.07
2.80	2.49	2.52
3.38	3.59	3.38
2.94	3.18	3.18
3.39	3.20	3.30
2.48	3.16	3.20
3.22	3.08	2.99
2.80	2.85	2.84
3.01	2.96	2.87
2.69	2.77	2.61
3.25	2.85	2.85
1.49	2.05	2.17

TABLE 29. CORRELATION MATRIX, 1970

Variable No.	*Variable No.* 1	2	3	4	5	6	7
1	1	−0.60738	0.85917	−0.72419	0.62598	0.73518	−0.5156
2	−0.60738	1	−0.84183	0.46966	−0.97823	−0.88958	0.64685
3	0.85917	−0.84183	1	−0.65032	0.86473	0.87426	−0.67753
4	−0.72419	0.46966	−0.65032	1	−0.46161	−0.4986	0.57914
5	0.62598	−0.97823	0.86473	−0.46161	1	0.87256	−0.62015
6	0.73518	−0.88958	0.87426	−0.4986	0.87256	1	−0.64141
7	−0.5156	0.64685	−0.67753	0.57914	−0.62015	−0.64141	1

TABLE 30. ARRANGEMENT OF VARIABLES IN ORDER OF IMPORTANCE BY METHOD 1, 1970
(Forward stepwise inclusion)

Variable No.	*Coefficient*			
	Partial correlation		*(Multiple correlation)²*	
3	r_{13}	0.85917	R^2_{13}	0.73818
5	$r_{15.3}$	−0.45518	$R^2_{1.35}$	0.79243
4	$r_{14.35}$	−0.35569	$R^2_{1.354}$	0.81869
7	$r_{17.354}$	0.2819	$R^2_{1.3547}$	0.83309
6	$r_{16.3547}$	0.29089	$R^2_{1.35476}$	0.84722
2	$r_{12.35476}$	0.12412	$R^2_{1.354762}$	0.84957

TABLE 31. MULTIPLE REGRESSION OF VARIABLE 1 ON VARIABLES 3, 5, 4, 7, 6 AND 2, AND ANALYSIS OF VARIANCE, 1970

Variable No.	*Coefficient*	*Standard error*	*"T" value*
3	0.053521	0.015707	3.4075
5	-0.0043201	0.010554	-0.40933
4	-0.029465	0.013075	-2.2535
7	0.015053	0.011358	1.3253
6	0.0091066	0.0064961	1.4019
2	0.0053956	0.010167	0.53072
Constant	1.4505		

Multiple correlation coefficient, $R = 0.92172$ ($R^2 = 0.84957$)

Analysis of variance

	Regression	*Error*	*Total*
Degrees of freedom	6	18	24
Sum of squares	2.7466	0.48632	3.2329
Mean square	0.45776	0.027018	
Standard error of estimate	0.16437		
F value	16.943		

TABLE 32. MULTIPLE REGRESSION OF VARIABLE 1 ON VARIABLES 3, 7 AND 4, AND ANALYSIS OF VARIANCE, 1970

Variable No.	*Coefficient*	*Standard error*	*"T" value*
3	0.046522	0.0085726	5.4267
7	0.017666	0.011334	1.5587
4	-0.033363	0.01287	-2.5922
Constant	2.0588		

Multiple correlation coefficient, $R = 0.8988$ ($R^2 = 0.80785$)

Analysis of variance

	Regression	*Error*	*Total*
Degrees of freedom	3	21	24
Sum of squares	2.61169	0.621215	3.2329
Mean square	0.870563	0.0295817	
Standard error of estimate	0.171993		
F value	29.4291		

TABLE 33. OBSERVED VALUES AND ESTIMATED VALUES, 1970

Observed value	*Estimated value: 1 on 2, 3, 4, 5, 6 and 7*	*Estimated value: 1 on 3, 7 and 4*
2.02	1.83	1.84
1.94	1.87	1.85
2.13	2.12	2.01
2.21	2.28	2.20
1.86	1.99	2.08
1.55	1.67	1.67
1.58	1.63	1.57
2.12	2.08	2.18
2.40	2.36	2.49
2.19	2.27	2.35
2.12	2.27	2.24
2.08	2.44	2.44
2.44	2.19	2.30
2.58	2.60	2.57
2.06	1.99	1.84
2.86	2.83	2.69
2.59	2.55	2.53
2.88	2.84	2.81
2.36	2.43	2.41
2.34	2.46	2.48
2.15	2.19	2.19
2.40	2.26	2.27
2.00	1.90	1.96
2.62	2.30	2.26
1.45	1.59	1.67

TABLE 34. MULTIPLE REGRESSION OF VARIABLE 1 ON VARIABLES 3 AND 4, AND ANALYSIS OF VARIANCE, 1970

Variable	*Coefficient*	*Standard error*	*"T" value*
3	0.040029	0.0077323	5.1769
4	−0.028389	0.0128670	−2.2063
Constant	2.1716		

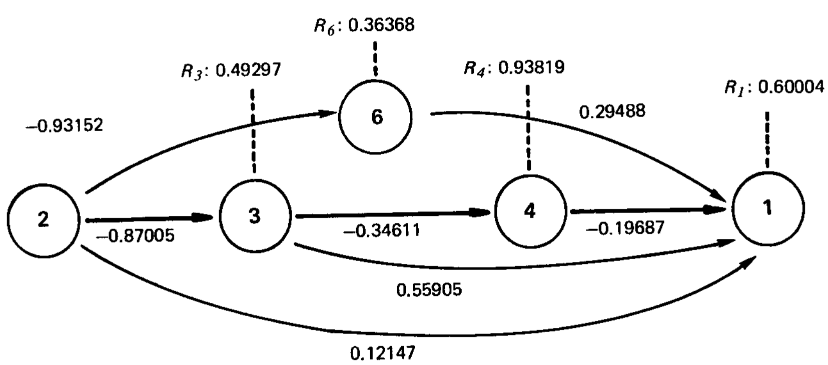
Multiple correlation coefficient, $R = 0.88635$ ($R^2 = 0.78562$)

Analysis of variance

	Regression	*Error*	*Total*
Degrees of freedom	2	22	24
Sum of squares	2.5398	0.69308	3.2329
Mean square	1.2699	0.031504	
Standard error of estimate	0.17749		
F value	40.31		

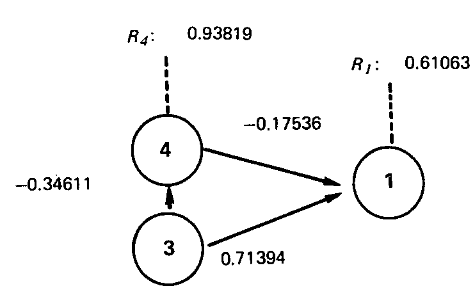
Lastly, a number of models were constructed for the relations among the variables, and the path analysis method was applied. In that analysis, variable 5, which was closely correlated with variable 2, was omitted.

The models for 1960 obviously do not include variable 7, coverage of the family planning programme.



Dependent variable	Independent variable 2	3	4	6	Multiple correlation coefficient R^2	Residue $\sqrt{1-R^2}$
1	0.12147	0.55905	−0.19687	0.29488	0.63996	0.60004
3	−0.87005	0	0	0	0.75699	0.49297
4	0	−0.34611	0	0	0.11980	0.93819
6	−0.93152	0	0	0	0.86774	0.36368

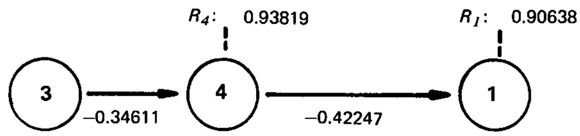
When the model was reduced to the relations of variable 1, 3 and 4, which were investigated by the multiple regression method, the following result was obtained:



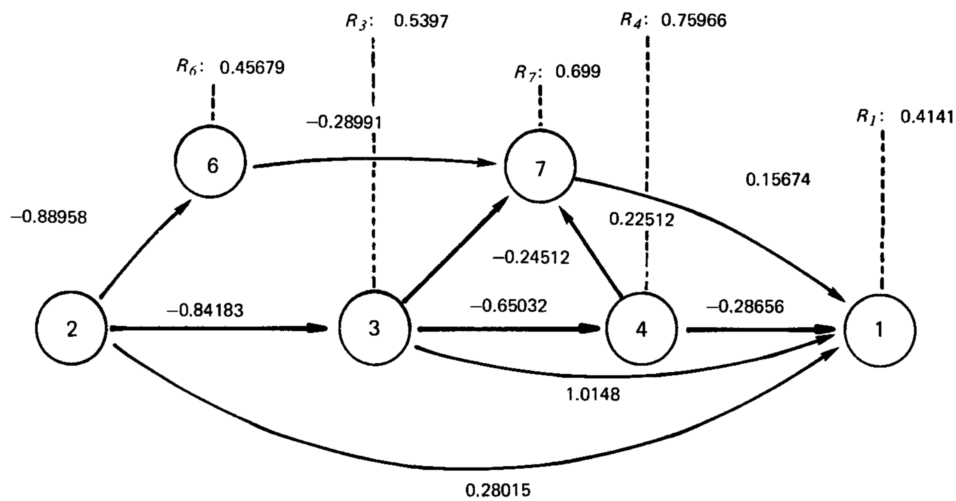
Dependent variable	Independent variable 3	Independent variable 4	Multiple correlation coefficient R^2	Residue $\sqrt{1-R^2}$
1	0.71394	−0.17536	0.62713	0.61063
4	−0.34611	0	0.11980	0.93819

It can be seen that, as in the analysis conduction by multiple regression, the residue of variable 1 does not vary significantly when the model is reduced to two independent variables.

In addition, another model with the same variables was proposed, in which there is a marked increase of the residue of variable 1.



The models constructed for 1970 are given below.



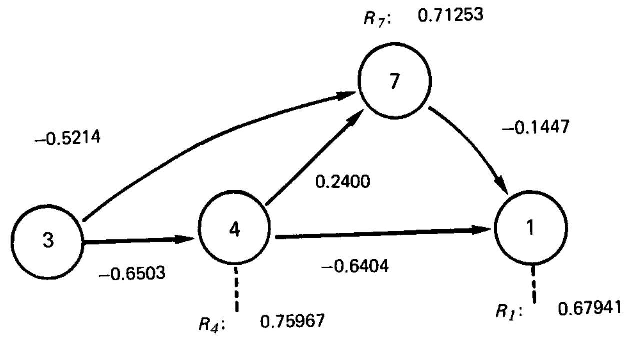
Dependent variable	Independent variable 2	3	4	6	7	Multiple correlation coefficient R^2	Residue $\sqrt{1-R^2}$
1	0.28015	1.01480	−0.28656	0	0.15674	0.82849	0.41410
3	−0.84183	0	0	0	0	0.70867	0.53970
4	0	−0.65032	0	0	0	0.42292	0.75966
6	−0.88958	0	0	0	0	0.79134	0.45679
7	0	−0.24512	0.27518	−0.28991	0	0.51139	0.69900

It can be seen that although the value of the residue for variable 1 suggests that the model formulated is a good one, the coefficients obtained are not what might be expected.

It is noteworthy, for example, that the coefficient between variable 2 (percentage of urban residents) and variable 1 (gross rate of reproduction), and the coefficient between variable 7 (percentage of coverage of the family planning programme) and variable 1 are both positive in sign. The direct relationships implied by these results are obviously inconsistent with reality.

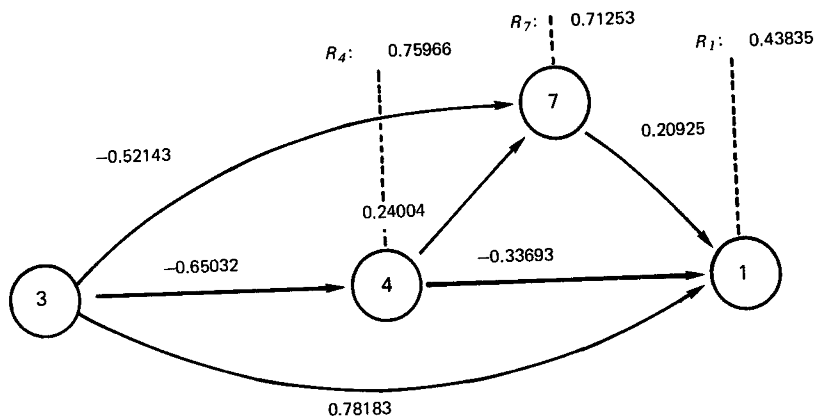
The reason for this result is, no doubt, that the model formulated in accordance with the present hypothesis of relations omits some direct effects of other variables on variable 1 and the intercorrelations of the variables considered.

Lastly, a model was formulated based on the results obtained in the arrangement of the variables in order of importance with the multiple regression analysis.



Although the residue for variable 1 is greater than in the first model for 1970, the signs of the coefficients agree with the signs expected from the relations.

The last model constructed was that given below. Except for the sign of the coefficient between variables 7 and 1, which one would have expected to be negative, this model appears to be fairly satisfactory.



Annex

BASIC DEMOGRAPHIC DATA FOR CHILE

TABLE 35. FEMALE POPULATION OF CHILDBEARING AGE, BY AGE GROUP, 1959-1974
(Thousands)

Year	Age group 12-14	15-19	20-24	25-29	30-34	35-39	40-44	45-49	12-49
1959	246.8	364.9	304.4	270.1	252.1	209.4	188.8	169.2	2 005.7
1960	253.1	376.8	311.3	270.8	260.4	212.4	191.1	172.2	2 048.1
1961	256.0	385.2	323.0	277.6	261.3	220.7	194.2	174.6	2 093.6
1962	258.9	395.6	334.6	284.4	262.2	229.1	197.2	177.1	2 139.1
1963	261.8	405.0	346.3	291.3	263.1	237.4	200.2	179.5	2 184.6
1964	264.7	414.4	358.0	298.1	264.0	245.8	203.3	181.9	2 230.2
1965	267.6	423.8	369.7	304.9	264.8	254.2	206.3	184.4	2 275.7
1966	279.3	430.6	379.1	316.6	271.7	255.2	214.7	187.5	2 334.7
1967	291.0	437.3	388.5	328.3	278.6	256.3	223.0	190.7	2 393.7
1968	302.6	444.1	398.0	340.0	285.5	257.3	231.4	193.8	2 452.7
1969	314.3	450.9	407.5	351.7	292.3	258.4	239.7	197.0	2 511.8
1970	326.0	457.6	416.9	363.4	299.2	259.4	248.1	200.2	2 570.8
1971	334.1	477.4	423.8	372.8	310.9	266.3	249.2	208.3	2 642.8
1972	342.2	497.1	430.7	382.4	322.5	273.1	250.3	216.5	2 714.8
1973	350.3	516.8	437.5	391.9	334.2	280.0	251.4	224.7	2 786.8
1974	358.4	536.6	444.4	401.4	345.8	286.8	252.5	232.9	2 858.8

Source: Chile, Oficina de Planificación Nacional (ODEPLAN), *Proyección de la población de Chile, por sexo y grupos quinquenales de edad, 1958-2000.*

TABLE 36. AGE-SPECIFIC FERTILITY RATES, CORRECTED, 1960-1974
(*Rates per 1 000 women*)

Year	12-14	15-19	20-24	25-29	30-34	35-39	40-44	45-49
	Age group							
1956	2.5	81.3	232.5	269.4	207.5	147.6	70.3	16.5
1957	2.2	83.0	237.6	280.9	217.1	155.0	71.5	16.8
1958	2.1	81.0	231.7	276.0	218.2	156.2	71.1	16.3
1959	2.6	79.6	228.7	269.5	223.5	155.2	69.0	14.3
1960	2.7	79.6	235.2	273.3	229.3	160.1	67.3	15.5
1961	2.9	82.1	230.7	275.6	241.0	158.1	64.4	13.6
1962	2.6	81.5	233.4	268.6	245.9	155.0	68.4	13.1
1963	3.0	83.8	241.3	253.5	234.7	155.2	68.5	11.8
1964	2.9	81.9	235.4	247.5	220.1	150.4	65.5	11.1
1965	2.9	82.8	227.9	242.7	207.4	149.5	64.5	11.4
1966	2.8	82.1	220.2	230.5	186.3	142.4	60.3	10.4
1967	2.9	82.5	215.0	213.2	168.8	130.4	55.3	9.9
1968	2.5	79.2	206.0	200.0	152.1	116.0	52.9	8.9
1969	2.7	78.9	194.8	186.1	142.6	105.2	48.0	8.4
1970	2.3	80.7	192.7	182.5	135.0	94.4	44.7	8.4
1971	2.8	84.3	200.1	186.0	136.2	88.7	42.4	7.9
1972	3.1	86.4	203.0	183.6	131.0	82.3	40.0	7.3
1973	2.8	83.5	202.5	175.5	123.9	76.4	35.4	6.3[a]
1974	3.1	79.3	191.3	169.4	116.9	73.4	33.5	6.0

Sources: Births per age group were calculated by applying the structure of recorded births to the total of births estimated by H. Gutiérrez, *La integridad del registro de nacidos vivos en Chile, 1953-1966;* and since 1967, from unpublished data. For 1956-1970, the information used was the birth data published by Instituto Nacional de Estadísticas in its journal *Demografía;* and for 1971-1974, those published by Servicio Nacional de Salud in its journal *Nacimientos.* The number of women of childbearing age was taken from data for 1960, 1965, 1970 and 1975 given in Chile, Oficina de Planificación Nacional, *Proyección de la población de Chile por sexo y grupos quinquenales de edad 1958-2000* (Santiago).

[a] The initial value obtained was 15.3, which is inconsistent with the trend of this specific rate: it indicates a defect in the basic information. The final value of 6.3 was obtained by projecting the ratio (5f45/5f40).

TABLE 37. ADMISSIONS PER ANNUM TABULATED ACCORDING TO INTRA-UTERINE DEVICE AND ORAL GESTAGENS, 1964-1974

Year	*Intra-uterine device*		*Oral gestagens*		*Total*	
	Number	*Percentage*	*Number*	*Percentage*	*Number*	*Percentage*
1964	11 264	96.0	471	4.0	11 735	100.0
1965	20 467	69.3	9 056	30.7	29 523	100.0
1966	33 686	67.1	16 515	32.9	50 201	100.0
1967	45 375	77.1	13 477	22.9	58 852	100.0
1968	46 422	77.1	13 788	22.9	60 210	100.0
1969	40 674	68.1	19 070	31.9	59 744	100.0
1970	43 602	71.3	17 528	28.7	61 130	100.0
1971	36 142	63.4	20 835	36.6	56 977	100.0
1972	35 167	60.1	23 331	39.9	58 498	100.0
1973	80 155	66.5	40 399	33.5	120 554	100.0
1974	121 879	66.1	62 450	33.9	184 329	100.0

Sources: For 1964-1966, M. Requena and T. Monreal, "Evaluación del programa de control del aborto inducido y planificación familiar en Chile", *Milbank Memorial Fund Quarterly,* vol. XLVI, No. 3 (July 1968), part 2, pp. 213-246; for 1969-1973, Z. Soto, *América Latina: situación de los programas de planificación de la familia hasta 1973,* CELADE Series A, No. 130 (Santiago, 1975); for 1974, Asociación Chilena de Protección de la Familia (APROFA).

Note: Figures for 1967 and 1968 are estimated. In the report by Soto, the totals are: for 1967, a total of 16,976, with no distinction as to method; and for 1968, a total of 102,086, grouped into intra-uterine devices (IUD) and oral gestagens. Since these figures are not in satisfactory agreement with the general trend, it was assumed that some of the admissions listed for 1968 were actually admissions for 1967. The number of admissions for the two years was therefore prorated in the same ratio as existed between the years 1969 and 1970. The resulting totals were distributed under the headings of admissions with IUD and admissions with gestagens, in accordance with the respective percentages recorded for the year 1968 in the document referred to, as no such information was available for 1967.

TABLE 38. PERCENTAGE DISTRIBUTION OF ACCEPTORS OF THE INTRA-UTERINE DEVICE AND ORAL GESTAGENS, BY AGE GROUP

Age group	*IUD*	*Oral gestagens*
15-19	4	8
20-24	30	34
25-29	29	36
30-34	21	15
35-39	12	4
40 and over	4	3
	100	100

Note: These percentages correspond to the average percentages of admissions for the periods January 1968-August 1971 and September 1971-June 1972 at the National Employees' Medical Service, Valparaíso, rounded off to make a total of 100 per cent. Centro Latinoamericano de Demografía, Series A, No. 115, November 1972, annex 2.

TABLE 39. CUMULATIVE RATES OF CONTINUATION BY METHOD, ACCORDING TO DURATION OF USE
(*Rates per 100*)

Duration of use (months)	*Cumulative rates* Intra-uterine device	Gestagens
1	98.84	98.26
2	97.04	94.61
3	95.26	91.35
4	93.20	87.50
5	91.35	83.76
6	90.10	81.35
7	88.29	78.06
8	87.08	75.45
9	85.66	72.11
10	83.93	69.21
11	82.72	66.56
12	81.16	64.06
13	79.72	60.42
14	78.61	57.24
15	76.91	54.45
16	76.35	52.23
17	75.51	49.70
18	74.75	47.23
19	74.34	44.49
20	73.06	41.77
21	72.02	39.40
22	71.06	36.13
23	69.89	33.79
24	68.49	31.56
30	62.12	20.12
36	59.06	13.12
42	54.99	8.36
48	48.85	2.79
54	42.74	2.79

Source: Chile, National Employees' Medical Service (SERMENA).

TABLE 40. HOSPITALIZATION FOR ABORTION AT THE NATIONAL HEALTH SERVICE AND RECORDED LIVE BIRTHS, 1937-1964

Year	*Abortions*	*Live births*	*Abortions per 100 births*
1937	12 963	153 354	8.4
1938	13 982	154 927	9.0
1939	14 730	163 589	9.0
1940	16 254	166 593	9.7
1941	18 265	165 004	11.0
1942	19 342	170 222	11.3
1943	20 009	172 095	11.6
1944	19 449	174 864	11.1
1945	21 581	178 292	12.1
1946	23 619	175 686	13.4
1947	24 535	186 784	13.1
1948	26 448	189 236	13.9
1949	28 514	189 719	15.0
1950	29 512	188 323	15.6
1951	30 571	191 332	15.9

TABLE 40 (*continued*)

Year	*Abortions*	*Live births*	*Abortions per 100 births*
1952	32 862	195 470	16.5
1953	33 862	211 808	15.9
1954	35 748	209 920	17.0
1955	39 340	225 352	17.4
1956	41 429	237 268	17.0
1957	44 945	262 746	17.1
1958	49 041	262 759	18.6
1959	49 448	249 799	19.8
1960	47 096	256 674	18.3
1961	49 195	263 985	18.6
1962	51 246	275 960	18.6
1963	49 772	280 167	14.9
1964	56 391	277 893	20.3

Source: Francisco Mardones Restat, Jorge Rosselot Vicuña and Lucía Lopez Cazenave, *Política y programa de regulación de la natalidad en el Servicio National de Salud de Chile* (Santiago, Servicio Nacional de Salud, 1967).

Part Two

REPORT OF THE EXPERT GROUP MEETING ON METHODS OF MEASURING THE IMPACT OF FAMILY PLANNING PROGRAMMES ON FERTILITY

INTRODUCTION

A. Background and purpose

The Expert Group Meeting on Methods of Measuring the Impact of Family Planning Programmes on Fertility was convened at Geneva, Switzerland, from 20 to 27 April 1976. The meeting was organized by the Population Division of the Department of Economic and Social Affairs, United Nations Secretariat, with the collaboration of the Committee on Demographic Aspects of Family Planning Programmes of the International Union for the Scientific Study of Population (IUSSP) and the financial support of the United Nations Fund for Population Activities (UNFPA).

The following persons served as members of the Expert Group: Albert I. Hermalin (United States of America), Chairman; Erica Taucher (Chile), Vice-Chairman; Jeanne Clare Ridley (United States of America), Rapporteur; William Brass (United Kingdom); S. A. Brogger (World Health Organization); C. Chandrasekaran (World Bank); Chen-tung Chang (Singapore); G. Edward Ebanks (Canada); Halvor Gille (United Nations Fund for Population Activities); Yolande Jemai (Tunisia); Suk Kap Koh (Republic of Korea); W. Parker Mauldin (United States of America); Robert G. Potter (United States of America); Ismail Sirageldin (United States of America); K. Srinivasan (India); Kalman Tekse (World Health Organization); H. Bradley Wells (United States of America).

Members of the Expert Group served in their individual capacity as scholars having special knowledge of and research experience with methodology of evaluating the demographic impact of family planning programmes on fertility and not as representatives of a Government or agency.

Léon Tabah, Director of the Population Division, represented the Secretary-General of the United Nations at the meeting. Gwendolyn Johnson Acsádi, Chief, Fertility and Family Planning Studies Section, Population Division, served as Technical Secretary of the meeting. Maurice Szykman, Population Affairs Officer, Fertility and Family Planning Studies Section, was Assistant Technical Secretary.

In convening this meeting, the Secretary-General acted pursuant to a recommendation by the Population Commission at its fifteenth session[1] and its eighteenth session[2] and endorsed by the Economic and Social Council in its resolution 1946 (LVII),[3] which supports the biennial (1976–1977) work programme of the Secretariat.[4] The Population Commission had considered that, in light of the growing numbers of Government-sponsored or -supported family planning programmes intended to regulate fertility, the Secretary-General should convene an expert group to consider whether the technology was adequate to determine whether and to what degree the family planning programmes were fulfilling their purpose.

The Expert Group was requested to review problems and issues in the application of methods and, in particular, in the interpretation of results obtained when various methods were utilized to evaluate the impact of family planning programmes on fertility. The objective was to establish guidelines for determining which measurement method or group of methods might be most appropriate in specific circumstances and to ascertain, in so far as possible, the relative validity of the results obtained by each method.

To facilitate the discussions, the United Nations commissioned three country case studies in which as many as possible of the methods used to evaluate programme impact on fertility were applied to the same set of data, with the view to analyzing the issues and problems involved and comparing the results obtained. Those studies were developed by national experts in Chile (ESA/P/AC.7/4); Karnataka State, India (ESA/P/AC.7/2); and Tunisia (ESA/P/AC.7/3). In addition, the United Nations Secretariat provided a background paper (ESA/P/AC.7/1) in which aspects of the issues and problems were set forth.[5]

To enrich the discussions further, each member of the Expert Group was requested to prepare a brief statement on one or more of the following topics: (*a*) problems of interpreting the validity of evaluation results; (*b*) problems of determining appropriate methods of evaluation; (*c*) needed data and research in these fields.[6]

[1] *Official Records of the Economic and Social Council, Forty-eighth Session, Supplement No. 3*, paras. 154–155.

[2] *Official Records of the Economic and Social Council, Fifty-eighth Session, Supplement No. 6*, para. 139.

[3] *Official Records of the Economic and Social Council, Fifty-eighth Session, Resolutions and Decisions, Supplement No. 1*, p. 10.

[4] "Biennial work programme (1976–1977), medium-term plan (1976–1979) and long-term perspectives", note by the Secretary-General (E/CN.9/317), paras. 36 and 72.

[5] These documents, which are listed in annex II, comprise part one of the present publication.

[6] The statements, which were originally presented at the meeting as conference room papers, make up part three of this publication. See annex II for list of papers.

The Expert Group considered the following substantive items:

- Methods of measuring the impact of family planning programmes on fertility;
- Major problems encountered in the case studies of methods of measuring the impact of family planning programmes on fertility;
- The major methodological issues;
- Comparison of measures: problems in interpreting validity of evaluation results;
- Problems in determining the appropriate method(s) of evaluation;
- Needed data and research;
- Recommendations.

B. Organization of the work

The Director of the Population Division, opening the meeting on behalf of the Secretary-General, welcomed the experts and thanked them for their assistance to the United Nations work programme in the field of population. He also expressed appreciation to the Committee on Demographic Aspects of Family Planning Programmes of IUSSP for its collaboration in the project. He stated the purpose of the meeting and outlined the essential issues and problems relating to evaluation of the impact of family planning programmes on fertility which required the attention of the Expert Group.

At the invitation of the representative of the Secretary-General, C. Chandrasekaran, Chairman of the Committee on Demographic Aspects of Family Planning Programmes of IUSSP, addressed the meeting. He noted the timeliness of the meeting and observed that the experts had before them an opportunity as a group to find means of dealing with many difficult and critical problems which had confronted them as individual researchers.

The members of the Committee on Demographic Aspects of Family Planning Programmes of IUSSP, which collaborated in the Expert Group meeting, are: C. Chandrasekaran, Chairman; Albert I. Hermalin, Secretary; Robert G. Potter; Ismail Sirageldin; K. Srinivasan; and Gwendolyn Johnson Acsádi (Associate).

Some members of the Expert Group, including members of the IUSSP Committee, had been invited to lead discussions of substantive items of the agenda. Those assignments are listed in annex I.

The Expert Group focused on the major methodological and related issues in the application of the principal methods currently used to evaluate the effect of family planning programmes on fertility and on the advantages and disadvantages of each of the methods. Those methods were: (*a*) standardization approach; (*b*) trend analysis; (*c*) couple-years of protection (CYP); (*d*) component projection; (*e*) experimental designs; (*f*) analysis of the reproductive process; (*g*) regression analysis (including path analysis); and (*h*) simulation models. In its discussions, the Expert Group considered the problems of applying those methodologies and of interpreting the results as set forth in detail in the background papers prepared for the meeting and the various aspects of those problems as outlined in the statements prepared by members of the Expert Group.

The Expert Group agreed that a number of the questions which related to the impact of the family planning programme on fertility were not appropriate to its discussions. For instance, it was recognized that alternative designs of family planning programmes were likely to have different effects on fertility, but the question which alternatives were likely to produce the largest impact was not among the concerns of the meeting. In some countries, family planning programmes were defined rather broadly, so that efforts to raise the age of marriage, for example, might be included among programme activities. However, the Expert Group decided that, in its discussions, family planning programmes would have the more usual, limited meaning, that of the delivery of contraceptive services and other fertility control methods by either governmental or private agencies.

It was also recognized that the impact of a programme could be considered from several perspectives: whether the programme was reaching and serving the persons it had been designed to serve; whether the programme was able to retain its clients as evidenced by increasingly long contraceptive continuation rates. But those questions also were not of immediate relevance to evaluation of the impact of a family planning programme on fertility and therefore, were outside the scope of the meeting.

I. THE METHODS

At the outset of the discussion of the eight methods commonly applied in the evaluation of the effect of family planning programmes on fertility, it was made clear that evaluation could usefully be undertaken only in light of what precisely was being measured and the time to which the measurement related.

To specify what was being measured was not a simple matter, for family planning programmes generated both direct and indirect effects; and those effects were not easily differentiated, nor was there a satisfactory method of separating the effects of a family planning programme from those of socio-economic development. A direct programme effect would be births delayed or forgone by a group of women who had become acceptors of a contraceptive method in a programme. An indirect programme effect would be births delayed or forgone by women who had adopted some contraceptive method obtained from a source other than the programme, but who had been influenced by the programme to do so. An effect of socio-economic development might be a reduction of fertility traceable to increases in the employment of women outside the home in non-familial activities or to improvements in the health and nutrition of women.

The relevance of time involved two perspectives: how long a programme must run before the measurement of impact was feasible; and whether an observed impact was of short- or long-term significance. Long-run impact was viewed as an effect on the completed fertility of women currently in the reproductive ages. Short-term impact might represent changes in period rates, due to spacing of births, with or without change in completed family size.

As stated above, the principal focus was on the methodological issues and the problems encountered in the application of the methods, along with certain advantages and disadvantages of the specific methods.

A. Standardization approach

In the standardization approach, an attempt was made to answer the question whether there had been a real change in the fertility of the population, by determining changes that were due to shifts in age structure and marital status. Use of that method indicated the proportion of change attributable to structural factors (e.g., age structure and marital status), and the proportion attributable to fertility behaviour. The method did not, however, separate the effects of a family planning programme from those of socio-economic factors, due to the interaction of the structural factors with the development variables.

The standardization approach was an essential first step in the measurement of the effect on fertility of family planning programmes, and it should be routinely applied in all evaluations. It might also be used to study factors that had affected trends in fertility prior to the introduction of a family planning programme. However, because the method produced only a residual after variables had been standardized for, it could not be used to determine how much of the change in fertility had resulted from the effects of the family programme and how much from socio-economic changes. Further, if the residual was, in fact, taken as a measure of the impact of the family planning programme, an over-estimate of impact would be likely. The results of the application of the standardization procedure in the case studies presented as background papers for the meeting indicated that in no case could all of the observed decline in fertility have been accounted for by the variables included in the standardizations, but that the decrease could be attributed to the effects of the family planning programme as well as to factors that had not been taken into account. In summary, standardization could only indicate that the programme might have had some impact.

The choice of variables to be controlled by standardization was a crucial step in application of the method. Age and marital status were considered essential, and rural-urban residence and socio-economic status indicators added refinements that provided invaluable information to programme administrators. However, the choice of variables was in part a theoretical problem, in that the determinants of fertility must be identified. Consequently, the variables must be selected according to hypotheses concerning their relation to fertility. It was noted that the statements by Ebanks[1] and Wells[2] deal directly with that problem.

One of the greatest problems encountered in using the standardization approach was the availability of data. In each of the three country case studies provided for the Expert Group meeting, standardization was carried out for age and marital status. The case study for Karnataka State, India, also standardized for urban-rural residence, while the Chilean study included area of residence, educational level and female labour force participation. The overriding role of availability of data as a determinant of evaluation procedures was seen in the fact that, in each study, the

[1] G. Edward Ebanks, "Needed data and research on the impact of family planning programmes on fertility" (Conference room paper 8).

[2] H. Bradley Wells, "Notes on causal relationships in measuring fertility change" (Conference room paper 9).

inclusion or exclusion of variables depended upon what statistics were available.

Data availability might prove a crucial factor if it were desired to undertake a more sophisticated application of standardization. For instance, if the intent was to standardize several factors simultaneously, the unavailability of required data could be a major hindrance. Interaction between various factors might be viewed as a problem. For example, the unobserved variables might influence those employed in the standardization, while, at the same time, the variables used in the standardization might interact with one another; and it was exceedingly difficult to determine their separate effects. In either case, it would be difficult to ferret out the amount of fertility change attributable, for example, only to changes in proportions married, if women had also advanced in education or had engaged increasingly in economic activities. For, improvements in education had tended to raise age at marriage and, consequently, to lower the average duration of marriages, while equipping women for non-familial activities.

B. Trend analysis

The Expert Group considered that trend analysis, the fertility projection technique, permitted some indication as to what might have happened without a family planning programme or in the absence of social, economic and demographic changes. Moreover, with that method account could be taken of changes in fertility levels which had occurred prior to the initiation of a family planning programme. For instance, it was possible to extrapolate into the future declines in fertility observed before a programme was begun. Consequently, the method indicated whether a change had taken place beyond what would be expected on the basis of past experience. But the results derived by the projection approach rested upon a very gross assumption, namely, that future fertility would be equivalent to what would have happened if past trends of fertility had continued. The impact of the programme was seen as the difference between projected fertility and recorded fertility.

The method required data on past trends in fertility and good observations on the trends in fertility following inauguration of a programme. But the results derived from it were even more uncertain than those of the standardization technique, which at least permitted the identification of some factors that had contributed to an observed change. Mention was made of the fact that there were a number of projection techniques, each of which differed somewhat in respect of the data required for application. For example, the statistics needed for the projection technique could be quite formidable if the number of components employed was large, as the requirements might include data on sex, age, marital status, residence, education and other characteristics over a considerable period. But owing to the great variety of problems encountered and the means of coping with them that were applied in each instance, the results could not be viewed as leading to any firm conclusions as to the efficaciousness of the method for gauging the impact of a programme.

C. Experimental design

The extent to which truly classical experimental design could be carried out was considered by the Expert Group to be the major issue involved in the application of that method. The opinion was that truly classical experimental design could be approached, but rarely achieved in family planning evaluation. However, that method was a powerful scientific tool, setting a standard against which all other methods must be judged.

In classical experimental design, the researcher selected randomly a member of the control group and one of the experimental group. But in evaluation, it was all but impossible randomly to allocate individuals or areas to experimental and control groups, and the problem would be different for areas and individuals. A drawback to application of experimental design was that it was implemented over a considerable period of time and thus was not suited to be a measure of programme impact after short intervals.

The difficulties of applying the method were considered virtually insurmountable, but there were various ways in which the classical experimental design might be approached. Matching studies were one means of affecting the more rigorous experimental design method. One variation would be to use individuals or areas matched on a number of basic characteristics. Retrospective data for the matching of the groups would be collected and the groups established at some point prior to a survey. Follow-up surveys would then produce the prospective data whereby changes could be assessed.

More generally, the full potential of surveys had not been realized in terms of analyzing the data with a view to applying the experimental design as the model. Further, the use of stratified sampling in survey design could assist greatly in achieving the basis for more rigorous analysis. Regression techniques also were important analytical tools that would permit an approach to the experimental design by taking advantage of the fact that family planning programmes were usually begun in phases, e.g., only some parts of a country were included in a programme when it was initiated and the programme was gradually extended to other parts of the country. If areas of a country could be randomly assigned prior to the time that the family planning programme was initiated, using as criteria the projected time when the programme would be phased in, fairly good measures of the impact of a programme after the elapse of a given time period could be obtained. Thus, something approaching a classical experimental design could be achieved.

An indication of the difficulties of carrying out experimentally designed studies was provided by the applications of the method in case studies of India and

Tunisia, which approached the method by utilizing available survey data. Both applications were considered good examples of what could be achieved with survey data. The utilization of the method in the Indian case study might be regarded as an example of the use of careful matching on a number of important reproductive characteristics. Both applications illustrated the problems of interpreting the results.

D. Couple-years of protection

The central issues in regard to couple-years of protection (CYP) was what it purported to measure. It was originally conceptualized as a prevalence index but became widely used, even by those who had developed it, as an index of births averted and, thus, as a method of assessing family planning programme impact. It was the consensus of the Expert Group that if that method was used for the assessment of the impact of a family planning programme on fertility, the formula should be modified to give a more precise measure both of couple-years of protection and of births averted.

Among the advantages of the CYP method was its relative simplicity in terms both of data requirements and of calculation. If data were available, many refinements might be introduced, such as a CYP measure specific for age, residence and education.

Nevertheless, since such refinements would require additional data, it was felt that a major drawback existed. If the measure was to be refined in a meaningful way, fairly detailed data on continuation rates (derived from a follow-up of acceptors) would be needed, and such data were not usually available. Further, the calculations were made, as a rule, for a calendar year, and realistic assumptions regarding continuation rates of contraceptive use on a calendar-year basis were particularly difficult to obtain.

Although, as noted above, the CYP method was capable of refinement, the additional calculations would require statistics cross-classifying women by fertility and by acceptance of specific contraceptive methods according to age, residence, marital status etc. Those data were not usually available, and without such refinements, interpretation of results of the method was very difficult. Further, lacking refinements based on reliable data, a number of gross assumptions must be made. Owing to the rough nature of the method, it might yield unreliable estimates of programme impact.

Each of the case studies applied the CYP method. A fairly detailed appraisal of the method was given in the Tunisian study. In the Indian case study, the method was calculated on only one set of assumptions regarding continuation rates although two different population distributions were used. But it was considered noteworthy that similar results had been obtained from applications of the component projection approach and the CYP method.

E. Component projection approach

The major issues raised concerning the component projection approach appeared to be its sensitivity and robustness. It was felt to be conceptually sound, workable and easy to apply, and was considered to hold the best potential for single-method evaluation.

(*b*) Mention was made of the following advantages: (*a*) age was taken into account; (*b*) the estimation of births averted for acceptors was possible; (*c*) it could be used for target-setting by providing estimates of the number of acceptors needed to achieve a specified reduction in the number of births; (*d*) it permitted the disaggregation of the population into components and, consequently, estimates might be obtained for important subgroups of the population (e.g., rural-urban, educational groups); (*e*) it also could be used to estimate the impact of a programme in the future under varying assumptions.

But the component projection approach also posed significant disadvantages, including the necessity of estimating potential fertility. The method also required continuation rates and, hence, a good follow-up of acceptors.

Although the method was thought to be conceptually good, many of the applications had merely been simplified versions of it, owing mainly to the various simplifying assumptions required in the absence of concrete data. Ideally, programme administrators should plan towards the use of the method by collecting the basic data at the time the programme was begun and periodically thereafter, for it was the lack of such bench-mark data that necessitated resort to a number of arbitrary assumptions.

The Expert Group observed that the statement by Potter,[3] which summarized the recent developments in the use of the method, pointed out that the component projection approach fulfilled a variety of evaluation needs.

The difficulties of obtaining the requisite data for that method were emphasized in each of the case studies. Estimation of potential fertility was also considered a major problem in those studies. For example, it had become clearly evident from those studies that because no clear guidelines existed as to what data or group of women the estimate should be based upon, the quality of evaluation depended on the skill and judgement of the evaluator; there were no standards for assessing those qualities.

F. Analysis of the reproductive process

Analysis of the reproductive process had not been widely used, owing mainly to the many detailed refinements incorporated into the approach. And it was considered that those refinements constituted the principal

[3] Robert G. Potter, "Component projection *versus* other techniques for assessing programme achievement towards a targeted fertility reduction" (Conference room paper 6).

issue with respect to the method. Questions were raised concerning what the conditions were that made the extra refinements incorporated into that approach really essential and when those additional adjustments were expendable. The refinements consisted of subtractions from woman-years of protection that allowed for secondary sterility, overlap of contraceptive practice with amenorrhoea and failures of contraception causing accidental pregnancies. In addition, alternative estimates of potential fertility were provided in order to represent a range of substitution effects.

The method provided a highly accurate estimate of women-years of protection and of potential fertility of acceptors, which were its major strengths; and it had proved a fruitful research tool in that it focused research on a number of significant factors affecting evaluation.

The Expert Group was of the opinion that a basic disadvantage of the approach was its extensive data requirements. Also, the assumptions for one population with respect to secondary sterility, post-partum amenorrhoea and time to conception in the absence of contraception, derived in Potter's application,[4] were not necessarily valid for other populations. Another problem was that, in the Potter application, the approach was designed to estimate births averted per insertion of an intra-uterine device (IUD). Some modification was necessary for application in respect of other contraceptive methods.

Analysis of the reproductive process as a method of evaluation, as formulated by Wolfers,[5] had been designed for evaluation of a post-partum family planning programme. Its application to other types of family planning programmes would necessarily require modification.

Theoretically, data permitting, the procedures of the analysis of the reproductive process approach could be incorporated into the component projection approach in order to measure family planning impact in a period population, though that merger had not yet been formally carried out.

Significantly, analysis of the reproductive process was the only method not applied in any of the three case studies, and the reason given in each case was unavailability of the requisite data. In light of the considerable problems of applying it, as it was currently formulated, that highly specialized method had limited applicability to general evaluation efforts and research was needed to facilitate its wider application.

G. Regression analysis

The Expert Group felt that regression analysis depended upon a conceptual model, and, therefore, some of the more important issues in regard to its application related to the variables that should be included. The method posed a number of statistical and theoretical problems, none of which could be easily resolved. Regression analysis permitted estimation of the relative impact of the programme, and it could be used with both macro and micro data, although in evaluation studies macro applications were the more common.

Among its disadvantages, regression analysis required adequate measures not only of fertility but of non-programme and programme variables. Because the choice of variables to be included was so crucial, a fairly detailed conceptualization was required. Most regression analyses dealt only with different moments in time, but the method could also be applied to time series data and thus could take account of time lags.

The attempted application of the method to Chilean data in the case study prepared for background material was cited as a good example of the difficulties encountered in applying it. In practice, there was not much choice in terms of variables to be included; the difficulty arose with respect to availability of data on those variables.

Additional problems confronted the researcher when the method was applied to areal units; a range of areal units was required, with variability among the units, for the method would break down if there was a lack of variance.

It was mentioned that in applying regression analysis, variation within the units must be accounted for, which was rarely done. The fact that the method required a fairly high level of statistical sophistication could lead to its misapplication.

H. Simulation models

The Expert Group felt that some question existed about the usefulness of simulation models for estimating the impact of family planning programmes on fertility, and the cost was considered to be high in relation to that of other methods. Simulation had an important role in research on evaluation, for simulation models could be useful in the study of factors that were not readily observable. Some of the basic biological factors were not amenable to direct observation and, consequently, simulation could aid in the study of such variables as age at sterility. One of its important uses had been—and was likely to continue to be—to indicate the role of chance. Simulation could also be used to estimate potential fertility, and the method could aid in testing estimating procedures under different condi-

---

[4] For a description of the method, see Robert G. Potter, A technical appendix on procedures used in manuscript "Estimating births averted in a family planning program", paper prepared for Major Ceremony V University of Michigan Sesquicentennial Celebration, 1 June 1967. See also Robert G. Potter, "Application of life-table techniques to measurement of contraceptive effectiveness", *Demography*, vol. 3, No. 2 (1966), pp. 297–304; and *idem*, "Estimatimg births averted in a family planning program", in S. J. Behrman, Leslie Corsa, Jr., and Ronald Freedman, eds., *Fertility and Family Planning: A World View* (Ann Arbor, Mich., University of Michigan Press, 1969), pp. 413–434.

[5] David Wolfers, "The demographic effect of a contraceptive programme", *Population Studies*, vol. XXIII, No. 1 (March 1969), pp. 111–141.

tions. It could be a powerful methodological tool for validating different methods of measuring the impact of programmes on fertility, and it should also be useful for studying such topics as measurement error, response error and sample size. Simulation models permitted the carrying-out of experiments and might in that way provide insights into the implications of various strategies in family planning programmes. Clearly, simulation was primarily a research tool, and it should be treated as such.

In the opinion of the Expert Group, a disadvantage of the models was that they required a great deal of data, some of which were practically non-existent. Moreover, many of the more sophisticated models (i.e., micro models, using the Monte Carlo method) required large-scale computers which, as a rule, were relatively expensive to utilize. In addition, the development of such models was highly expensive, and, for that and other reasons, the models currently available should be more fully exploited before any new models were developed. Because the structure of many of the models was relatively complex, potential users must devote a great deal of time to understanding the definitions or assumptions built into them. It was pointed out that not only was the development of simulation models costly, but the costs of running experiments with them were high.

Only one of the case studies, that of Karnataka State, India, included an application of a simulation model. That model, it was noted, had already been developed and thus was available for use in the case study. The results obtained with the model indicated the possibility that the CYP method over-estimated the number of births averted.

II. THE MAJOR METHODOLOGICAL ISSUES

Application of the various methods of measuring the effects of family planning programmes on fertility involved a number of critical methodological issues. Although it was not possible for the Expert Group to resolve those issues, it dealt with a variety of them and emphasized that awareness of the issues by those who attempted to use the various methods was essential if the limitations of the methods were to be fully appreciated. Moreover, an understanding of these methodological issues should lead to a better assessment of the results obtained with the different methods.

A. Data requirement problems

The Expert Group considered problems associated with the selection, use and interpretation of data to be among the more difficult of those attending measurement of the impact of family planning programmes on fertility. The principal difficulties related to the types of data needed, their sources and quality. The type of data needed depended upon the particular method to be applied and, in that respect, some methods were more demanding than others. Essentially, statistics on the composition of a population, fertility data of reliable quality and quantitative information on the scope of the programme were indispensable. Certain methods also required data on mortality and on biological factors. Frequently, a method that might be considered the most appropriate could not be applied because of lack of relevant statistics.

The data were drawn from four principal sources: population censuses; vital statistics registration systems; demographic or other special surveys; and family planning programmes (service statistics).

The opinion was that there was no substitute for reliable data based on population censuses. Census data, along with post-enumeration surveys and special surveys, could provide the essential data on the demographic, economic and social characteristics of a population and on various other non-programme factors affecting fertility, such as health and nutrition. The Expert Group recommended that post-census surveys, in which more detailed social and economic statistics were collected along with data on fertility, should be carried out as a standard procedure. A substantial sample enumeration at the time of the census would also provide the requisite statistics on a number of factors not associated with the family planning programme.

It was advantageous to collect the required statistics through sample surveys, not only because the surveys could be designed to meet specific data needs but because they could provide fertility data specific for other variables considered important. Surveys could also produce data on characteristics of acceptors, and they could be conducted more often than censuses, which provided data only at substantial intervals.

It was recognized that systems of vital registration were generally poor in most developing countries and it would therefore often be necessary to rely upon population census and sample survey data as substitutes for vital registration statistics.

The family planning service statistics system provided data on acceptors and their characteristics, and on other programme factors, including contraceptives supplied, use of specific methods and so on. Those statistics were a by-product of the records usually required for delivery of service by a family planning programme, and as they were not, in most cases, designed or intended to serve evaluation purposes, the statistics that they produced were frequently inadequate for scientific evaluation. That short-coming in the quality of service statistics was frequently a major source of error in evaluation efforts, and the system should be improved. However, for a majority of developing countries, demographic and other related data also were generally of inferior quality. Indeed, the inadequate quality of the service statistics was a reflection of the absence in such countries of a history of good data collection systems. Certainly, improvement of data on the activities of family planning programmes should be a primary goal. It would be appropriate gradually and incrementally to build a system for collecting adequate data for evaluation.

The amount of information routinely recorded at family planning clinics for each client should be only that necessary for the delivery of services, which should be the primary emphasis at the clinic level rather than data collection. For research purposes, it was not essential to have service statistics for each client; a sample would frequently suffice. At any rate, the service statistics system should not be regarded as the sole source of detailed data on acceptors (i.e., their characteristics, their contraceptive practices and their continuation of contraceptive use).

Careful thought should be given to the statistical data required from the family planning programme and the means of collecting it. The statistics needed for studying the impact on fertility of a family planning programme might be obtained from samples of family planning clinics or service units or through the use of periodic surveys of women in the childbearing ages. In that way, the burden for the collection of additional or special types of data of a service statistics nature could be placed on the staff of the evaluation unit of the programme.

There was not consensus among members of the Expert Group with respect to the role of family planning programmes in providing certain essential data. In opposition to the view expressed above was the opinion that service statistics represented a third major data source in countries with family planning programmes and that efforts should therefore be devoted to obtaining and improving the quality of such data. However, service records on use of conventional contraceptives, for example, were probably not needed. Techniques should be devised for obtaining such data directly from the sources supplying the conventional methods.

One solution to the problem of data on the use of conventional contraceptives was to make provisions for the collection of such data when a programme was established. Bench-mark surveys were an example of what could be done. Knowledge-Attitude-Practice (KAP) surveys yielded useful information, particularly at the beginning of the programme and at its early stages.

The Expert Group recommended that data should be collected by a census or a sample survey at least once every five years. Where sample surveys were carried out, provision should be made for the collection of detailed areal and socio-economic data needed for family planning programme evaluation. The samples of surveys designed for evaluative purposes should be of adequate size to provide data both on an areal basis for a country and on significant subgroups in the population.

It was observed that in various respects, censuses, vital registration statistics and service statistics were all limited. Surveys must be designed to provide the types of data not usually included in those more limited sources. Although one could question whether there existed an over-emphasis on surveys, some sample surveys were indispensable if the data necessary for evaluative purposes were to be obtained. A well-designed and carefully planned periodic survey programme could obviate the need for many hastily designed surveys and might even reduce the total number of surveys in some countries.

Other aspects of data requirement problems considered by the Expert Group were the types of problems encountered in the collection and classification of data. A great deal of error in evaluation efforts could be traced to the types of errors that arise in data collection and classification.

The evaluator was required to adjust for two categories of errors. First, there were non-random errors and biases that either affected all units similarly or were distributed in such a way as to have differential effects. Those errors included, *inter alia*, errors of non-response, recall lapse, and digit preferences in age reporting. An excellent example of the latter type of error commonly encountered was reported in the case study of Karnataka State, India.[1]

Well-known techniques were available for identifying and adjusting for such errors. For example, a number of techniques were available for correcting faulty age distribution due to the misreporting of ages. The choice of the proper method of adjustment, however, could be a major problem. As indicated in the Indian case study, the application of two different procedures for smoothing the age distribution yielded different evaluation results.

In general, the identification of the magnitude and the direction of non-random errors posed a particularly difficult problem; even for experienced statisticians, who were fully cognizant of the ramifications of such problems, solutions were not always easy. Non-random errors might affect all types of variables. Because of their nature and the difficulties in observing and measuring them, the Expert Group felt that such errors were potentially the most hazardous encountered in the application of the various methods of evaluation.

The second category of errors encountered in data collection and classification that constituted serious impediments to obtaining valid results from evaluation efforts were those termed "random". Such errors could originate either from sampling procedures or from other sources. The random errors that arose from sampling procedures could be dealt with in a satisfactory manner, as standard statistical theory provided a means for handling such errors. But adjustments of the random errors that resulted from chance factors not easily identifiable and that frequently had a large cumulative effect could be made only by the most complex procedures and frequently with limited results. Methods of evaluation that involved matching and regressional analysis were particularly susceptible to the influence of random errors.

It was considered that the extent to which it was possible to deal satisfactorily with the various types of error bore directly on the interpretation of the results of evaluation studies. Reports of such studies should, therefore, include a description of the techniques employed to assess measurement errors and other types of errors as well as the types of adjustments made. If that policy were followed, analysts would invite greater confidence in the results obtained in studies of the impact of family planning programmes on fertility.

B. Minimum data needs

The Expert Group recognized that the problem of data requirements was ever present in attempts to measure the impact of family planning programmes on fertility. Indeed, as previously mentioned, the

[1] In the section of that paper which deals with data problems, see figure entitled "Raw and smoothed population distributions of Karnataka State, based on 1971 census"

evaluator's choice of method might frequently be dictated not by the purposes of the evaluation or the validity of results that might be derived from the method but by the type and amount of data at hand. Given a wide variety of high-quality demographic data and a full complement of reliable family planning service statistics, it would be possible to apply in one evaluation effort any or all of the methods currently used for that purpose. Moreover, researchers should be able to sharpen the measures and to develop new, and possibly more reliable, methods of evaluation. But such an ideal situation never existed, and evaluators were compelled rather to work with scarce resources, making adjustments and assumptions where data were inferior or altogether lacking.

The Expert Group undertook, therefore, to list the minimum statistical data required for each of the eight methods of evaluation that were being considered during the meeting. Table 1 summarizes the types of data that were thought to be essential for application of each of the eight methods of evaluation. The minimal data requirements shown in table 1 were detailed as described below.

Standardization approach

The standardization approach required: (*a*) crude birth rates of the population at two points in time; (*b*) age-sex-marital status distributions of the population at two points in time; (*c*) general fertility rates at two points in time; (*d*) marital age-specific fertility rates of the population at one point in time; (*e*) if possible, data on (*a*), (*b*), (*c*) and (*d*) separately for rural and urban areas.

Trend analysis

The trend analysis approach required one of the following: (*a*) crude birth rates; (*b*) general fertility rates; (*c*) marital age-specific fertility rates and proportions married; (*d*) gross reproduction rates; (*e*) distribution of women by number of children born (parity); and (*f*) age distribution of the population during the period of the programme and before the start of the programme.

Experimental designs

The experimental-design approach required the following data for both the experimental and the control groups: (*a*) socio-economic variables (e.g., income, educational level, labour force participation, religion, ethnic characteristics, residence); (*b*) demographic variables (e.g., sex, age, marital status, parity); (*c*) biological variables (e.g., foetal mortality, amenorrhoea); (*d*) general fertility rates and rates for both groups at the beginning and at the end of the observation period.

Couple-years of protection

The couple-years-of-protection method required the following data:

(*a*) Programme variables by method: (i) for IUDs,

TABLE 1. ESSENTIAL DATA NEEDED FOR EACH EVALUATION METHOD

Data needed	*Evaluation methods*							
	Standardization	*Trend analysis*	*Experimental designs*	*Couple-years of protection*[a]	*Component projection*	*Analysis of reproductive process*	*Regression analysis* *Areal*	*Regression analysis* *Micro*
Fertility								
Crude birth rate	X[b]	X[b c]	N	N	N	N	X[c]	X[c]
General fertility rate	X[b]	X[b c]	X	X	N	X	X[c]	X[c]
Marital fertility	X[b]	X[b c]	N	[d]	X	N	X[c]	X[c]
Gross reproduction rate	N	X[b c]	N	N	N	N	X[c]	X[c]
Parity	N	X[b c]	N	N	N	N	X	X[c]
Programme variables								
Number of acceptors, by method	N	N	N	X	X	X	N	(X)
Continuation rates, by method	N	N	N	X	X	X	N	(X)
Number of users	N	N	N	(X)	(X)	(X)	N	N
Measures of programme input	N	N	X	N	N	N	X	X
Non-programme variables								
Demographic variables	X	N	X	N	[d]	X	X	X
Socio-economic variables	[d]	N	X	N	[d]	N	X	N
Cultural variables	N	N	X	N	[d]	N	N	N
Mortality rates	N	N	N	N	X	X	[d]	N
Biological variables	N	N	X	N	N	X	N	N
Other	N	N	N	N	N	X	N	(X)

Note: X = minimum requirements;
(X) = estimates from other data;
N = no application for data in method.

[a] Modified to give births averted.
[b] Data needed for two or more points in time; any one fertility measure would suffice.
[c] Any fertility measure would suffice.
[d] Optimum data.

number of users or number of acceptors and continuation rates for two points in time; (ii) for pills, number of users or number of acceptors and continuation rates for two points in time, or number of cycles distributed and estimate of wastage; (iii) for injectables, number of injections; (iv) for sterilization, number by age and sex; (v) for conventional contraceptives, number of units distributed and estimate of frequency of use and wastage;

(*b*) Fertility measures: fertility rates or age-specific fertility rates if users were known by age.

Component projection approach

The component projection approach required the following data: (*a*) programme variables by method as listed above for CYP but by age of user; (*b*) pre-programme age specific marital fertility rates of acceptors; and (*c*) age/sex-specific mortality rates.

Analysis of reproductive process

Potter method

The following data were needed for the Potter method[2] of analysis of the reproductive process. In the case of the IUD: $B = I/D$, where $I = F(R - A - PW)$ and D = average birth interval per birth required in the absence of programme contraception (i.e., the reciprocal of age-specific pre-acceptance fertility rate classified by never-use and use, or combined never-use and use of contraception during the pre-acceptance period of from three to five. In the formula to obtain I:

F = proportion of couples not sterile at time of acceptance deduced from L. Henry's estimates of age-specific sterility proportions (no special data required);

R = mean time programme contraception was used among couples not sterile at time of acceptance (deduced from the two parameters of the modified negative exponential distribution, which required data for an ordinary life-table continuation function as well as mortality data to calculate an age-specific annual risk of marital dissolution by death of a spouse and as judged necessary, data on divorce to calculate the annual rate of divorce; with the risks of discontinuation, marital dissolution by death of a spouse, and of divorce treated as independent within age classes);

A = mean overlap of contraception and post-partum anovulation (which required data on mean intervals from birth to acceptance of the method and on estimates of mean length of post-partum anovulation);

P = estimates by age class of the proportion of IUD terminations attributable to accidental pregnancy (which required data on reasons for discontinuation that included accidental pregnancy);

W = an estimate of the mean number of fecundable months required per conception in the absence of IUD, by age (usually a rough deduction using Perrin-Sheps model[3] from the values of F and anything known about level of pregnancy wastage and post-partum anovulation).

Wolfers method

The data needed for the Wolfers[4] method of reproductive analysis were thought to be about the same as those required for the Potter method, except that the estimates of potential fertility were obtained directly from questions related to birth intervals. Both cases involved the calculation of a life table following month-by-month post-acceptance experience. The table could be constructed for as long a duration as the data would permit. The results of the life-table calculations gave the proportions of acceptors who were still using a method and the number of births that were being averted. Separate tables could be calculated for different subgroups of the population (e.g., age, socio-economic group, type of contraceptive method used). In respect of the Wolfers method, corrections for duration of contraceptive use, post-partum amenorrhoea, accidental pregnancy and proportions who were becoming sterile were applied monthly.

Regression analysis

Application for areas

The application of regression analysis for areas required the following data for each area: (*a*) one current fertility measure other than parity (e.g., crude birth rate); (*b*) some measure of programme inputs (e.g., number of clinics, number of family planning personnel) and/or other programme variables that would summarize conditions prior to the date to which the measure of fertility related; (*c*) measures of non-programme variables having a direct or indirect effect on fertility (e.g., percentage of women in reproductive ages married, urban-rural residence and socio-economic variables).

Application for micro data

Regression techniques for micro data required: (*a*) one current fertility measure; (*b*) some programme variables (e.g., acceptors by method, continuation rates by method, a measure of programme input); and

[2] For a detailed exposition of this method, see R. G. Potter, "Estimatimg births averted in a family planning program".

[3] Edward B. Perrin and Mindel C. Sheps, "Human reproduction: a stochastic process", *Biometrics*, vol. 20 (1964), pp. 28–45.

[4] For a detailed exposition of this method, see D. Wolfers, loc. cit. Further, for a proposed synthesis of the Potter and the Wolfers methods, see David Wolfers, "Births averted", in C. Chandrasekaran and Albert I. Hermalin, eds., *Measuring the Effect of Family Planning Programs on Fertility* (Liège, International Union for the Scientific Study of Population for the Development Centre of the Organisation for Economic Co-operation and Development, 1975), pp. 163–214.

(*c*) some measures of non-programme variables (e.g., age of acceptors or other demographic characteristics, urban-rural residence, educational level); biological variables, such as foetal mortality, amenorrhoea and menopause status of women.

In the opinion of the Expert Group, it could not be overstressed that, in all cases, it was important to have accurate statistics and that data of at least moderately high quality were essential. The requirements specified by the Expert Group were far from optimal and, accordingly, some analysts might wish to impose higher standards. For example, for the estimation of couple-years of protection, statistics often were not available on users of conventional contraceptives; and, therefore, a crude estimate could be made from the number of pieces distributed to clients if it were possible to derive reasonably good estimates of frequency of use and of wastage.

However, the data suggested by the Expert Group represented minimum requirements; and, ideally, a greater variety of statistical information would be useful. In addition, it should not be assumed that the module of data needs presented in the present report was the only one that would serve; ultimately, that factor would depend upon the evaluation model applied. Further, it was of importance that although data on some of the variables might be needed for evaluation at one time, on another occasion a different set of data might be required.

C. Potential fertility

The Expert Group considered the estimation of potential fertility to be one of the most difficult problems encountered in the application of several of the methods. Potential fertility was a basic component of six of the eight techniques distinguished in the main background paper (ESA/P/AC.7/1) prepared by the Secretariat for the Expert Group meeting. The only exceptions were standardization and perhaps regression analysis. In regression analysis, however, potential fertility might be viewed as the fertility predicted by the regression equation when the independent variables which represented programme input were set at zero. In trend analysis, the projected birth or fertility rate was a measure of potential births. The projected (potential) births less the observed births yielded the number of births prevented. In an experimental design or matching study, the fertility of the control group defined potential births. The fertility of the control group minus that of the experimental group defined programme impact. In the CYP approach, couple-years of protection multiplied by an assumed rate of potential fertility yielded births averted. The same basic paradigm figured in the component projection approach. In the analyses of the reproductive process, a highly corrected mean couple-years of useful protection per acceptor was divided by a potential birth interval to give births averted per acceptor. In the case of simulation models, fertility rates with and without family planning were obtained. The latter measure of fertility was defined as the level of potential fertility. The simulated fertility rates with and without family planning were based on hypotheses about the value of various biological and family planning parameters. The derived fertility was then differenced as an estimate of the impact of a family planning programme.

Since either the measure of potential fertility or the observed fertility rate was being differenced or, alternatively, women-years of protection were being multiplied by a potential fertility rate, the resulting estimate of the impact of a family planning programme was extremely sensitive to the estimate of potential fertility.

The Expert Group noted that a distinction existed between total family planning programme impact and net programme impact. Total programme impact measured the effects of contraceptive use and other measures of fertility control regardless of whether the users adopted a method through the programme or adopted one outside the programme. Net programme impact took into account both the substitution effect (i.e., the extent to which the observed fertility behaviour would have occurred in the absence of the programme) and the catalytic effects of the programme (i.e., the extent to which the programme had stimulated fertility control that had at that time been provided from non-programme sources).[5]

Although potential fertility was non-observable, probable ranges of potential fertility might be estimated. The Expert Group took note of the statement in the background paper: ". . . estimating potential fertility consists of determining a particular fertility level that did not materialize. It is not possible to know with certainty what such a fertility level would have been and this problem cannot be solved. The purpose of the procedure is thus to compute reasonable estimates" (ESA/P/AC.7/1, section B).

Even if one dealt with total family planning impact and even if one treated potential fertility as natural fertility, the measurement of the potential fertility of contraceptive users was not straightforward. Users were to some extent selected for higher than average fecundity and were non-random with respect to physiological status, as there was less chance that they were pregnant or secondarily sterile than a married woman drawn at random from the same age class. If interest was focused on net programme impact, the Expert Group felt that there were a number of additional issues: what alternative contraceptive methods existed that the programme clients might have used if there was no programme; how much of contraception adopted by women outside the programme was attributable to the catalytic effect of the programme.

The basic significance that estimates of potential fertility had in most measures of family planning im-

[5] See Ronald Freedman and Bernard Berelson, "The record of family planning programs", *Studies in Family Planning*, vol. 7, No. 1 (January 1976), pp. 1–40.

pact and the difficult problems encountered when an attempt was made to estimate potential fertility had been previously recognized in a critique of the births-averted concept.[6] The Expert Group believed that progress would be facilitated if the concept of potential fertility, the estimation techniques and the data requirements were related to the classification of approaches presented in the main background paper (ESA/P/AC.7/1, section B). At the empirical level, progress could be achieved if provision for collecting data relevant to estimating potential fertility were built into study designs. Current designs tended to focus on acceptance and continuation measures of clients' behaviour in a family planning programme and to neglect the problem of estimating potential fertility, especially the problems of definition and measurement.

Problems of definition

Use of the term "potential fertility" to describe the level of fertility that might have characterized a population if its past reproductive pattern and the conditions that affected it had remained unchanged was not an entirely satisfactory means of identifying the concept. And although "expected fertility" was thought to be a reasonably clear term, since it had a well-defined meaning and was clearly understood to imply what would be expected under certain conditions, its meaning in demography was well established as applying to values derived in the standardization method. An alternative term might be "hypothetical fertility", which would imply the unrealistic nature of the measure. Moreover, the term "hypothetical" would distinguish the concept from that of "expected" fertility. It would be unfortunate if potential fertility were to be equated with expected fertility.

The Expert Group cautioned that potential fertility was not to be confused with "natural fertility", which was the level of fertility exhibited by a population that made no deliberate attempt to control its fertility. Natural fertility was not the biological maximum. In most populations, a number of socio-cultural checks tended to operate in such a way as to depress fertility well below what was believed to be near the biological limit. Therefore, defining potential fertility at the level of the biological limit would produce an overstatement of potential fertility in most populations. As previously stated, in the framework of family planning programme evaluation, potential fertility referred merely to what would have happened to the fertility of a population in the absence of the programme.

Another aspect of the definitional problem was that of determining the group of women on whose fertility performance estimates of potential fertility should be based. The Expert Group discussed the question of the various groups of women for whom potential fertility could be calculated and suggested possible time dimensions for the analysis. The results of their discussions all shown below in table 2. The time periods chosen indicate that measures of potential fertility could be calculated for a period either prior to the initiation of a programme or after a programme had begun to operate. If the measure of potential fertility to be used was for some period prior to the initiation of the programme, it could be based either on observed fertility rates, which were preferable, or on some estimates of those rates. But for many countries, the requisite data were not available. Thus, the measures of potential fertility would have to depend upon some estimates of the fertility level. Alternatively, that measure could be calculated for the period following the beginning of the programme. As indicated in table 2, either observed or estimated fertility could be used as the measure of potential fertility for the first three categories of women. If the measure was based on observed data, such a measure would, however, presumably produce underestimates of potential fertility since acceptors would be included in those groups.

If the measure of potential fertility desired was for acceptors only or for other more specially defined groups of women (as suggested in the lower part of table 2), it would be necessary to estimate what the fertility of those groups of women would be if no programme had been in existence. It would be required to base the estimates on a number of assumptions, and the type of assumptions that should be made posed difficult problems. If potential fertility was calculated for acceptors, it could be based on specific methods of contraception; women selecting different methods differed in terms of their motivation; and, for that reason, their potential fertility would differ. Another dimension that might be introduced into the table was that of other characteristics of acceptors (e.g., educational level or place of residence). Thus, other columns could be added to the table to show the measures of observed fertility that might be subtracted from the potential fertility measures.

There was great difficulty in classifying women as acceptors. Many women entered a programme, left it and then re-entered. To facilitate research on evaluation, acceptors might be classified into various subgroups, such as ever-users and current users. The term "use" implied prevalence and the term "acceptor" implied incidence, and the Expert Group felt that there was no need to link the term "acceptor" with that of "user".[7]

Problems of measurement

Opinions varied as to the appropriateness of measuring the potential fertility of different groups of women

[6] W. Parker Mauldin, "Births averted by family planning programs", *Studies in Family Planning*, vol. 1, No. 33 (August 1968), pp. 1–7.

[7] The general definitional problem of "acceptors", "users" etc. has been discussed in detail by another United Nations Expert Group, and the definitions recommended by that Group should be followed. See *Assessment of Acceptance and Effectiveness of Family Planning Methods*, report of an Expert Group meeting, Bangkok, 11–21 June 1968; Asian Population Studies Series, No. 4 (United Nations publication, Sales No. E.69.II.F.15).

TABLE 2. POSSIBLE POTENTIAL FERTILITY MEASURES FOR VARIOUS GROUPS OF WOMEN

Group of women	Time period: Pre-programme initiation ($t-x \ldots t-3$ $t-2$ $t-1$)	t 0	Post-programme initiation[a] ($t+1$ $t+2$ $t+3 \ldots t+x$)
All women aged 15-44 (49)	Observed or estimated		Observed or estimated
All married women aged 15-44 (49)	Observed or estimated		Observed or estimated
All fecund women aged 15-44 (49)	Observed or estimated		Observed or estimated
Acceptors: Method 1, Method 2, Method 3, . . . etc.	Observed or estimated		Estimated
Women by postpartum status	Observed or estimated		Estimated
Other special populations (ever-users, current users etc).	Observed or estimated		Estimated

[a] The time periods might also relate to intervals after the programme was initiated, such as from t_1 to t_3, and not merely to a period from the beginning of the programme to some years after, such as from t_0 to t_1, from t_0 to t_2 etc., as specified in the table.

(as shown in table 2). One view held that it was not of importance which group was used to derive a measure of potential fertility and that it was not necessary to take account of the higher fertility of acceptors, since that fertility quickly regressed to the mean. Although it might be appealing to estimate potential fertility of acceptors on the basis of their previous fertility, it could be viewed as a transient phenomenon and would not persist, as the regression to the mean was extremely strong. However, it was of interest to know the number of years required for that regression to the mean to occur. After all, it was only by chance that the women fell into two groups, i.e., whether they had high or low fertility. Furthermore, even the differences among various groups of women in continuation rates would tend to disappear after a certain period. Thus, it was unnecessary to calculate the potential fertility specific for various contraceptive methods. Any selectivity of women that operated at the beginning of a programme would tend to wash out and acceptors would become more homogeneous. At higher ages, however, the higher fertility of acceptors might be a problem; but because an increasing number would become sterile, that factor might not pose too serious a problem. If it were possible to use only fecund women, some of the major problems of selectivity might be overcome. That grouping, however, was difficult to arrange, inasmuch as some acceptors, even in a post-partum programme, were sterile when they accepted.

It was noteworthy, however, that it had not been shown conclusively that differentials in fertility and continuation rates, in fact, regressed to the mean, particularly in the short run. Rather large differentials in fertility as well as contraceptive practice continued to be observed in developing countries where contraception had long been practiced.

Simulation models might aid in answering some of the questions posed regarding the groups for which potential fertility should be calculated. Perhaps, in certain conditions, it would be more advantageous to use one group than another. Simulation would, perhaps, give some indication of which groups should be used. It was pointed out that, in the view of some researchers, the exact value of potential fertility was not essential and that a range of values within which potential fertility probably lay might be sufficient.

Opinions differed also regarding the necessity for introducing corrections for such factors as the overlap of contraceptive practice with the post-partum non-susceptible period, accidental pregnancies and the effect of acceptors in a programme merely substituting for other contraceptive practice that might have occurred anyway. On the one hand, considerable error could be introduced if corrections were not made for those factors. Imposing a penalty per accidental pregnancy could have a significant impact on calculated protection time when considerable emphasis was placed on conventional contraceptives, with irregular practice leading to many failures. Given a post-partum programme of contraception, especially in a context of lengthy lactation, it was critical to subtract from credited protection time the average period of overlap between practice of contraception and lactation amenorrhoea. Lowering the estimate of useful protection time in order to reflect the proportion sterile might be important for the older age-classes of acceptors (e.g., ages 35–39 and older). If mortality or divorce levels were high, it might be worth while to correct continuation rates for marital dissolution by death of a spouse and by divorce or separation.

On the other hand, a number of those corrections would not influence the results to any important extent, and the lack of data on which to base the correction factors was, in any case, a major hindrance. In addition, in many instances, relatively poor data were being subjected to rather refined corrections and thus such corrections were not warranted. Only when there was good evidence that a particular factor, such as a particularly long post-partum period resulting from lactation, was operating, should certain correction factors be introduced.

A number of other factors could be identified as affecting the measurement of potential fertility. The effect of many of those factors would be to produce increases in the level of fertility while the effect of other factors would be to produce declines in fertility. Accordingly, the operation of some factors would offset the effect of other factors, and, therefore, no change in fertility levels would be observed. In either instance, the estimation of potential fertility could be particularly difficult.

If the birth rate remained stable as a result of a number of offsetting factors, such as changes in population structure, the impact on fertility of a family planning programme might not be discerned. That situation would be particularly true if the effect of a number of factors was in the direction of increasing fertility while the effect of the family planning programme was to offset the potential increases in fertility, with the result that no change in the birth rate would be observed. In such a circumstance, to conclude that the family planning programme was having no effect would be erroneous.

Among the factors that could have a positive effect on fertility were the removal of the traditional checks on fertility, changes in socio-economic conditions, changes in mortality and migration, and changes in marital patterns.

The weakening of traditional checks might take such forms as decreases in the practice of breast-feeding or a shortening of its length, the removal of taboos on intercourse during certain periods and the decline in prolonged visits by the wife to her mother's home. Such factors might have contributed to a rise in potential fertility in one state of India. In urban areas of many developing countries, the decline in breast-feeding practices might be especially significant. In general, however, the effect of changes in those factors was considered to be highly conjectural, as very few studies had been carried out that clearly demonstrated their effects.

Improvements in socio-economic conditions, which had an over-all effect on levels of living, could act differentially in increasing or decreasing fertility. For example, improvements in nutrition were believed to increase fertility, especially when previous levels of nutrition had been very low.

Changes in mortality and migration also played some role in affecting the level of fertility. Declines in infant mortality were not believed to have much effect on decreasing fertility in the short run,[8] but decreases in the mortality of spouses had the effect of increasing marital durations and might therefore contribute to rises in fertility. On the other hand, increased migration that involved separation of husbands and wives for relatively long periods could have a depressing effect on fertility.

D. Correlated variables and interaction; uncontrolled variables and independence of method

The Expert Group recognized that the effects of a family planning programme on fertility must be viewed as a complex system of many interconnected parts, a point noted explicitly in the statements by Ebanks[9] and Brass.[10] Both writers stated that the evaluation of family planning programmes should be viewed in the larger context of studying the determinants of fertility.

The problem then was to identify the boundaries of the system, its component parts and the linkages among these components. Stated in other terms, that task was in the realm of theory building or model construction. Such efforts must involve the identification of the key concepts or variables and the form of their relationships. The Expert Group noted that the papers by Sirageldin,[11] Wells[12] and Srinivasan[13] took up that problem in some detail.

The need existed for a conceptual model of the underlying process regardless of the analytical technique used—whether it was regression or simple tabulations. Also, problems of data availability should be divorced from those of model construction. If the needed data were not available, substitutes should be devised or the model modified. In that way, it would be possible explicitly to recognize what variables were not included.

The Expert Group recognized that a theory was an arbitrary construction, and it was best viewed not as true or false, but as fruitful or less fruitful. Theories sought to explain known relationships; if possible, they did so in a parsimonious manner, and they led to new observations. Theories might be developed by beginning with a few basic concepts and adding factors

---

[8] See Committee for International Co-ordination of National Research in Demography, *Seminar on Infant Mortality in Relation to the Level of Fertility* (Paris, 1975).

[9] G. E. Ebanks, *op. cit.*

[10] William Brass, "Comments on comparison strategies for the evaluation of family planning impact" (Conference room paper 14).

[11] Ismail Sirageldin, "Some issues in determining appropriate methods of evaluating the fertility impact of family planning programmes" (Conference room paper 4).

[12] H. B. Wells, *op. cit.*

[13] K. Srinivasan, "Interaction of socio-economic changes with family planning programmes: an assessment model" (Conference room paper 7).

found to be important—an incremental strategy. On other occasions, one might have a large number of factors and proceed by removing those found not to be salient in explaining the phenomena under consideration. Although it was desired to avoid bias, a strategy of including every possible variable was often not wise; the introduction of many variables might lead to other statistical artifacts equally serious.

Given that a complex system was involved, there was also the likelihood of correlated variables, i.e., effects produced by two or more factors which were highly interrelated. The implication of that possibility was that one could not hope to partition out the variance of a dependent variable explained uniquely by those "independent" or causal factors. Some portion of the variance would be due to the compounded effect of two or more factors. Though the nature of that limitation was most clearly formulated and quantified in terms of regression analysis, it was a common limitation running through all the methods and it arose from the complexity of the system under study. For example, in standardization, when the difference between two rates was decomposed, it was often found that some portion was due jointly, say, to age and marital status, and could not be attributed uniquely to either factor. The correlated variable problem might be more severe with residual techniques, since there was clear omission of variables known to be important and which might be correlated with those included. Unlike some of the other methods, the regression model at least attempted to include the family planning inputs explicitly.

Interaction effects could be seen as a function of the form in which the relationships were expressed, to use the terminology of regression analysis. If a linear additive model applied, it was possible to say how much change in the dependent variable was associated with a unit change in an independent variable, taking into account the other independent variables. If there was interaction, so that the effect of variable A depended upon the level of variable B, then that deduction is no longer possible. Although interaction was often suspected, it was worth while to test the adequacy of the simpler model (i.e., without interaction) to gain some of the benefit of the unambiguous allocation of effects.

Further, it might be desired to distinguish operational interaction—where a programme varied the type of input by the characteristics of the area, e.g., urban-rural distribution or income levels—from statistical interaction. With regard to the latter factor, interaction might not occur over part of the range of the variables involved, but might appear at another part of the distribution.

The uncontrolled variable problem arose when one or more factors that affected the dependent variable were omitted from the model. The penalty depended upon the importance of the omitted variable to the dependent variable and whether it was correlated with the independent variables included in the model. When a variable was omitted, there occurred a loss in explanatory power that was proportional to the importance of that variable. If that variable was not correlated with the variables in the model, then the effects attributed to the included variables were not biased or misleading; but if there was correlation with the omitted variable, then the measured effects would be biased.

It was desirable to explore ways of determining the degree of independence of methods. Just as the variables in any one model were correlated, there was likely to be correlation among the methods of evaluation, arising from common data, assumptions etc. It might be useful to envisage an observation arising from the application of a method as due in part to what was being measured, in part to random error and in part to method variance, i.e., systematic error attributable to the data and procedures employed. If a number of observations were obtained from different methods applied over a sufficiently large number of settings, it might be possible to decompose the relative magnitude of each of those factors attributable to each method. That point was discussed in detail in the statement by Hermalin.[14]

Many of the problems of correlated variables, interaction and uncontrolled variables were discussed by the Expert Group in the context of regression techniques. One current problem with areal regression models was that often it was not known what variables should be included, or it was unclear what the variables included represented. In addition, the results appeared to be very sensitive to the inclusion or omission of particular variables. The results, therefore, were very difficult to interpret.

As one solution, the problem of proxy and surrogate variables might be alleviated by including a number of indicators of an underlying concept in such a way that they would be recognized as indicators. Developments with "unobserved variable" models and confirmatory factor analysis were also suggested as promising in that respect.

An advantage of regression might be that the method bypassed the problem of potential fertility and births averted, in that it was not necessary to develop measures of these concepts explicitly. But that viewpoint did not find consensus among researchers, as some would maintain that it was a disadvantage and that measures of births averted provided more validity than did the coefficients obtained with regression techniques.

The method used depended upon the purpose of the evaluation; if the purpose was to understand the relative importance of family planning vis-á-vis development, the appropriate procedure was regression analysis. In contrast, if the aim was to trace out the effects of a regimen of contraceptive use, births-averted analyses making detailed use of data on various aspects of the reproductive process were prefera-

[14] Albert I. Hermalin, "Avoiding an embarrassment of riches" (Conference room paper 10).

ble. However, it might be possible to study the effects of development on certain aspects of the reproductive process through regression techniques.

It was pointed out that several difficulties arose with the regression approach. Areal regression typically did not separate programme acceptance from non-programme acceptance, but sought to establish the over-all effect of the programme in comparison with other factors that were operating. However, data on non-programme acceptance could be utilized in a regression model. Sometimes by partitioning the areas at the beginning of the programme by different levels of fertility, it was possible to achieve controls for non-programme contraceptive use and implicitly for potential fertility.

The allocation of effects in temporal terms was very difficult in regression analysis, even if one worked with a model containing various lags. If there were strong, interareal correlations among the areal structural variables, then the results would not be too sensitive to the particular lags used. But that fact did not make temporal allocation any easier. Methods that made use of a series of cross-sectional data might be a means for improving the understanding of the time dimension.

Multiple regression was a form of multiple comparison and one should contemplate other types of multiple comparison, which could get at more detailed aspects of the process. It was possible, for example, to set up multiple comparisons from individual data and to study the effect of non-programme factors, such as breast-feeding. That procedure could serve to validate some of the results for areal regression and also to answer additional questions. For example, some aspects of development might increase fertility as well as reduce it; different educational groups might react differently to a family planning programme. Because it was difficult to observe such differentials through regression analyses, there was a need to give more attention to the reliability of regression results, e.g., by splitting the sample in half or by other means.

Areal regression results were likely to be sensitive to the number of observations and the particular measures used, as well as to a number of other factors. Thus, it was important to replicate results with different models and different measures, as well as to validate the results with alternative techniques. The results of a regression analysis should be compared with the results of other analytical methods, such as controlled cross tabulations. Cross tabulations made less severe demands on the nature and form of data. Attention should also be given to studying the effect on the results of areal regression at different levels of aggregation. Distinctions between categorical and interval data usually were not as important as they had been. Recent developments permitted the testing of linear models for categorical data.

In so far as areal regression was hampered by lack of data, improvements could come through the use of the community-level data obtained in surveys. Also, it should be possible to aggregate the individual level data of surveys and to form areal measures that would not ordinarily be available from the usual statistical sources.

III. COMPARISONS OF EVALUATION METHODS

A. Problems of interpreting the validity of evaluation measures

Some essential questions concerning evaluation which arose in the discussions were whether a method permitted the logical progression from the results of measurement to identification of programme impact; whether the measures obtained provided an adequate explanation or description of family planning efforts or acceptance; and whether the methods yielded a valid estimate, for example, of potential fertility, if the purpose was to quantify that phenomenon.

The techniques currently in use were generally capable of quantifying the phenomena that they had been designed to measure, and there had been some success in obtaining useful explanations through application of some methods. However, serious shortcomings existed with respect to the logic of some of the methods; and as validity of reasoning was an indispensable attribute, additional research was required to correct those inadequacies. An illustration of the problem was that infant and child mortality were often regarded as variables that directly influenced fertility; but much, if not all, of that relationship had been shown to come from the effect of fertility on mortality, which raised the issue of how associations were to be interpreted in causal terms.

In demonstrating the dimensions of the problems of validity, use might be made of a physical analogy of a cloud of particles of different colours and sizes, the net effect of whose interaction was movement; the movement of the cloud was the trend in the birth rate and the particles were the variables influencing the trend in the birth rate.

The researcher was interested in the cause of that movement. To determine the cause, one of three courses was open. First, the researcher could focus on only the red particles (i.e., the family planning programme variables) and a few other surface or large ones, such as age and marriage, and determine their net impact, ignoring the effect of "small" or less evident particles thought to have less influence. Another approach would be to break down the cloud (fertility behaviour) into subvolumes, tabulate types of particles within each and in a regression analysis relate them to the movement of the component parts of the cloud. Or, the cloud could be divided into sections and an analysis made of the way in which the family planning particles interacted with the others in a few chosen sections in varying circumstances and under whatever conditions of control might prove feasible. Those courses corresponded with the three main procedures for assessing family planning impact.

Each approach or evaluation method involved a certain amount of speculation about some part of the process, and each took account of some important variables, while it ignored others. Evident also was the fact that with respect to each method, a speculative "leap" was required from the results obtained to an explanation of the impact of family planning programmes on fertility. But that "leap" was guided by some model or theory about the nature of the forces at play.

Ordinarily, the results of two or three methods would test validity, with a similarity of findings adding persuasion as to acceptability, though repetition of results did not really validate results; plausibility was also a necessary criterion of validity.

However, the third approach cited above could provide the crucial link between programme activities and accompanying changes or non-programme variables. After all, to elucidate the factors affecting change in aggregate fertility, it was necessary to determine how both non-acceptors and those who accepted a contraceptive method from the programme reacted to the changes in society, as well as how they reacted directly to the family planning programme. That determination would require application of an experimental design or a quasi-experimental design, because pure control, although desirable, was virtually impossible. However, "opportunistic observation" might be a more appropriate term than quasi-experimental design, as the aim should be to examine any relevant process, whenever and wherever possible.

Regression analysis should be undertaken to ferret out the impact of family planning programmes on fertility from that of other interrelated variables. It would be advantageous, for example, to divide the population by area according to the level of inputs of the family planning programme, and to subdivide those units according to population characteristics. The areas could then be examined in a parallel fashion to determine the way in which the family planning programme variables were related to changes in fertility. In that connexion, some ratio of medical and paramedical personnel to population size might prove a useful index. As noted previously, some research in that area was sorely needed.

The regression method was also a means of adjusting or controlling for various factors in an experimental or quasi-experimental situation. However, it was pointed out that some researchers held the view that experimental design yielded more reliable results.

Not enough attention had been given to the appropriateness of a method to determine achievement of a specified objective. In addition to the measurement of net impact, there was the objective of target setting which had a future orientation. Simple areal regression analysis had not been designed to do that task; for that purpose, the component projection approach was more useful. More attention should be given to evaluation of the effects of the intermediate factors on fertility.

In order to measure the influence of non-programme factors on fertility, it was necessary to break down the birth rate into the factors that influenced it, for example: (*a*) impact of the programme on acceptors; (*b*) induced effects of the programme on non-acceptors; (*c*) social and economic or community effects on acceptors; (*d*) social and economic or community effects on non-acceptors; (*e*) effects of interaction between programme and non-programme factors on acceptors and non-acceptors. The problem was to seek out the methods best adaptable for determining each of those influences and to ascertain which methods were independent and which were related.

The desirability of replication in evaluation both at different places and at different points in time could not be over-emphasized. In addition, some work should be undertaken in which there were parallel studies in the more developed and less developed areas of a country, so that both programme factors and social economic factors could be related to fertility.

Measurement of the impact of a family planning programme on fertility was not necessary for all programmes, although planners wanted to see a causal relationship. In many cases, however, the establishment of the order of magnitude of the effect of the programme on fertility should suffice. It was emphasized that the claim that declines in fertility were due only to non-programme factors, i.e. development only, was unproved. The evidence necessary to establish the impact on fertility of development was as difficult to marshall as that needed to establish the impact of family planning programmes.

It was recognized that there were many imperfections in the methodology in that sphere, and that no true experiments could be made. Each method required assumptions, as none could be handled entirely with observations. The incremental approach to building up the current knowledge in evaluation would obviously be the most fruitful.

B. Problems in determining appropriate methods of evaluation

Several criteria might be applied in the selection of appropriate methods for evaluating a particular family planning programme. Those criteria included: (*a*) the objectives of the evaluation; (*b*) the particular population for which the evaluation was needed; (*c*) the availability of data; (*d*) the independence of the method; (*e*) the validity and reliability of the methods; and (*f*) the cost of the method. It was pointed out that a detailed statement of the various criteria that might be applied was contained in the paper prepared by Sirageldin.[1]

An important criterion in the choice of the method of evaluation was whether the objective was to evaluate the short-term effects of a programme or whether the concern was with the long-run impact. Further, the frequency with which evaluation must be carried out dictated to a large extent the type of method that might be employed for evaluation. If the need was for a monthly indicator of the impact of a programme, none of the methods being considered could be said to be appropriate. Moreover, little data of any kind were available on a monthly basis. If the need was for an evaluation on a yearly basis or only once every five years, the choice of methods was much greater. Measures of the impact of a programme on fertility were not feasible for a period shorter than one year because of the biological requirements of reproduction. An obvious error would be to equate one abortion with one birth averted. Since most abortions were performed in the early months of pregnancy, one woman could experience up to three abortions in one year, and it would be inaccurate to state that that woman had averted three births in one year.

Since none of the methods considered at the meeting could be applied for a period shorter than a year where evaluation covered a period of that length, resort could be made to other indicators. The proportion either of acceptors or of users in a population might serve as an appropriate indicator for short-term evaluation. Such proportions at the early stages of a programme should give some indication of the probable eventual impact on fertility of a programme. Data on the number of acceptors and continuation rates of users currently available for Indonesia might be seen as indicators of a probable decline in the birth rate in that country.

Of the methods discussed at the Expert Group meeting, the couple-years of protection method was considered to be more appropriate for evaluations of a period of less than five years than for longer periods. In general, CYP was only a very rough measure, and it should be used only if the data were not available to calculate other methods. When the CYP method dealt with supplies, account should be taken only of those supplies which actually reached the people. The component projection approach was generally more satisfactory than the CYP method. The Expert Group deemed the component projection method the method of choice over the CYP method when the requisite data were available.

When the focus of the evaluation was on a longer period (i.e., for a period of five years or longer) standardization was a particularly appropriate method. Indeed, if the data were available, there also would be no reason not to use it for periods of less than five years; standardization was a basic method of assessing whether a programme might be having any effect. The component projection method was also highly appro-

[1] I. Sirageldin, *op. cit.*

priate for long-term evaluations. Trend analysis was highly appropriate if there was evidence of a strong trend prior to the beginning of a programme. However, various difficulties arose in its application.

In terms of the population for which the evaluation was needed, some methods were considered more appropriate than other methods. For instance, if the evaluation was concerned only with acceptors and the required data were available, analysis of the reproductive process would be appropriate. Only in rare instances, however, would the requisite data be available. Thus, that approach had limited utility in most evaluations, and, moreover, it was a very complex, detailed method, involving a large number of calculations that posed additional difficulties in any application. Experimental designs were particularly suited for evaluating the impact of a programme on acceptors. The usefulness of that method was especially appropriate with respect to small countries, as the necessary controls could be more easily introduced for them than for larger countries. High priority should be assigned to the use of experimental designed studies.

Since the couple-years-of-protection method usually could yield only births averted for acceptors in a programme, it was a limited tool for assessing the impact of a family planning programme on national fertility. In contrast, a major strength of the projection method was that it yielded estimates of births averted by all women rather than past acceptors. Regression analysis also was considered to be a highly appropriate technique, when the requisite data were available, but the data demands for that method were heavy. The need for clear conceptualization and a relatively high level of statistical competence also were major barriers to its use in many situations.

Simulation should not be used for evaluating family planning programmes. Rather, simulation models should be utilized as research tools, for resolving many methodological problems that arise in connexion with application of the various methods of measuring the impact of family planning programmes and, in particular, with interpretation of the results. The Expert Group noted that the statement prepared by Ridley[2] provided pertinent information on such uses of simulation models, including their appropriateness for validating the efficaciousness of different methods. Since the validity of the various methods was essentially unknown, an especially strategic utility of simulation models was in the area of validity studies. Other possible uses of simulation included quasi-experiments that aided in the selection of various programme strategies and application of the models to assist in target setting. However, simulation was a highly expensive tool and thus could be used only selectively.

A schematic summary of the most appropriate methods of evaluation deemed most appropriate for different time periods and groups of women is given below (table 3). Since it was the most appropriate method that the Expert Group was designating, analysis of the reproductive process approach and simulation models are not indicated as being appropriate.

[2] Jeanne Clare Ridley, "A note on two applications of micro-simulation models to the problems of evaluating family planning programmes" (Conference room paper 5).

TABLE 3. MOST APPROPRIATE METHODS OF EVALUATION FOR DIFFERENT TIME PERIODS AND GROUPS OF WOMEN

Method of evaluation	Time period								
	One year or less			*One–five years*			*Five or more years*		
	Total population	*Women aged 15-49*	*Acceptors only*	*Total population*	*Women aged 15-49*	*Acceptors only*	*Total population*	*Women aged 15-49*	*Acceptors only*
Standardization approach	N	N	N	A[a]	A[a]	I	A	A	I
Trend analysis	N	N	N	I	I	I	A	A	I
Experimental designs	N	N	N	I	I	A	I	I	A
Component projection approach	N	N	N	I	A	A	A	A	A
Analysis of reproductive process	N	N	N	I	I	I	I	I	I
Regression analysis	N	N	N	A	A	A	A	A	A
Simulation models	N	N	N	I	I	I	I	I	I

Note: A = appropriate;
I = inappropriate;
N = no application for data in method.

[a] Data rarely available, as census statistics are required.

IV. NEEDED RESEARCH

National family planning programmes had been implemented almost exclusively in developing countries, which were generally deficient in demographic and other statistics. Lacking a tradition of quality data collection, processing and analysis, most of those countries had not succeeded in producing family planning service statistics of good quality nor in carrying out the needed research. Consequently, a great deal of research must be done if the impact of family planning programmes on fertility in developing countries was to be reasonably accurately measured and well understood. Lack of quality data was a serious handicap to research; but, apart from improving the supply and reliability of the data, it was required to sharpen and refine the research tools and to present them in such a way as to permit their wider use. In that connexion, attention was drawn to the statement by Mauldin.[1]

In the social sciences, methodological research was not popular; that unpopularity extended to methodological research relevant to the efficiency and impact of national family planning programmes and such research did not ordinarily receive donor support. Yet, studies of research method could reduce the cost and improve the value of other research. In recognition of that possibility, the Expert Group emphasized and recommended strongly that Governments and international agencies should give greater support to methodological research, which had real value for evaluation studies.

The Group specified the endeavours listed below as being among the more urgent research priorities:

(*a*) Research methodology should be improved and greater resources be devoted to the evaluation of programme impact on fertility;

(*b*) Studies should be undertaken in which the socio-economic variables were associated with the family planning programme variables, so as to account for as many as possible of the factors influencing fertility change. In that connexion, lack of an acceptable measure of programme input or effort constituted a major impediment to skillful and productive analysis;

(*c*) Extensive research should be undertaken on sources of error in the data used for evaluation. Studies of the types of data for which errors were common, how the errors arose and their extent were legitimate topics for research; the problem of errors in data merited far more attention than it had been given. It would be of great value, for example, if the World Fertility Survey would take a subsample of women and determine the validity of their responses. In any case, the determination of reporting errors common to certain groups should form an important part of future research relevant to evaluation of the impact of family planning programmes on fertility. There was a need to validate the results of one-time surveys that went beyond the follow-up and other surveys taken for evaluation purposes. Specifically, researchers should deal with the large question of whether one-time surveys yielded valid, dependable data;

(*d*) Not enough was known of the relevance of family planning programme variables to the influence that the programmes had on fertility. Studies should be made, for example, of the way in which the range of family planning methods offered by the programme and the availability of those methods affected the impact of the programme as a whole and in different social and economic settings;

(*e*) The advantages of the simulation technique had not been fully exploited in the area of evaluation research. It could be used very profitably to determine the sensitivity and robustness of the different methods of measuring impact of family planning programmes on fertility, as well as to investigate the effect of other measures designed to influence fertility. A particularly promising endeavour would be the application of simulation to the measurement of potential fertility for short-range projections;

(*f*) Additional research was needed on analysis of the reproductive process as a method of measuring family planning programme impact;

(*g*) The entire area of target setting had not been given sufficient attention, and there was an urgent need to examine the hypotheses and methodologies which related to establishment of goals in family planning programmes;

(*h*) In some programmes, research should be undertaken which would lead to assimilation of knowledge on the question when a family planning programme should be terminated. For example, with due regard to the human element, there should be experimentation with the termination of the programme in a variety of socio-economic settings;

(*i*) Among the many other important areas to which attention should be given was the question whether it was necessary that the impact of every family planning programme be evaluated. Another question to which

[1] W. Parker Mauldin, Needed research for measuring the impact of family planning programmes on fertility" (Conference room paper 12).

research should be directed was whether family planning programmes met the needs and wishes of the people that they were intended to serve. And there should be extensive research on the extent to which fertility change was a function of development. The Expert Group was of the opinion that those and related questions were a part of the matrix of issues surrounding evaluation and should receive a fair measure of the resources available for research on the impact of family planning programmes on fertility.

V. RECOMMENDATIONS

A. Data collection and data quality

The Expert Group repeatedly expressed concern for the deficiencies in the supply and quality of data required for evaluation of the impact of family planning programmes on fertility. Vital statistics registration is inadequate for nearly all the developing countries and non-existent for a majority of them. In its view, years of international effort have failed to produce statistics of the quality needed to measure short-term changes in fertility, and researchers must rely mainly upon results of population censuses, which are generally short on fertility data and which, in any case, are taken at intervals too infrequent to meet essential data needs. Moreover, family planning service statistics also are often inadequate for evaluation purposes. Accordingly, the Expert Group made the following recommendations:

1. Governments and international agencies should give the highest priority to improvement of the supply and quality of data required for the few basic measures of fertility needed to assess family planning programme impact;

2. Countries with a good census-taking system but with poor vital statistics systems should give serious consideration to conducting a population census or large-scale demographic survey every five years;

3. Decennial censuses should be followed by post-censal surveys that would allow for more intensive interviewing on a broader range of topics than is feasible in a population census. These post-census fertility surveys should obtain interviews using family planning modules that would incorporate, along with a fertility history, programme policy variables as well as major socio-economic variables. The sample design should facilitate a community stratification system that would allow for analysis of the interaction between individual and community variables;

4. As a supplement to family planning programme statistics, or in lieu of them where they are inadequate, sample surveys at reasonable intervals should be employed to obtain the information on acceptors and other programme variables that is required for evaluation purposes. The scheme for data collection in these surveys should provide for the classification of respondents by their family planning status, i.e., ever-user (of a method from the programme, a method from another source), never-user, current user and so on,[1] and some measure of their fertility;

5. Intensive effort should be made by national programmes and by international agencies to improve the quality of family planning service statistics and to increase the supply of information on acceptor characteristics by geographical areas to facilitate special analytical studies. These service statistics should be published periodically, and the reports should include some estimate as to the quality of the data.

6. While family planning service statistics systems should be made to realize their full potential as sources of useful statistical data, the systems should not be so overloaded as to impede the delivery of the services;

7. The United Nations should use its influence to urge and assist the World Fertility Survey in continuing to give high priority to inclusion of the family planning module in surveys conducted under its auspices where national family planning programmes exist. Further, the United Nations should urge the World Fertility Survey, wherever feasible, to invite its participating Governments to include the community-level data module, so as to facilitate research on relationships between individual behaviour and relevant conditions in the society. In addition, it is of the first order of importance that the World Fertility Survey should be encouraged by United Nations to undertake methodological research on the reliability and validity of data obtained in fertility surveys and that funds from bilateral and multilateral sources be sought for this purpose.

B. Improvement of methodology

The Expert Group is of the opinion that although researchers have made much progress in refining the techniques of evaluating the impact of family planning programmes on fertility and in coping with many of the critical methodological issues, there remains, none the less, a need for considerable work in this sphere. To facilitate greater progress, it recommended that:

1. A concerted effort should be made to determine how the validity of different methods of evaluation could be established. Among other things, several methods, including areal regression analysis where feasible, should be applied simultaneously in evaluation as an aid to determining the reliability of the results achieved by each; multiplication of evidence would serve as a basis for greater confidence in the methodology;

2. Although pure control permitting experimental designs in their classical forms may be impossible to establish for family planning evaluation purposes, experimental designs should, none the less, be viewed as

[1] *Assessment of Acceptance and Effectiveness of Family Planning Methods*, p. 7.

an ideal evaluation procedure, and every effort should be made to approximate them. Currently, quasi-control groups serve as a proxy for true control in evaluation studies. Important advances in knowledge of the impact of family planning programmes on fertility would result from such quasi-experimental studies of the relations between family planning programmes and fertility change, where an effort is made to control or randomize other possible determinants; such studies should be provided for in both data collection and evaluation research schemes. However, unless subjects or areas have been randomly assigned to different levels or types of programme input, it would be preferable to designate groups used in analysis as "comparison" rather than as control or experimental groups. In studies of the impact of types or levels of programme impact on fertility change, to the extent possible, areal subgroups should coincide with official administrative (statistical) units;

3. Evaluators should determine and report range of error within which estimates of programme impact lie. Moreover, the estimates themselves should be related in terms of order of magnitude and not as precise numbers, which are less realistic. This requirement applies also to estimates of potential fertility. In addition, high priority should be given in evaluation studies to determining the source and extent of errors;

4. The original formula for estimating couple-years of protection did not allow for a demographically reliable estimate of births averted, a necessary part of the assessment of family planning programme impact. If this method is applied, it should be modified so that when couple-years of protection have been estimated, births averted can be determined;

5. More work should be undertaken in development of analysis of the reproductive process as an approach to determining programme impact on fertility. Accordingly, there should be empirical measurement of its key parameters, and the methodology should be integrated theoretically with the more general issue of measuring fertility change;

6. While simulation is not a method of evaluation of the impact of family planning programmes on fertility, simulation models are a powerful tool for investigating the methodology of various approaches to determining programme impact on fertility, and, for this purpose, greater use should be made of them.

C. Family planning programme activities and administrative arrangements

There has recently been a growing interest among researchers in the importance of the quantity as well as quality of family planning programme input as a variable in fertility change. In light of this interest and owing to the need for national programmes to place greater emphasis upon evaluation, the Expert Group made the following recommendations:

1. The evaluation of the impact of family planning programmes on fertility should be undertaken by evaluation units and these units should be viewed as complementary to the family planning effort and hence given high priority;

2. Inasmuch as the impact of a family planning programme depends at least in part upon the nature and intensity of the family planning efforts, or inputs, it is necessary to have a valid method for classifying and measuring such efforts. Accordingly, research should be undertaken to classify and measure the types and intensity of family planning efforts, including specifically studies of the feasibility of utilizing some measure, such as number or distribution of service personnel by type of service and time spent on family planning, as an index of programme input. The measure adopted should be suited for comparison between areas of countries and/or between countries. Researchers should address the general problem of developing a yardstick suited for use within a country and invariant to time and areas;

3. Efforts should be made to improve methods of obtaining information on quantities of contraceptive supplies to clients, especially pills and condoms, to permit greater precision in determining the numbers of couples or individuals using these methods.

D. Special needs for international training and research

The Expert Group considered the questions given below to be the more crucial unsolved problems remaining in evaluation research. Because of its unique position, the United Nations was considered to have greater opportunities than other bodies to contribute solutions. The Group made the following recommendations:

1. In view of the intrinsic merit of evaluation in the improvement of programmes and the methodological value involved, the United Nations is urged to promote efforts in training personnel from various countries in the evaluation of such programmes, including assessment of births averted by the programme. The United Nations is also urged to assist in the production of training materials required for the purpose. In this connexion, the United Nations should develop a manual utilizing aspects of the results of this Expert Group meeting and other existing knowledge, classifying the methods, specifying when to use them and illustrating how to use them, and indicating the data needed for each method of assessing the demographic impact of family planning programmes on fertility;

2. The United Nations should develop additional methodological case studies in which the population studied should be subdivided and examined in parallel fashion by areas classified by the degree of family planning programme input, to determine whether the input relates to fertility and its change. Such studies should be verified by repetition;

3. The United Nations should, within a reasonably short period of time, select the most critical unsolved

issues emerging in this meeting and invite a small group of experts to consider them and the methodology for solving them. The United Nations should follow its usual procedure in exploring the issues in some suitable manner to provide documentation for the proposed group of experts;

4. The World Health Organization should be encouraged and supported to continue expediting the collection and dissemination of data on family planning in health services, including the inputs provided through those services to family planning activities.

ANNEXES

Annex I

MEETING TOPICS AND DISCUSSION LEADERS

Topic	*Discussion leader*
Methods of measuring family planning programme impact on fertility	C. Chandrasekaran
Major problems encountered in the case studies: presentation of the case studies	
India	K. Srinivasan
Chile	Erica Taucher
Tunisia	Yolande Jemai
The major methodological issue:	
(*a*) Potential fertility	Robert G. Potter
(*b*) Data requirement problems	K. Srinivasan
(*c*) Correlated variables and interaction	
(*d*) Uncontrolled variables	Albert I. Hermalin
(*e*) Independence of method	
Comparison of measures: problems in interpreting validity of evaluation results	William Brass
Problems in determining appropriate methods of evaluation	Ismail Sirageldin
Needed data and research	W. Parker Mauldin
Recommendations	C. Chandrasekaran
Adoption of the report	Jeanne Clare Ridley

Annex II

LIST OF DOCUMENTS

Document No.	*Title and author*
	BACKGROUND PAPERS
ESA/P/AC.7/1	Methods of measuring the impact of family planning programmes on fertility: problems and issues United Nations Secretariat; revised by Albert I. Hermalin
ESA/P/AC.7/2	Methods of measuring the impact of family planning programmes on fertility: the case of Karnataka State, India K. Srinivasan
ESA/P/AC.7/3	Methods of measuring the impact of family planning programmes on fertility: the case of Tunisia Yolande Jemai and Hedi Jemai
ESA/P/AC.7/4	Methods of measuring the impact of family planning programmes on fertility: the case of Chile Erica Taucher and Albino Bocaz
ESA/P/AC.7/5	Provisional agenda and annotations

CONFERENCE ROOM PAPERS

Paper No.	*Title and author*
1	Officers and list of participants
2	Schedule of meetings
3	Statement on the validity of evaluation results of family planning programmes, the problems of determining appropriate methods of evaluation and needed data and research Erica Taucher
4	Some issues in determining appropriate methods of evaluating the fertility impact of family planning programmes Ismail Sirageldin
5	A note on two applications of micro-simulation models to the problems of evaluating family planning programmes Jeanne Clare Ridley
6	Component projection *versus* other techniques for assessing programme achievement towards a targeted fertility reduction Robert G. Potter
7	Interaction of socio-economic changes with family planning programmes: an assessment model K. Srinivasan
8	Needed data and research on the impact of family planning programmes on fertility G. Edward Ebanks
9	Notes on causal relationships in measuring fertility change H. Bradley Wells
10	Avoiding an embarrassment of riches Albert I. Hermalin
11	Some aspects of determining appropriate methods of evaluation for measuring the impact of family planning programmes on fertility in the Republic of Korea Suk Kap Koh
12	Needed research for measuring the impact of family planning programmes on fertility W. Parker Mauldin
13	Measuring the impact of family planning programmes on fertility Chen-tung Chang
14	Comments on comparison strategies for the evaluation of family planning impact William Brass
15	Problèmes posés par les données nécessaires à l'évaluation de l'impact des programmes de planning familial sur la fécondité Yolande Jemai
16	List of proposed recommendations
17	Minimum data needs by method of analysis

Part Three

STATEMENTS ON METHODS OF MEASURING THE IMPACT OF FAMILY PLANNING PROGRAMMES ON FERTILITY

Submitted by members of the Expert Group

COMMENTS ON COMPARISON STRATEGIES FOR THE EVALUATION OF FAMILY PLANNING IMPACT

*William Brass**

The background paper, "Methods of measuring the impact of family planning programmes on fertility: problems and issues" (ESA/P/AC.7/1), systematically reviews the area of study. The following notes are of particular interest in relation to recent applied research. No attempt is made to examine the topics in detail; rather, points are raised for discussion.

Perhaps the most significant element in the development of family planning evaluation over the past few years has been a change of attitude. In the earlier periods, there was a strong tendency for the subject to be treated in isolation as a well-defined subdiscipline. Currently, to an increasing extent, it is seen as an integral part of fertility analysis. Thus, family planning is one of a variety of factors underlying childbearing behaviour at different levels of "depth". The explanation of the part it plays cannot be separated from the interrelation with the other influences. The consequences of this change of attitude have not yet been fully realized. Among them are the need to move the focus from the programme itself towards the broader determinants of fertility, which in turn has a whole series of implications about the resources which are justified for information gathering about individual programmes and the nature of the data required.

Research stategies are also affected if the aim is to measure the contribution that a family planning programme is making to fertility change individually or, more plausibly, synergistically with other variables. In particular, there is the necessity for comparative materials for a range of situations in which these variables enter either explicitly or implicitly in combination with different degrees and/or types of family planning effects. This need sets the subject within the broad conceptual and methodological field of observational interpretation which has a much longer history of study within other disciplines, for example, epidemiology. The principles and techniques that have evolved in these disciplines can be translated and adapted for family planning evaluation.

By their nature, observations of an uncontrolled process cannot provide the same kind of certainty about cause-and-effect relationships as can well-designed experiments. Quasi-experimental approaches, for example, the use of comparative groups but without treatment randomization, can never fully eliminate the unknown factor correlations which make invalid quantitative specifications of belief in terms of probabilities. Subjective elements must enter, which makes the investigation a search for plausibility. However, there is sufficient agreement about the kind of evidence that establishes plausibility to most informed observers for a consensus to be reached.

Internal comparisons of fertility changes in appropriate subgroups of a population are much more powerful than assessment in terms of theoretical or model comparisons and the returns from the latter are weak and limited. The reason is the great variability in reproductive determinants and behaviour over populations. Models are a form of simple averaging and their use as standards against which observed measures are assessed must ignore many influences which may be critically relevant. Of course, models can sometimes contribute towards an essential adjustment of direct measures which are not immediately comparable, but reliance on such uses must be very tentative.

Within-population comparisons have the immense advantage that an important part of the determining forces will be common to all subgroups, permitting the remainder, including family planning programme effects, to be defined more precisely. (Of course, inter-population studies are needed to illuminate the common factors.) Several broad types of comparative design are possible but some principles are applicable to all. One such principle is the need for replication, that is, the repetition of the examination of measure differences related to a characteristic, in a variety of circumstances. Another is the importance of timing since influences are directional; conclusions must be governed by considerations of the order in which changes occurred and not only by their size. Although there is little possibility of strict randomization because control over systematic effects is impracticable, it may be useful to include some randomization to guard against the introduction of subjective biases.

These points are general but they will be illustrated by reference to particular types of comparative investigation. Multiple replication can be achieved by the relation of measures of fertility change in subareas of a country or region to indexes of family planning effort, use or effect, and other putative determinants normally of a socio-economic nature. But the replication is often much less than would appear because of strong dependence of the family planning indexes upon the socio-economic indexes. In addition, it is difficult to

* Professor of Medical Demography, University of London; and Director, Centre for Overseas Population Studies, London School of Hygiene and Tropical Medicine, London, England.

take satisfactory account of the underlying relationships because of restrictions on the data available (frequently from censuses) and the complexity of the possible lag intervals. The results of such multivariate analyses are essentially superficial as there is not the detail of outcome and characteristics to permit the closer linking which could throw light on how the measurement of effects is distorted by dependence.

Some improvement may be gained by quasi-experimental approaches in which comparative design classifies in the same strata subareas "homogeneous" with respect to fertility determinants. Within one stratum or several strata, family planning programme differences are established preferably, but not necessarily, subsequent to the stratification. In the extreme, but most usual, situation there may be one stratum and two subgroups; that is, family planning activity is concentrated in one area and the other serves as a control. For example, in an experimental programme in the Matlab region of Bangladesh, the family planning area is being "saturated", household by household, with supplies of contraceptives. The degree of replication in this design is very thin for subareas and for strata. Both are important, but for slightly different reasons than those operating in the classical randomized-block experiment. The provision of many comparisons in varied situations guards against the random effects both of external influences and of possible dependence due to the "contamination" of the control by the operational areas, because such dependence is likely to be variable.

Quasi-experiments can be assessed by the same kinds of general summary indexes as used in the multivariate analysis of observations for subareas (birth or fertility rates, percentage of family planning acceptors etc.), but more detailed and sophisticated measures are preferable. Replication, in fact, need not be done by complete subpopulations in separate areas, but can also be through particular groups by age, ethnic origin, social class, education and so on. In other words, the need is for a series of comparisons of the effects of a factor in different contexts with evaluation based on the consistency of patterns. But the nature of the reasoning about the impact of family planning must now be altered because the biosocial subgroups cannot be expected to respond in the same way to other influences (apart from random or more properly erratic distortions) which is the hypothesis for designated and control subareas. The criterion must be the agreement of the differentials by the family planning indexes with a coherent set of concepts of how a genuine impact would operate, as opposed to a situation where the programme use was a substitute for action to achieve the same ends by other means. Of course, there is certain to be a subjective element in any such set of concepts and ideas are not static, but a broadly acceptable view appears to be attainable. Thus, responses of subgroups (by age, education, type of community, religion and so on) to a family planning programme to an extent that appears to reflect the allocation of effort rather than the social classification can be taken as valid evidence of impact. The quantitative measurement must, however, be extremely uncertain.

The greater the detail of significant classification, the more incisive is the evaluation since it expands the opportunities for the effects of uncontrolled factors to be registered. But detail has another important advantage. "Nuisance" influences can be eliminated and the sensitivity of the comparisons increased by concentration on the refined measurements which will show the relevant indications more strongly. This detail may, in some cases, include basic biological features which can interact with family planning adoption. Thus, in Nigeria, survey studies are being made of "modernizing" and traditional groups with the objective of separating out the effects of family planning and biosocial factors, particularly breast-feeding practice, on fertility change. The component of the birth interval associated with level of breast-feeding is so large that modifications in practice can interfere greatly with the attempt to examine other influences.

The logic of the evaluation by detailed comparison of measures in subpopulations differently affected by a family planning programme can be extended to the situation where the designated group consists of the entrants to the programme itself. The controls are those who have not entered (or a representative sample) but who are subject to the other influences towards lower fertility and who have some opportunity to respond. Of course, it is always possible to argue that entrance to the programme constitutes such a powerful selective process that the bias is overwhelming. The argument has great force in some situations but much less in others. The most favourable case is where a substantial proportion of the potential clients has entered, build-up is rapid, there is considerable social diversity and access to contraception outside the programme is relatively easy. It is also necessary to assess the change in fertility of all entrants to the programme (or a sample) regardless of whether they have continued to be clients, since, otherwise, the biases are insuperable.

The measurement of fertility decline of the entrants and of the controls on a consistent basis raises difficult problems, even if it is assumed that behavioural biases are sufficiently modest. This difficulty arises because the entrants are selected for past fertility in complex ways. They are not pregnant and differ from the average in exposure to risk, proportions sterile, intervals to last birth, births in previous periods and fecundity. Much study has been given to the problems of adjustment for these selection biases, usually under the topic heading of "births averted". The results are theoretically not too satisfactory, but the range of error is often small enough to be unimportant in the conditions where impact can reasonably be evaluated. The major difficulties are two. Some effects can be calculated quite precisely but only if there is sufficient information, e.g., on the times of last pregnancy of entrants and also on the biological components of the birth intervals (post-partum amenorrhoea, delay to conception etc.). Others depend upon characteristics that cannot be measured directly, in particular, fecundity and sterility. Only the outcome in fertility can be observed and it includes a large stochastic element which is part of the selection bias.

The adjustments for selection biases have often been based on model constructs of the reproductive process. However, there are increasing doubts about the regularity of the measures in different populations, in particular, the relation of post-partum amenorrhoea to lactation and cultural factors and the pattern of sterility with age (and recognition of sterility) in different nutritional and social conditions. The need for comparisons within the populations becomes more and more evident. It is here that the chance elements in fertility are particularly troublesome because the measurements reflect to an unknown extent the persistent (which is relevant to the comparisons) and the transient (which is not). For example, women aged, say, 35 years enter a programme with, in general, a higher parity, greater fertility in the past five years and shorter interval to the last birth than the average for married women of the same age in the population. Part of the difference is due to sterility (which will continue), part to risk exposure and fecundity (which can change) and part to chance (where the expected future excess is zero). Barrett and Brass[1] have shown that the chance effects are large and a recent calculation from historical data shows a very moderate correlation of about half between fecundity in two successive birth intervals. Thus, the attractive idea of establishing the fertility potential of the entrants by their past performance has major snags.

Nevertheless, the advantages are so considerable, in terms of the minimization of model elements in the measurements, that more study of procedures for reducing the stochastic biases is indicated. The most promising approach is through the specification of the most effective controls, either in reality or conceptually. The stochastic biases can be eliminated from the comparisons by appropriate matching of controls to entrants so that both groups have the same chance distortion. Presumably the right criteria for matching depend upon recent birth intervals but more research is needed on the best choice and its sensitivity. If the control matching is conceptual, that is, a computer simulation of the stochastic biases in a population with the "persistent" characteristics of the entrants, further problems arise of the proper specification of features, such as fecundity variation. Brass and Barrett produced a conceptual matching for the evaluation of the data from the Mauritius family planning programme and have recently developed more general results.

---

[1] J. C. Barrett and W. Brass, "Systematic and chance components in fertility measurement", *Population Studies*, vol. XXVIII (November 1974), pp. 473–493.

MEASURING THE IMPACT OF FAMILY PLANNING PROGRAMMES ON FERTILITY

*Chen-tung Chang**

In measuring the impact of a family planning programme on fertility, it is generally assumed that a birth averted is identical with every other birth. In a sense, this assumption is indisputable; if the same number of births has been avoided over the same period of time, the same fertility impact has been achieved. However, if one wants to go beyond the mere statement of the number of births averted and try to relate it to prospective population growth for planning purposes, the untenability of the assumption becomes all too apparent. The same problem also arises when one attempts to compare the fertility impacts of two or more family planning programmes. Is a birth averted in a programme equivalent to a birth avoided in another? The answer hinges upon the question whether the two births have been averted for the same reason.

For the present purposes, the most relevant way to classify reasons for family planning is into spacing and limiting. To have a proper assessment of the fertility impact of a programme, it is important to ascertain whether the acceptors have come for purpose of delaying or stopping childbearing. The same number of births averted over the same given period of time, especially when the period is short, as when a programme is evaluated on an annual basis, may have substantially different implications for prospective fertility if the predominant reason for family planning differs between acceptors of two programmes. For child-spacing, by definition, implies merely a temporal redistribution of live births over a period of several years. Its impact on fertility, as far as the quantity dimension is concerned, is therefore of only a temporary nature. On the other hand, births averted because of the decision to limit family size may be avoided forever.

It will be observed from table 1 that the practice of child-spacing is very popular in Singapore, with a substantially high proportion of delayed births among all births averted in a given year.

There is no question about the fertility impact of the national family planning programme in Singapore. The national programme began in 1966, about a decade after the inception of the transition from high to low fertility. However, evidence strongly indicates that the programme has significantly accentuated the downward trend of fertility. Especially in the first few years, actual fertility was much lower than expected fertility, however it is estimated. The question is that despite the continuing recruitment of new acceptors, the fertility reduction decelerated, and there was actually a slight increase in the total fertility rate in 1972.

TABLE 1. PERCENTAGE OF ACCEPTORS [a] IN SINGAPORE PRACTISING CHILD-SPACING, BY NUMBER OF LIVING CHILDREN, MID-1967 – END-1970

Number of living children	*Acceptors [a] practising child-spacing*
Zero–one	86.8
Two	84.8
Three	61.8
Four	29.4
Five or more	11.0
TOTAL	60.0

[a] Excluding acceptors who are undecided or whose reason for family planning is not known; the proportion of such cases is small in both programmes. Including acceptors of all methods, 50.5 per cent of whom are pill acceptors.

The most likely explanation of this pattern of fertility change is that the initial impact of the programme resulted in the recruitment of large numbers of women practising birth control for the purpose of spacing. Because many of these delayed births appeared several years later at a time when there was a decline in the number of new acceptors, there was an upsurge in births and hence the stagnation of the downward fertility trend and even a slight upward turn.

The availability of birth-order data permits a more careful examination of this process. It will be observed from table 2 that in 1971 and again in 1972, whereas birth rates for births of the fifth order and above continued to decline, lower-order birth rates showed increases for many age groups of women. This pattern of changes indicates that as the reductions of higher order birth rates reflected the intention to limit family size, the increasing practice of birth control by high-parity women continued to result in the reduction of the birth rates. On the other hand, much of the reduction of lower order birth rates was due to the wide practice of child-spacing in response to the inception of the national family planning programme in 1966. As postponement of childbearing is usually only for a short period, many of the births delayed in the first few years were appearing several years afterwards. As a consequence, despite the continued spread of the practice of birth control, in 1971–1972, increases were observed in many of the lower order birth rates.

* Lecturer, Department of Sociology, University of Singapore.

TABLE 2. SINGAPORE: BIRTH RATES BY AGE GROUP AND BIRTH ORDER, 1970-1972

Year and age group	*Age-specific birth rate*	*Birth order*					
		First	*Second*	*Third*	*Fourth*	*Fifth*	*Sixth and over*
1970							
15-19	26.1	18.6	6.0	11.3	00.2	0.0	0.0
20-24	139.1	61.8	43.0	21.5	8.6	2.6	1.1
25-29	208.8	49.0	53.8	42.1	27.9	17.5	17.7
30-34	138.0	14.3	23.7	25.0	20.2	16.1	38.2
35-39	74.6	3.6	5.1	6.8	7.2	6.8	44.9
40-44	26.7	0.6	0.7	1.3	1.5	1.9	20.8
45-49	4.7	0.1	...	0.2	0.2	0.2	4.1
TOTAL	618.0	148.0	131.3	98.2	65.8	45.1	126.8
1971							
15-19	25.9	18.8	5.8	1.1	0.1	0.0	0.0
20-24	143.4	65.6	44.9	21.3	7.9	2.2	0.8
25-29	191.6	49.1	52.3	39.3	24.6	13.4	11.9
30-34	142.6	15.5	26.3	29.6	21.9	15.9	32.6
35-39	67.9	3.1	6.0	8.4	8.0	6.7	35.3
40-44	22.4	0.6	0.9	1.1	1.4	1.4	17.0
45-49	3.4	0.1	0.1	0.1	0.1	0.2	2.8
TOTAL	597.2	152.8	136.3	100.9	64.0	39.8	100.4
1972							
15-19	26.1	19.7	5.2	1.0	0.2	...	0.0
20-24	140.4	65.9	43.8	20.2	9.1	2.1	0.6
25-29	200.6	55.8	56.2	41.3	23.8	12.7	9.8
30-34	148.0	15.4	28.3	31.6	23.2	17.4	31.4
35-39	64.0	3.6	6.4	8.2	7.9	7.0	30.5
40-44	21.2	0.7	0.9	1.3	1.3	1.5	15.4
45-49	2.9	0.1	0.1	0.1	0.1	0.2	2.3
TOTAL	603.2	161.2	140.9	103.7	63.6	40.9	90.0

Note: A few births of unknown age of mother or of unknown birth order were excluded in the computation of birth rates. The underscored figures are those which show an increase over the previous year.

It is interesting to note in this connexion that in a study[1] on the fertility impact of a family planning programme in one country, a comparison was made between the number of births averted by the programme estimated from use of programme methods and that from differences between expected and actual fertility over the period 1965–1971. The expected fertility was derived from the fertility trend in 1959–1963 because 1963 was the year when the programme gained momentum through expansion. The comparison shows that the calculation, which was based on the difference in expected and actual age-specific marital fertility, over-estimates the number of births averted for the period prior to 1967 and underestimates the number after 1967. The author's interpretation is that in the absence of the programme, the expected fertility trend after 1963 should be curvilinear, with a higher rate of decline at the beginning and a lower rate at the end, so that the difference between expected and actual fertility would be smaller at the beginning and larger at the end. An alternative interpretation, which appears to be more plausible, is that the reduction of fertility was due in part to the practice of child-spacing, and the births originally to have appeared in the early years were postponed to the later years. In other words, a phenomenon occurred which is similar to what has happened in Singapore.

The two examples serve to make a point that if evaluation is conducted as a routine operation on an annual basis, the results must be interpreted carefully for their fertility implications. Otherwise, the elation at the initial impact of the programme or the introduction of new schemes may soon be followed by alarm when the downward trend of fertility is arrested, which, when misinterpreted, may in turn lead to the development of unnecessarily drastic measures. The problem arises from the failure to make a proper distinction between the short-term and long-term effects of a programme. Strictly speaking, the practice of child-spacing also has its long term effect, inasmuch as it affects the tempo dimension of fertility. For all practical purposes, however, its annual implication is negligible when the effect is dispersed thinly over the length

[1] T. H. Sun, "The impact of fertility of a family planning program", in C. Chandrasekaran and Albert I. Hermalin, eds., *Measuring the Effect of Family Planning Programs on Fertility* (Liège, International Union for the Scientific Study of Population for the Development Centre of the Organization for Economic Co-operation and Development, 1975), pp. 427–504.

of a generation.[2] The fertility impact along the quantity dimension is, therefore, the primary concern here.

The proper assessment of programme impact, with due recognition of the distinction between short-term and long-term effects, can be approached if the number of couple-years of protection (CYP), for example, is measured for each parity group of acceptors specifically. The number of births averted is thus estimated together with their birth-order composition. Since higher order births are averted most probably because of child-limiting, they can be given more weight than those of lower order in assessing the "permanent" fertility impact of the programme. In other words, provided with information not only on the total number of births averted but on their birth-order composition, it is possible to have a more "realistic" interpretation, on a short-term basis, of the impact of a programme on the fertility of a population.

[2] Nevertheless, ". . . restricting the measurement of program impact to change in *unwanted* fertility would exclude the demographic and other impact of *spacing*, the improvement of which may turn out to be one of the principal contributions of U.S. family planning programs." Frederick S. Jaffe, "Issues in the demographic evaluation of domestic family planning programs", in J. R. Udry and E. E. Huyck, eds., *The Demographic Evaluation of Domestic Family Planning Programs* (Cambridge, Mass., Ballinger Publishing Co., 1975), pp. 19–30.

The same applies to the fertility projection approach. If in comparing the actual with the expected fertility both are examined in more detail by birth order, one can then gain some idea of the proportion of their difference that is due to changes in fertility of higher parity women. Again, since the effects on higher order and lower order fertility are assessed separately, one can maintain a more balanced perspective on the over-all impact and be better guarded against the misinterpretation of the fertility implications of programme impact.

Obviously, such a more differentiated approach to assessing fertility impact is especially necessary for programmes where child-spacing is popular among acceptors or where the proportion of acceptors practicing spacing changes significantly because of frequent programme changes or policy revisions. In Singapore, for example, child-spacing has been popular from the beginning of the national programme. Moreover, in the six years from 1967 to 1973, the proportion of acceptors practising spacing increased from 53 per cent to 84 per cent, while the proportion with one child or none at all more than tripled, rising from 21 per cent to 67 per cent. For a programme like that in Singapore, it is necessry to know more about the births averted, especially the birth-order composition, in order to interpret the fertility impact of the programme on a firmer basis.

NEEDED DATA AND RESEARCH ON THE IMPACT OF FAMILY PLANNING PROGRAMMES ON FERTILITY

*G. Edward Ebanks**

There is no necessity to labour the point concerning the need for data and scientific research in the area of evaluating the impact of family planning programmes on fertility. In this short paper, the following assumptions are made:

(*a*) There is, or there is about to be launched, a family planning programme, national or otherwise;

(*b*) Because of emphasis, and rightly so, upon delivery of services, data collection within the everyday operation of the programme is minimal, e.g., age, marital status, parity, living children, occupation, education and religion;

(*c*) There is a separate research and evaluation unit with adequate funds for conducting research;

(*d*) Use of an electronic computer is available;

(*e*) There is no need to make assumptions about trends in fertility coinciding with or following upon the initiation of the programme;

(*f*) Because of the nature of the programme within a single national unit, it is not possible to set up an experiment with experimental and control groups. The programme, whether national or subnational, is assumed to contaminate to greater and lesser extents the subunits of the national population.

Since using service statistics as sufficient for evaluating the impact of the programme on fertility is implicitly ruled out and since a true scientific experiment is also ruled out, one is left with two appropriate research possibilities. The first is use of sample survey techniques and the second is an extension of this method to panel studies. Also ruled out is the matching of clients and non-clients because of the many problems associated with that type of design.

Sample surveys, if well planned and executed, have the potential to yield the type of data necessary for an evaluation of the impact of a family planning programme on fertility. The evaluation of the effectiveness of a family planning programme in terms of its impact on fertility is best done within a broader framework, that of establishing the determinants of fertility within a specific socio-cultural system. In other words, a study directed towards assessing the impact of a programme on fertility can be extended to an examination of the determinants of fertility, another area in which there is a need for data and research. One can, using the Blake-Davis[1] formulation of 11 intermediate variables and Freedman's[2] extension of it, establish the links between the institutional variables and the intermediate variables, and between those variables and fertility. One intermediate variable, use or non-use of contraception, may take on significance all the way from just a minor one of the 11 variables up to its being the single most important determinant of fertility and hence making nought of most of the other 10. The family planning impact at this extreme can then be seen as an institutional variable influencing the intermediate variables. Properly designed and executed sample surveys of the types presented below can assess the intermediate variables, as well as the socio-cultural variables operating through them to influence fertility.

Approach 1: Periodic field surveys

A periodic field survey should be conducted about every five years. Assuming that each of these periodic sample surveys is intended to cover a period of evaluation of, say, 5, 10, 15 or 20 years, the initial survey may try to assess the impact of the programme over the past period of, say, anywhere from 5 to 20 years. The longer the time period to be covered, the more difficult is the problem, the larger the sample required and the greater the problems of reliability and validity. By way of an example, one may assume that the initial survey is directed towards evaluating the impact of the family planning programme over the first 10 years or the most recent 10 years. The appropriate age range of women to be included in the sample is 15–59 years, or should, in the present author's view, be that range. At the beginning of the period, the women currently aged 59 were 49, the upper limit of the childbearing ages. These women and those aged 45–48 years are included in order to have completed fertility for a five-year age cohort at the initiation of the programme. The 15–19 age group was aged 5–9 years at the initiation of the programme. They are now included because they have just entered the childbearing ages. By at first including the age range 15–59, one has women aged 15–49 for all the 10 years for which the evaluation is being conducted.

* Associate Professor, Department of Sociology, University of Western Ontario, London, Ontario, Canada.

[1] Kingsley Davis and Judith Blake, "Social structure and fertility: an analytical framework", *Economic Development and Cultural Change*, vol. 4 (April 1956), pp. 211–235.

[2] Ronald Freedman, "The sociology of human fertility", *Current Sociology*, vol. X-XI, No. 2 (1961–1962), pp. 35–121.

The aim here is to be able to calculate age-specific fertility rates for each of the 10 years under consideration. This calculation should be done for single years of age, which can then be aggregated into five-year age groupings for each of the 10 years under consideration. The calculation of age-specific fertility rates is seen as a necessity for an evaluation of the impact of a family planning programme on fertility.

If the evaluation is to cover a period of five years, one would select into the sample women aged 15–54 years, which would give an age range of 10–49 at the beginning of the period. Similarly, if it is for a period of 15 years, the sample could still include women aged 15–59 years, which would give at the beginning of the period an age range of 0–44, thus missing the 45–49 age group at the beginning of the investigation, an omission which could be accommodated. If it is for 20 years and one includes women aged 15–59 years at the beginning of the period, one would have women aged 0–40 years, thus missing only the two last five-year age groups within the childbearing ages.

Since the objective is the calculation of age-specific fertility rates, one must be concerned with the number of women at each age over the period of investigation. For argument's sake, one may say that 100 women are required for each single year of age. Because one is including ages 15–59, that is, 45 single years of age, 4,500 women would be required for the sample. However, one would at least want to compare ever-users of contraceptives with never-users, and hence twice as many women, say 9,000, would be needed. If one wants to compare the age-specific fertility of clients of the programme, non-clients but users of contraceptives and non-users, one should again increase the sample size by a significant number, even if not by another 4,500. If one wants to subdivide further the sample units into, say, active clients, inactive clients, active users who are non-clients, inactive ever-users who are non-clients, and never-users, the sample size should be even larger. To cover adequately all these subgroups, one would need a very large sample, which then becomes expensive. However, research is not cheap and this type of evaluative research, if it is to accomplish its goal, is even more expensive.

A problem not yet discussed which is related to size of the sample is one of identifying the subsamples mentioned above, at each point in time. That is, in the case of the 10-year evaluation, one needs to be able to identify the subgroups in each of the 10 years, so that a woman who is a client in year x may have been a non-client in year $x - 1$, and an inactive client in year $x + 1$. The recording of the retrospective data should afford this opportunity and this factor should be considered when deciding on the necessary sample size.

A national probability sample is proposed because in so doing, it is hoped that the other independent variables will be randomly distributed, making it less necessary to control for them and hence further necessitating large sample sizes. On the other hand, a national probability sample must consider the proportion of women in the childbearing ages who are members of the family planning programme, since this will affect the number of possible clients included in the random sample.

The minimal amount of desirable data is a contraceptive history, a pregnancy history and sexual union history. Other useful histories would be a work history and a migration history. The questionnaire should be designed in such a way as to make integration of these histories relatively easy and also to facilitate the identification each year of the various subgroups of interest, so that age-specific fertility rates may be calculated for each of them.

Repeated at about five-year intervals, this research procedure would, in addition to evaluating the effectiveness of the family planning programme, allow the researchers to examine changes in the intermediate variables as well as their societal determinants and hence to provide data on the determinants of fertility.

Approach 2: the panel study

Approach 2, the panel study, is seen as not very different from approach 1 and hence is an alternative to it. The data to be collected are the same. The aim is the same: evaluation of the effectiveness of the family planning programme through its impact on fertility. Once again, it is done by calculating annually the age-specific fertility rates for single years of age, for some or all of the subgroups identified under approach 1.

One should begin with a national sample covering women in the age interval 15–54. In so doing, one has at first two five-year age cohorts for which one can calculate completed fertility. Data obtained on this group should be similar to those obtained in approach 1. This initial national survey is the base-line study and should preferably coincide with the launching of the programme. However, if the programme is already in existence, it would mark the beginning of the panel study and would retrospectively look at fertility and contraception, among other things. Although covering the above-mentioned age range, there should be oversampling of the 15–19 age group, as those women should become the cohort which will be followed in order to record their social, economic, demographic and contraceptive statuses and changes as they go through the childbearing ages.

These women aged 15–19 years should be followed for 30 years, being reinterviewed every two years, and hence interviewed a total of 15 times. It is proposed that with the two-year interview interval one would include each time women aged 15 years and over. So at the second interview, that is, the interview following the base-line study, one would interview the panel members, who would currently be 17–21 years of age, and also a sample of women aged 15 and 16 years. At the third interview, the original panel members would be 19–23; and one would reinterview them, plus a sample of the women aged 15–18 and so on, until the fifteenth interview, 30 years from the beginning when the original panel members would be 45–49 years of age and one would include an additional sample of women aged 15–44. In so doing, one would have the benefit of a panel study and also each time of being able to calculate age-specific fertility rates for the 15-year to highest age among the original panel members.

A variation on this technique would be to keep adding each time the younger age groups as additional panel members rather than as described above, where there is an independent sample each time except for the original 15–19 age cohort of panel members. Another variation is to take every two years an independent sample rather than following the panel members. That is, at the second interview or the one after the base-line study, a sample of women aged 15–21 years is drawn and interviewed. On the next round, two years later, another independent sample is chosen of women aged 15–23. This approach enjoys much of the benefits of the first panel type described and avoids the high attrition rates of panel studies as well as the sensitization effect. It is also less costly and more manageable.

Regardless of the format of the panel study, it has several advantages. By following the panel members and interviewing them at regular intervals, one is more likely to obtain reliable information. In all cases of this type of research, validity is a major problem. In following the panel, one can assess the effect of the programme on a group of women passing through the childbearing ages. Both the direct and indirect effects of the programme are measurable within this approach.

Summary and conclusion

Both approaches briefly described above have much in common, although they differ in certain important aspects. In both approaches, one must be concerned with problems of reliability and validity. Both require the same type of data but get at it in slightly different ways. They are both costly and time-consuming. But in the present author's opinion, there is no cheap or easy way of evaluating the impact of family planning programmes on fertility. In addition to evaluating the family planning programme, the approaches described above provide micro data for the study of the determinants of fertility. While this study is going on, the macro changes in the society could be carefully observed and recorded so that macro analyses relating to the determinants of fertility could be undertaken. Combining the micro and macro approaches would provide all the necessary data for whatever evaluation techniques may be desired. One would then have data for technical analyses as well as data for the policy-makers and programme personnel.

Both approaches have the potential for providing data on such areas as voluntary childlessness, infecundity, and changing ideals, norms and values as they relate to childbearing and family sizes. Family planning evaluation is best viewed within the broader framework of the determinants of fertility and these approaches afford the opportunity to do so.

In both approaches, as the aim is the calculation of annual age-specific fertility rates for various subgroups, sample sizes must be large and the possibility of interrelating the various data components must be built in. Only brief mention was made of the types of data required and the techniques for obtaining them because of the limited space and because examples of these needs and methods exist in the literature. The space limitation contributed to the lack of clarity on some of the points made and also to the omission of some of the underlying assumptions.

AVOIDING AN EMBARRASSMENT OF RICHES

*Albert I. Hermalin**

It is a well-known reflex among researchers when faced with limitations of inference from a body of data, to call for more data. Though the limitations encountered may be real enough, less clarity usually attaches to the way in which the additional data will relieve the problem. Analysts of cross-sectional data, for example, often send out a plea for panel or longitudinal data but usually without a detailed statement of the structure of the new analysis which will incorporate the desired data. Those concerned with measuring the effect of a family planning programme have been in a similar position, often restricted in what can be said from a single method applied to a single instance and wishing for multiple approaches in a variety of settings. Thanks to the efforts of the United Nations, these wishes have been granted in part, and it is to be hoped that further material of this nature will soon be forthcoming. If one is to avoid an embarrassment of riches, however, the Expert Group meeting might well devote some attention to how best to analyze this increased data base. The purpose of this brief note is to suggest some directions this discussion might take without at all pretending to resolve the matter. More specifically, the aim will be to review some developments about reliability and validity from the socio-psychological literature which appear isomorphic to the concerns of this meeting as a first step in appraising their potential utility. It is assumed in the following discussion that the type of data available to this meeting will be enlarged to more trials so that it makes sense to talk about the correlations over settings of variables emerging from the application of various methods for evaluating the effect of family planning programmes.

Elementary concepts

On the most general level, reliability refers to the consistency of a method or test, while validity relates to the ability of a method to measure what it intends to measure. In the "classical" approach, reliability is established through "the agreement between two efforts to measure the same trait through maximally similar methods", while validity is seen as "the agreement between two attempts to measure the same trait through maximally different methods".[1] It is recognized, of course, that independence of methods is a matter of degree and an alternative formulation of this distinction will be developed below. Though interest centres on validity, reliability is important because it limits the degree of validity that may be established. It can be shown, for example, that "the validity of a test with respect to any criterion cannot exceed the index of reliability"[2] (where validity is viewed as the correlation of a test score with a second measurement, and reliability is defined as the correlation of a test score and its true score). Attenuation formulae that attempt to estimate the correlation between the true scores of two tests involve the observed score correlation as well as the index of reliability of each test.

Establishing the reliability of methods for evaluating programs

One may thus ask as a first step how the reliability of the various methods for measuring the effect of family planning may be established. Most of the well-known methods, such as test-retest, parallel forms and split-form, have been developed in the field of mental testing. Some of these methods have little applicability to the subject-matter of this meeting, but others may be adaptable to the present concerns. For example, as an analogue to the split-half technique, it should be possible to determine contraceptive continuation rates for random subsamples of acceptors or to carry out areal regression for subsamples of areas, where the number of subdivisions is large enough. If this procedure were done over sufficient instances, the correlation between the measures obtained from each subsample would provide an estimate of reliability.

The parallel-form technique suggests that alternative approaches for obtaining the same measure should be considered. For example, continuation rates may be determined by follow-up of acceptors or by obtaining the necessary information from a KAP survey. Again, the correlation between measures derived from each source would speak to the issue of reliability. One can go further and sort out the various methods for measuring the impact of a programme according to their degree of similarity by analyzing the measures produced as well as the variables employed. One would expect greater correlation among similar measures than among those more dissimilar. Turning

* Director, Population Studies Center, University of Michigan, Ann Arbor, Michigan, United States of America.

[1] Donald T. Campbell and Donald W. Fiske, "Convergent and discriminant validation by the multitrait-multimethod matrix", *Psychological Bulletin*, vol. 56 (March 1959), p. 82.

[2] Frederic M. Lord and Melvin R. Novick, *Statistical Theories of Mental Test Scores* (Reading, Mass., Addison-Wesley, 1968), p. 72.

around the inequality between validity and reliability, mentioned earlier, it is also true "that the index of reliability cannot be less than any validity coefficient of a test".[3] Thus, the degree of correlation between methods deemed similar should provide a lower bound to the reliabilities of the methods in question.

Concepts of validity

Various approaches to validity have been advanced in the literature: content validity; predictive and concurrent validity; construct validity. The opinion here is that construct validity holds the major interest for this meeting, though the other concepts have some relevance. One may note in passing that it makes sense to judge whether the proposed methods of evaluation bear sufficiently on what they are intended to measure. For predictive validity, it is clear that an independent criterion of the effect of a program is not available, for if it were, all these efforts would not be required. Nevertheless, if a programme is evaluated as having a strong effect over a period of years, it is reasonable to expect that this effect will be reflected in an appropriate fertility measure, suitably lagged and adjusted for any demographic confounding effects. This would appear to be a necessary, but clearly not sufficient, condition to establishing the validity of the method in question.

Lord and Novick hold that "for scientific purposes the most important characteristic of a test is its construct validity".[4] Another source states: "Construct validity involves relating a measuring instrument to an overall theoretical framework in order to determine whether the instrument is tied to the concepts and theoretical assumptions that are employed."[5] As such, construct validity is not established in a single step but requires a chain of reasoning in which hypotheses concerning the construct are derived from theory and tested. Campbell and Fiske[6] have advanced the understanding of this process by noting that the validation must be convergent as well as discriminant. That is, the results of a test or method should correlate appreciably with other results that the hypotheses suggest as related (for example, other methods designed to measure the same construct), while there should not be appreciable correlations among variables that theory identifies as relatively independent. In carrying out these steps, one must be alert to spurious correlations which may inflate the correlations observed and lead to incorrect inferences with respect to both the convergent and the discriminant criteria. Spurious correlations may arise because two methods "measure in part something other than the construct of interest".[7] This statement introduces the important concept of method variance as a non-random factor affecting observed measures.

Campbell and Fiske state: "Each test or task employed for measurement purposes is a *trait-method* unit, a union of a particular trait content with measurement procedures not specific to that content. The systematic variance among test scores can be due to responses to the measurement features as well as responses to the trait content."[8] Lastly, one must recall that the reliabilities of each method come into play when assessing the level of correlation between two variables.

The multitrait-multimethod matrix

In order to carry through the ideas on construct validity described in the previous section, Campbell and Fiske state that more than one variable and more than one method must be used. They suggest that to separate the relative contribution to observed scores of trait variance and method variance, it is convenient to set up a multitrait-multimethod matrix in which all of the intercorrelations resulting from the measurement of several traits by each of several methods are displayed.[9]

In the context of this meeting, one might think of a situation where for a reasonably large number of countries, three methods of evaluation have been applied, each of which produces a measure of the impact of the programme on fertility (X), a measure of the effect of a socio-economic variable on fertility (Y), and the effect of some possibly confounding demographic variable, such as age structure or nuptiality (Z). Thus, one has three variables, each measured by three different methods. In this case, the multitrait-multimethod matrix can be visualized as was done by Alwin[10] in his excellent explication of this approach (see figure I).

Carrying on the example suggested, X_1 would be the measure of the effect of the programme obtained by method 1, X_2 by method 2 etc., so that $r_{x_1y_1}$ in figure I would be the correlation between programme effect and socio-economic effect obtained by method 1, while $r_{x_1y_2}$ would be the correlation between programme effect measured by method 1 and socio-economic effect measured by method 2. Figure I indicates two different blocks of correlations:

(1) The monomethod blocks which show the intercorrelations among variables obtained by a single method, the diagonals of which represent the reliabilities for each measure for that method $r_{x_1x_1}$, $r_{y_1y_1}$

[3] *Ibid.*, p. 72.

[4] *Ibid.*, p. 278.

[5] David Nachmias and Chava Nachmias, *Research Methods in the Social Sciences* (New York, St. Martin's Press, 1976). See also L. J. Cronbach and P. E. Meehl, "Construct validity in psychological tests", *Psychological Bulletin*, vol. 52 (July 1955), pp. 281–302.

[6] Loc. cit.

[7] F. M. Lord and M. R. Novick, op. cit., p. 279.

[8] D. T. Campbell and D. W. Fiske, loc. cit., p. 81.

[9] *Ibid.*

[10] Duane T. Alwin, "Approaches to the interpretation of relationships in the multitrait-multimethod matrix", in Herbert L. Costner, ed., *Sociological Methodology, 1973–1974* (San Francisco, Cal., Jossey-Bass, 1974), pp. 79–105.

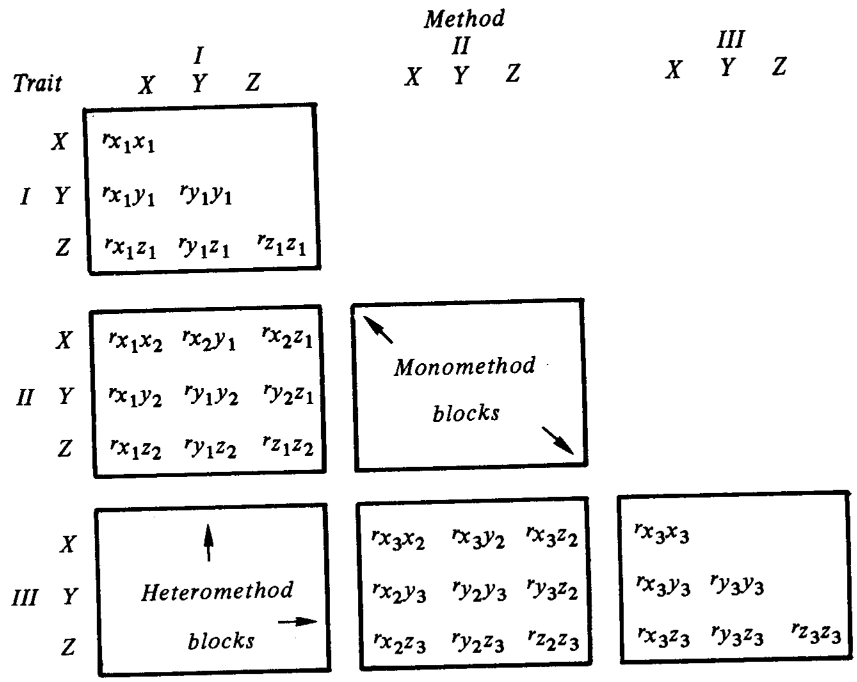


Figure I. Multitrait-multimethod matrix for three traits and three methods

Source: Duane T. Alwin, "Approaches to the interpretation of relationships in the multitrait-multimethod matrix", in Herbert L. Costner, ed., *Sociological Methodology, 1973–1974* (San Francisco, Cal., Jossey-Bass, 1974), p. 81.

Note: Values in validity diagonals (MTHM) are in bold-face type.

etc.); the correlations in the lower triangle are called heterotrait-monomethod values since they represent correlations among measures obtained through a single method.

(2) The heteromethod blocks which display the correlations among variables measured by different methods. The diagonals of the heteromethod blocks give the correlation between two different methods of measuring the same variable($r_{x_1x_2}$, $r_{y_1y_2}$ etc.) and for this reason are referred to as validity values; the correlations in the lower triangle of these blocks are called heterotrait-heteromethod values since they represent correlations between different variables measured by different methods.

Given the matrix of values in figure I and their characteristics, Campbell and Fiske[11] outline several criteria needed to establish convergent and discriminant validity, which are condensed by Alwin as follows:

"The entries in the validity diagonal should be significantly different from zero and sufficiently large to encourage further examination of validity. This requirement is evidence of convergent validity. Second, a validity diagonal value should be higher than the values lying in its column and row in the heterotrait-heteromethod triangles. That is, a validity value for a variable should be higher than the correlations obtained between that variable and any other variable having neither trait nor method in common. . . . A third common-sense desideratum is that a variable correlate higher with an independent effort to measure the same trait than with measures designed to get at different traits which happen to employ the same method. For a given variable, this involves comparing its values in the validity diagonals with its values in the heterotrait-monomethod triangles. . . . A fourth desideratum is that the same pattern of trait interrelationship be shown in all of the heterotrait triangles of both the monomethod and heteromethod blocks."[12]

General causal model for analyzing the multitrait-multimethod matrix

The scheme proposed by Campbell and Fiske is clearly a promising first step in establishing construct validity and for analyzing a series of results obtained by alternative methods. Evaluation of this proposal, however, has detected a number of weaknesses. For example, although Campbell and Fiske note the relevance of reliability, they provide no systematic way of taking this factor into account.[13] As stated by Siegel

[11] Loc. cit., pp. 82–83.

[12] D. T. Alwin, loc. cit., p. 82.

[13] *Ibid.*, p. 83.

and Hodge,[14] it is desirable to develop models that would simultaneously address questions of validity and reliability. Of the alternative approaches to the Campbell-Fiske methodology well reviewed by Alwin, outlined below are the features of the general causal model since it clearly brings out this distinction.

As against the definitions of reliability and validity given above in the section on elementary concepts, one may reformulate the distinction in terms of the unobserved construct one is attempting to measure and the factors that cause the observed values to depart from it. From this standpoint, questions of reliability focus on the random errors (e) that cause observed scores (O) to depart from the "true" scores (T) represented by the formula:

$$O = T + e$$

Questions of validity arise when variation in the observed scores is seen as arising also from other factors—the methods of measurement or other constructs—which are non-random or systematic. Under certain conditions,[15] the effects of both types of variation can be estimated through the factor-analytical techniques developed by Jöreskog.[16]

In the formulation relevant to this discussion, each measured variable would be treated as arising from the unobserved construct, the method used to obtain the measure, and random variation. Figure II represents this situation for the hypothetical situation posited earlier of three methods each used to measure three constructs. The correlations obtained are the basis of estimating a series of parameters α_i and β_i for structural equations of the form:

$$X_1 = \beta_1 X + \alpha_1 M_1 + u_1 \text{ etc.}$$

in which "each measured variable has non-zero loadings on one trait factor (β) and one method factor (α) but zero loadings on all other trait and method factors".[17] It should be noted that this model allows both trait and method factors to be correlated, and these correlations to be estimated.

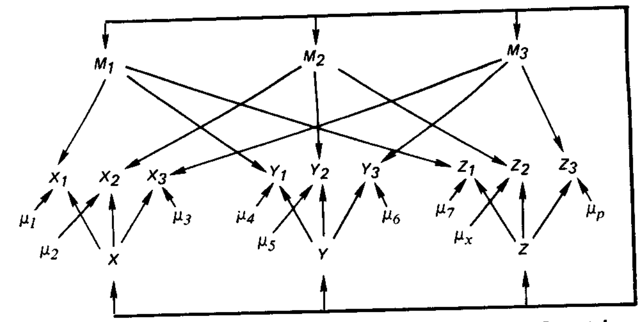


Figure II. Path diagram for the multitrait-multimethod matrix ($p = 3$, $m = 3$)

Source: Duane T. Alwin, "Approaches to the interpretation of relationships in the multitrait-multimethod matrix", in Herbert L. Costner, ed., *Sociological Methodology, 1973–1974* (San Francisco, Calif., Jossey-Bass, 1974), p. 84.

Some final observations

The foregoing observations were designed to show that developments in the study of reliability and validity, particularly through a causal model in which each source of variation is identified, are structurally similar to the type of data that may become increasingly available from the evaluation of family planning programmes, It would appear to be worth while, therefore, to carry out further analysis of the potential applicability of these methods to these data. Can one identify three or more methods of evaluation which produce comparable measures on three or more constructs thought to influence fertility? If not, is it possible to alter the analytical framework to fit the type of data produced? Or, alternatively, is it more promising to consider modifications of current methods of evaluation so that these fairly powerful techniques can be utilized? It may turn out that these methods for treating reliability and validity prove more useful for studying alternate ways of measuring the effect of a programme, not on fertility, but on variables antecedent to it, such as the proportion of the population currently using contraception and continuation rates. For example, prevalence of contraceptive use along with characteristics of users may be obtained through surveys or derived from service statistics and the general causal model outlined above may be applicable to examining these alternative approaches to measuring these key aspects of a programme. Questions of this type might well occupy some of the attention of this meeting if increased data on evaluation are to be analysed to maximum advantage.

---

[14] Paul M. Siegel and Robert W. Hodge, "A causal approach to the study of measurement error", in H. M. Blalock and A. B. Blalock, eds., *Methodology in Social Research* (New York, McGraw-Hill, 1968), p. 56.

[15] The conditions relate to the question of identifiability, which may be thought of as the number of unknown parameters to be estimated in relation to the number of observed correlations.

[16] K. G. Jöreskog, "A general approach to confirmatory maximum likelihood factor analysis", *Psychometrika*, vol. 34 (June 1969), pp. 183–202; *Idem*, "A general method for analysis of covariance structures", *Biometrika*, vol. 57, No. 2 (1970), pp. 239–251.

[17] D. T. Alwin, loc. cit., p. 84.

SOME ASPECTS OF DETERMINING APPROPRIATE METHODS OF EVALUATION FOR MEASURING THE IMPACT OF THE FAMILY PLANNING PROGRAMME ON FERTILITY IN THE REPUBLIC OF KOREA

*Kap Suk Koh**

The recent demographic history of the Republic of Korea is well known to those interested in the study of human fertility. The total fertility rate in the Republic of Korea declined from an estimated 6.17 in the early 1960s to 3.9 in the early 1970s. The crude birth rate declined by approximately 45 per cent between 1960 and 1975, falling from 42 to 24. This decline in the fertility of women in the Republic of Korea took place in conjunction with several other important developments. Over-all economic development has increased greatly since the early 1960s. During the period 1965–1975, the gross national product grew at an average annual rate of approximately 11 per cent. *Per capita* income increased from $83 in 1960 to $513 in 1975. In addition, urbanization increased. Approximately 50 per cent of the population currently live in cities of 50,000 or more; in 1960, the comparable figure was 28 per cent.

Of more immediate interest is the fact that in 1961, the Government of the Republic of Korea established a national family planning programme. Since then, approximately 8 million women have received programme-supported services. In addition, the age at marriage has risen steadily and the use of induced abortion increased considerably throughout the period 1960–1975.

As a consequence of these developments, the Republic of Korea has become something of a battleground for those arguing about the impact of national family planning programmes on fertility. Because of the obvious importance of the effect of family planning programmes on fertility, as well as the interest in the experience of the Republic of Korea, this meeting is especially welcomed by experts from that country. It is unfortunately the case that those experts do not know as much as they would like to about how to measure the impact of family planning programmes on fertility. It is hoped that this meeting will clarify at least some of the issues involved and lead to increased efforts in this important area.

Several of the methods discussed in the background papers for this meeting have been employed with data of the Republic of Korea. Cho and Retherford[1] have used the standardization approach. Smith and Koh[2] have carried out a trend analysis. Yang[3] has reported on an experimental study. Lee and Isbister[4] employed data of the Republic of Korea in a report. The Korean Institute for Family Planning[5] has calculated couple-years of protection (CYP). Analyses have been made of the reproductive process, following the leads of Potter and others. Mode and Littman[6] have worked on a simulation model. Analysis along these lines continues at several institutions both in the Republic of Korea and elsewhere.

All of the studies done so far have suffered from various short-comings. It is impossible to make specific comparisons here of the results obtained with the various procedures because of the different data sources and temporal reference periods employed in each study. Instead of detailing particular failings of past research, the present author would like to point out some of the problems that are most troublesome in the continuing analysis of the impact of the family planning programme on fertility in the Republic of Korea.

The selection of a given evaluation procedure depends upon one's orientation towards certain measurement problems and the availability of the necessary data. In the Republic of Korea, the existence of a body of high-quality survey data and the problems with programme statistics limit the approaches that can be employed in a straightforward fashion.

---

* Director, Evaluation Division, Korean Institute for Family Planning, Seoul, Republic of Korea.

[1] Lee-Jay Cho and Robert D. Retherford, *Comparative Analysis of Recent Fertility Trends in East Asia*, Reprint No. 51 (Honolulu, Hawaii, East-West Population Centre, 1973).

[2] D. P. Smith and K. S. Koh, "Population projections (1960–2001) for the Republic of Korea", Seoul, Population Council, September 1970 (mimeographed).

[3] Jae Mo Yang, "Fertility and family planning in rural Korea", *Proceedings of the World Population Conference, Belgrade, 30 August-10 September 1965*; vol. II, Selected Papers and Summaries: *Fertility, Family Planning, Mortality* (United Nations publication, Sales No. 66.XIII.6), pp. 309–312.

[4] B. J. Lee and John Isbister, "The impact of birth control programs on fertility", in Bernard Berelson and others, eds., *Family Planning and Population Programs: A Review of World Development* (Chicago, Ill., University of Chicago Press, 1966), pp. 737–758.

[5] Korean Institute for Family Planning, "1974 family planning programme evaluation, based on service statistics", Seoul, 1975.

[6] C. J. Mode and G. S. Littman, "A study of impact of age at marriage, sex preference, abortion, contraception, sterilization on population growth in Korea by computer simulation", Philadelphia, Pa., Drexel University, Institute for Population Studies and Department of Mathematics, n.d. (mimeographed).

Measurement methods

The standardization approach, at least as it has been utilized to date, has failed to come to terms with an important problem. It has missed the influence of socio-economic factors which are probably crucial determinants of fertility, both directly through their effect on marital fertility and indirectly through their impact on fecundity and the probability of having a live birth and a given level of infant mortality. One aspect of this problem is that observed changes in marital fertility are difficult to allocate to specific causal factors, such as the family planning programme. Although the standardization procedure is useful as a measure of changes in fertility due to such factors as age structure, age at marriage, proportions marrying and marital fertility, it is less helpful in specifying the contribution of the family planning programme to changes in these variables, most importantly marital fertility. In short, there is need to take account of the complexity of demographic change and socio-economic development and their interrelations. Standardization procedures probably can be modified to take into account additional factors. These modifications are necessary if one is to measure the impact of family planning programmes. A hopeful sign is the increased availability of sample data obtained from censuses that can be manipulated in a way that allows one to control to some extent the impact of socio-economic factors and to study changes in fertility among various subpopulations.

The projection approach has not been popular in the Republic of Korea. Some informal attempts have been made to utilize this procedure but the results have not been widely distributed. In part the projection approach suffers from the same short-comings as the standardization procedure. For the most part, available projection models have no completely adequate way to incorporate socio-economic and development variables that may influence the level of fertility independent of the family planning programme.

The specification of hypothetical fertility trends has usually been an unsatisfactory exercise. The complexity of recent demographic developments in the Republic of Korea means that the trends one assumes would have existed if there had been no family planning programme must be very carefully specified. In a country where the family planning programme has been influential, as is the case for the Republic of Korea, it is extremely difficult to specify what would have happened in the absence of the programme. Attributing changes in fertility to programme and non-programme factors is a major problem. For example, induced abortion,[7] in part supported by the programme but also widely available through non-programme channels, complicates any analysis. It is especially difficult to estimate what the impact of abortion would have been without the programme. The question arises whether the programme has prevented a higher level of abortions through the provision of contraception or whether it has helped to create the atmosphere and provide the resources that have made abortion so popular.

The Republic of Korea has been the location of several well-known experimental studies.[8] The difficulty with the experimental-design approach is that it is nearly impossible to assume that control and experimental areas are not influenced by unmeasured or even unknown factors. The "Hawthorne" effect is probably important in treatment areas where villagers are not used to and do not know how to deal with the considerable attention given them by programme workers. The rapid social changes of the past 15 years makes it extremely difficult to evaluate the differences between experimental and control areas. Moreover, because of the irregular pattern of change it is hard to judge the extent to which changes in one area are a reflection of changes in other areas of the country. Lastly, it must be pointed out that the experimental approach, although perhaps the most ideal procedure in theory, is extremely expensive and time-consuming in practice and is obviously limited by ethical considerations.

Couple-years of protection has become the central methodology in the official evaluation of the programme in the Republic of Korea. The CYP approach has been used for evaluation purposes since 1974. Currently, family planning acceptor targets also are calculated on the basis of the CYP approach. Several problems have arisen with this approach. Among the most troublesome are the accuracy of service statistics and the limited and too quickly outdated data on contraceptive continuation. Measuring the impact of abortion is also difficult because it is heavily influenced by the national programme. The most severe problem encountered in the Republic of Korea concerns the accuracy of the data collected through problem channels. What is needed is some way to measure both the reliability of service statistics and the continuation of contraceptive use. Several surveys along these lines have been conducted in the past. However, there is reason to believe that recent changes in target-setting procedures and payment to workers may have changed the situation sufficiently so that new surveys are needed. The case in the Republic of Korea is one example of the general problem of data availability and reliability, and the changes in these over time that limit the application of the CYP procedure.

The last four procedures, component projection, analysis of the reproductive process, regression analysis and simulation models, have received less attention in the Republic of Korea. The component projection approach, although it has been employed with national data, is in some respects unattractive because of the difficulty of directly including a measure of the influence of socio-economic factors. This

[7] On this subject, see S. H. Han, *The Study on Induced Abortion* (Seoul, Korean Institute for Family Planning, 1973).

[8] J. M. Yang, loc. cit.

method suffers from the same problem as the procedures discussed earlier. The transformation of the Republic of Korea over the past 15 years makes the inclusion of socio-economic factors and some estimate of the impact of their change on fertility an important element in any analysis of the change in fertility.

The modes of analysis developed by Potter[9] and Wolfers[10] are useful in the Republic of Korea because they allow the analyst to maximize the information obtained from a series of surveys carried out at regular intervals since the early days of the national programme. In addition, this approach appears to be better suited than others for handling the impact of socio-economic factors independent of the family planning programme.

Like the CYP approach, those developed by Potter and by Wolfers estimate programme impact directly from the number of acceptors. This is a limitation since the data available are known to be faulty with respect to both the number of acceptors of each programme method and their continuation. A special advantage of this approach is the fact that it incorporates a way of dealing with non-programme contraception. This factor is important in the Republic of Korea, where the combined influences of economic development, rising personal income and urbanization have increased the private market for contraception.

Regression analysis is another approach that is useful for the analysis of changing fertility in the Republic of Korea. A major problem that has limited the application of this method is the lack of relevant data. New efforts have recently been initiated to collect longitudinal measures of both programme activity and socio-economic changes for small areas. It appears likely that before long researchers will be able to provide new insights into the factors that have influenced the fertility decline, especially in the rural areas of the country. A particular advantage of the regression approach is becoming clear even during the preliminary stages of analysis. Because of the variety of measures of social and economic change that can be calculated from the data collected by various government agencies, it appears possible not only to estimate, albeit in a crude way, the impact of social and economic changes but to differentiate between the impact of socio-economic changes of various types (as indicated by different measured variables).

The last approach covered in the background papers, simulation models, has been employed only in a limited way. Mode and Littman[11] provide a useful analysis, although limited because of its inability to take account of the complexity of factors influencing fertility in the Republic of Korea. The model does incorporate sex preference as a key variable determining the use of contraception. In this sense, it represents an improvement over the techniques that either cannot handle attitudinal factors or ignore their impact. Given the current levels of interest and technological skill it is unlikely that simulation models will develop rapidly within the Republic of Korea as a means to measure the impact of the family planning programme.

Summary and conclusions

All of the methods to be discussed by this meeting could be usefully employed in the context of the Republic of Korea. However, because of the rapid development of the country, it is most desirable that procedures capable of explicitly treating the impact of socio-economic changes be developed and applied.

The complexity of the interaction between programme and non-programme factors remains the most significant barrier to a completely satisfactory estimate of the impact of the family planning programme in the Republic of Korea. Sophisticated experimental designs can measure fertility decline and programme impact in a precise way. However, national programmes in family planning and development cannot be experimental. One must settle for approximations often obtained through indirect observation. It will not be possible to solve the "substitution" problem although it is hoped that it can be dealt with in a more satisfactory fashion. (One advantage of the analysis of the reproductive process and of various regression procedures is that they represent improved ways of dealing with this "substitution" problem.) In the Republic of Korea, it is known that acceptance and use through programme channels has been substantial and that there has been a substantial fertility decline due in large measure to a reduction in marital fertility. The problem of how to relate these two facts deserves more attention.

Two additional problems should be mentioned in closing. The measurement of the impact of family planning on fertility must be done in a way that makes sense to administrators and others. Rapid feedback to administrators and those voting on programme budgets is very important. Likewise, being able to apply methods quickly and within the context of limited resources is extremely important. One should not concentrate on measurement problems to the extent that one forgets about programme workers and the importance of the programme impact for management and over-all social and economic development.

---

[9] Robert G. Potter, "Application of life table techniques to measurement of contraceptive effectiveness", *Demography*, vol. 3, No. 2 (1966), pp. 297–304; *idem*, "Estimating births averted in a family planning program", in S. J. Behrman, Leslie Corsa, Jr. and Ronald Freedman, eds., *Fertility and Family Planning: A World View* (Ann Arbor, Mich., University of Michigan Press, 1969), pp. 413–434.

[10] David Wolfers, "The demographic effect of a contraceptive programme", *Population Studies*, vol. XXIII, No. 1 (March 1969), pp. 111–141.

[11] C. J. Mode and G. S. Littman, op. cit.

NEEDED RESEARCH FOR MEASURING THE IMPACT OF FAMILY PLANNING PROGRAMMES ON FERTILITY

*W. Parker Mauldin**

Family planning programmes probably have been more carefully studied and examined than any other large-scale programme effort in history. During the past 10–15 years, a great deal has been learned about such programmes, and in the process, social science research has become more sophisticated as well as contributing to the improvement of programmes. The following discussion seeks to identify some of the more important research needs related to measuring the impact of family planning programmes on fertility. A question is also posed whether enough is currently known about potential fertility so that it can be set aside as an issue.

Research needs

The substitution effect

Since the earliest days of discussion and analysis of the effects of family planning programmes on fertility, there has been recognition that:

(*a*) Users of fertility control methods prior to initiation of a national or large-scale family planning programme sometimes switch to the family planning programme for supplies and services. Such a switch could involve adoption of:

(i) A less reliable method, though that is thought to occur rarely; or

(ii) The same method because programme supplies are more convenient or less expensive; or

(iii) A more reliable method.

In spite of recognition of this type of substitution and a considerable amount of work that has gone into the problem, no really satisfactory accounting procedure has been developed to allocate appropriate proportions of such acceptors to the programme;

(*b*) A closely related problem is how to treat non-users at the start of a programme who subsequently have the number of children they desire and then accept a programme method. The question is how one estimates the proportion of those persons who would have adopted a fertility control method regardless of whether there had been a programme, and at what point in time they would have become users. The programme could, for example, influence one to become a user months in advance of what would have happened in the absence of a programme. One may then ask whether these programme and non-programme effects can be allocated;

(*c*) Another similar problem concerns the way in which to treat programme users who discontinue programme use but continue the use of fertility control. Such use is sometimes called the "extended use-effectiveness" of a given method (and sometimes is calculated on the basis of the first method adopted even though the person switches to another method);

(*d*) There is the reverse problem, or what might be called the catalytic effect of a programme. Many observers note that some national and large-scale family planning programmes create an awareness of family planning through their information and education programmes and through the activity of field-workers. Partially as a result, many persons adopt fertility control methods but do not become a statistic in the programme because their adoption takes place outside of the formal programme itself.

The two principal approaches to these problems, particularly to those outlined in (*a*) and (*b*), have been to question a subsample of acceptors within the family planning programme as to whether they were using fertility control methods before coming into the programme, and, if so, what method. This information permits an analyst to subtract such cases from the total number of acceptors and to exclude them in calculating births averted by the programme, though this is rarely done. Another method is regression analysis, which presumably takes account not only of users prior to initiation of the family planning programme but of non-users at the beginning of the programme who subsequently accept a programme method but would have accepted such a method in the absence of such a programme.

Fertility as a function of development

Perhaps the major theme of the World Population Conference held at Bucharest in 1974 was the interrelatedness of fertility and development, and the World Population Plan of Action recommends that "countries wishing to affect fertility levels give priority to implementing development programmes and educational and health strategies which, while contributing to economic growth and higher standards of living, have a decisive impact upon demographic trends, in-

* Senior Fellow, The Population Council, Inc., New York.

cluding fertility."[1] This is a needed area of research, and the guidelines for such research are spelt out in paragraph 32 of the Plan, which indicates that specified development goals, enumerated in that paragraph, generally have an effect on the socio-economic context of reproductive decisions that tend to moderate fertility levels.

Similarly, there is a need for research to determine the extent to which fertility is affected by family planning programmes in different socio-economic settings. Ideally, one needs to know the effect of a specified type of family planning programme administered with a specified intensity in different settings. More conventionally, one needs to know what the effect will be on fertility, within a specified period of time, of a family planning programme costing *X per capita*, at different levels of socio-economic development, assuming that: (*a*) all other factors are constant; or (*b*) socio-economic development continues at the same pace as in the past; or (*c*) it proceeds at some other specified rate. Research along these lines is particularly pertinent at this time because the United Nations Fund for Population Activities (UNFPA) is considering what strategy it should apply in allocation of its funds and basically is considering allocations on the basis of need within a country rather than on the basis of cost effectiveness or the ability of the country in question to use such funds effectively.

Impact as a function of availability

To what extent is the success of a family planning programme affected by the availability of different fertility control methods to large segments of the population? About two years ago, the present author asked a number of colleagues and associates to rank all of the countries with national family planning programmes on the extent to which various methods of fertility control were generally available to the population. Five classes of methods were used: conventional methods; orals; the intra-uterine device (IUD); male and female sterilization; and abortion. The method of rating availability of methods was subjective and crude, but there was, none the less, a remarkable relationship shown between the number of methods generally available and the extent of change in fertility by country. Observers differ on their evaluations of how much difference it would make if fertility control methods were generally available to the population, but the strategy of the largest single donor in the field is predicated on the assumption that the more generally available fertility control methods are, the more likely fertility is to decline, and, as a consequence, a great deal of money is being invested on this important assumption. Research is needed as a guide to Governments and donors in the allocation of resources and the development of programme strategies.

[1] See *The Population Debate: Dimensions and Perspectives*, Papers of the World Population Conference, Bucharest, 1974, vol. I (United Nations publication, Sales No. E/F/S.75.XIII.4), p. 160, para. 31.

When to stop

A much debated issue is how far a family planning programme will take one, and related to this issue is the question when such a programme becomes self-sustaining. It is in the self-interest of a Government and of the donor community in general to know when to stop investing in family planning programmes. Donors typically decide that they will no longer support activities in a given country when that country has reached a certain *per capita* gross domestic product, and in the family planning field, when it has developed a moderately mature programme and the growth rate has declined to the order of 2 per cent. This writer has suggested to programme personnel in one of the more successful programmes that an experimental design should be worked out and that programme activities should be stopped in different socio-economic settings, and the results measured over a period of time. Results should be measured, of course, in terms of changes in fertility, which is the crucial question with reference to the impact of family planning programmes designed to reduce rates of population growth. In addition, however, the assessment should take into account the extent to which underprivileged persons are either denied access to fertility control methods because of the change in policy or are faced with an undesirably heavy burden because of either the cost or the inconvenience of obtaining services and supplies. There are ethical issues involved here as well as demographic and programmatic issues, but it is the present author's assumption that wherever family planning efforts have been moderately widespread and successful, the abandonment of a programme, or the application of welfare criteria, would not mean cessation of supplies and services to that population. If such a programme were undertaken, it would be important for those involved to ensure that supplies and services should be available in the private sector, possibly including the provision of some types of contraceptive at subsidized prices through community distribution schemes or other more conventional commercial outlets.

Validation of births averted

Is there a gap, an inconsistency, between the reported number of users by method, continuation rates, the calculation of births averted, and actual changes in and levels of fertility? There has been a marked improvement in models relating to the calculation of births averted in recent years, but it is notable that each of these models, in effect, pays little or no attention to *coitus interruptus*, abstinence or rhythm. Each, in effect, assumes that a discontinuer is returned to the pool of non-users, but it is known from a number of studies that such persons have very different characteristics than never-users and that they more effectively control their future fertility than do never-users. Also, a recent study indicates that with current contraceptive methods it is not possible to attain a very low crude birth rate without rather widespread use of abortion. This latter conclusion has a number of implications, both ethical and programmatic. Accordingly, it is suggested that this area deserves considerable research attention.

One of the major areas that needs investigation is the definition of use of a fertility control method. If a person has an IUD inserted or is sterilized, clearly that person is a user of a fertility control method; but if a person accepts a gift of pills or condoms or even purchases those contraceptives, the individual is not necessarily a user. Surveys ask some questions relating to use, but this is a sensitive area and has not been widely researched. Some little attention is needed in this area. As was mentioned above, not much attention is paid to *coitus interruptus* or abstinence and relatively little to rhythm, to breast-feeding and perhaps also to other practices.

Biological determinants

Mahalanobis wrote in the foreword to *Couple Fertility*[2] that malnutrition was sufficiently widespread in India in the mid-1950s that one could expect a rise in fecundity with an improvement in economic conditions and that that increase would tend to offset the effects of a family planning programme. This thesis is currently receiving increasing attention, particularly by Rose Frisch and her colleagues, but also by the World Fertility Survey, which has developed a module in this field. The basic argument is that in many countries of the world malnutrition is so widespread that fecundity has been appreciably impaired even though fertility rates are moderately high. The expectation, according to this thesis, is that as food supplies become more plentiful and more equitably distributed, fecundity will be less impaired and fertility will increase. In most of the countries where malnutrition is widespread, there are national family planning programmes and the effect of such an improvement in nutrition could be to mask the effect of the national family planning programme which is in the direction of lowering fertility. Consequently, a country could come to the conclusion that its programme was being ineffective when, in fact, it was reducing expected fertility appreciably.

Potential fertility

Potential fertility is listed here as an area needing research although a colleague of the present author states that potential fertility is understood rather well now and that the problem is that the analyses do not sufficiently take into account what is known. The background paper on problems and issues (ESA/P/AC.7/1) contains a good section on potential fertility, and it states that simulation studies have shown that chance factors in fecundity operating on a group with identical probabilities will produce a subgroup with above-average fertility which is thereby disposed towards acceptance. Also, there has been some work that shows that a tendency exists for high fertility to regress towards the mean (and excluding subfecundity, the reverse; i.e., for persons with low fertility to regress towards the mean and thus to have higher fertility). The work of Brass and Wolfers is noted,[3] appropriately so at that point. In addition, the work of Sivin[4] has refined this concept a little further, and he has shown the expected fertility of women who accept a contraceptive method during the post-partum period and subsequently discontinue a method. This analysis shows that it takes several years to return to "expected" fertility for a group of such women. There remains, however, a residual question whether there is sufficient understanding of potential fertility and how to analyze it; perhaps this is one issue that can be resolved at this meeting.

Early indicators of change

There is a very considerable lag between the occurrence of events, their collection and tabulation, and analysis. It would indeed be helpful if it were possible to establish a short-cut to the conventional method of registration of vital events and the subsequent tabulation of such events. The attention that is being paid to birth intervals is a move in this direction. The query is whether other measures could be developed, and with the application of several of those methods, whether one could tell at an earlier time than is currently the case if fertility had changed and, if it had, also could give an order of magnitude of that change.

Validity of one-time surveys

There have been many improvements in surveys during the past decade or two and some of the methods of analyzing such data have been markedly improved. The World Fertility Survey is the largest undertaking of its kind and it is based exclusively on one-time surveys. None the less, a question remains as to how valid these surveys are.

A number of studies undertaken a few years ago indicated that one-time surveys often, perhaps almost always, underestimated fertility levels, and as a con-

---

[2] A. Das Gupta and others, *Couple Fertility*, National Sample Survey Report No. 7 (India, Ministry of Finance, Department of Economic Affairs, 1955).

[3] William Brass, "Assessing the demographic effect of a family planning programme", *Proceedings of the Royal Society of Medicine* (London), vol. 63 (November 1970), pp. 29–31; David Wolfers, "Births averted", in C. Chandrasekaran and Albert I. Hermalin, eds., *Measuring the Effect of Family Planning Programs on Fertility* (Liège, International Union for the Scientific Study of Population for Development Centre of Organization for Economic Co-operation and Development, 1975), pp. 163–214.

[4] Irving Sivin, "Fertility decline and contraceptive use in the international post-partum family planning program", *Studies in Family Planning*, vol. 2 (December 1971), pp. 248–256.

sequence, dual-purpose surveys were introduced. Indeed, the POPLABS project at the University of North Carolina (United States of America) was begun and is being carried forward because of the belief of a group of scholars associated with it that single-purpose surveys are typically deficient and unreliable. The National Sample Survey of India has instituted a dual collection system for vital statistics and those efforts clearly indicate that any one method by itself understates vital events appreciably. A question remains whether a more sophisticated analysis of the results would permit one to arrive close to the values produced with the dual system, but, to the best of this author's knowledge, such investigations are not being carried out.

There is the intriguing likelihood that the World Fertility Survey will show a remarkable decline in fertility because of two factors. First, historically one-time surveys have understated fertility; and, secondly, the Potter or domino effect results in a heaping of births about five years previously and gives an artificially high level for that period, but more important, a lower figure for the current period, thus indicating a recent decline in fertility. At a minimum, a serious analytical work is needed to analyze results in recent years and to acquaint scholars in the field both with the results and, to the extent that those results indicate that one-time surveys are satisfactory, with the specific analytical methods that should be used in the application of such surveys.

Satisfying demand

Most of the research relating to the impact of family planning programmes has concerned itself with the extent to which stated targets have been met or the extent to which fertility has fallen. More important is the extent to which family planning programmes have provided information, supplies and services that are needed and wanted by the population. As a part of a national family planning programme, a real effort should be made to determine what couples want, what women want and the extent to which those needs and wants are being met. A part of this relates to the extent to which coverage is being adequately supplied, including the location of facilities within some convenient time-distance frame, and also to the extent to which information and services are provided reasonably promptly and courteously. Cost also is an important factor. There is a tendency to say that the large-scale family planning programmes have invested so much in time, money and personnel that enough is being done in this direction; in fact, however, very little research has been undertaken to assess the extent to which such wants are being met. It is known that abortion is moderately widespread in many areas under appalling conditions, and this fact would appear to indicate that fertility control methods are in no sense generally available to the population. This is a major need, one on which very little research has been done to date.

Fertility of the entire population

When one focuses on measuring the impact of family planning programmes on fertility, there is a tendency to look only at acceptors within the programme and within limits this tendency is desirable. More important is what is happening to fertility in the population at large. If one assumes, as one must, that the family planning programme is only one element, albeit an important factor, in fertility reduction, it is important to know what is happening to fertility as a whole because presumably not only the family planning programme is working but social and economic development is occurring, and there is an important interaction between the two. If one looks only at the impact of the family planning programme, or rather looks only at acceptors within the programme, one may well be missing an important element of the programme effect, which is the interaction of the programme with social and economic factors.

Data needs

Because of the length of the foregoing section, the author had intended to omit any specific references to data needs other than those implicit in research needs, but two particular needs should be identified:

(1) *Improvement in vital statistics*. It is recognized that the improvement of vital statistics is a long-run process and perhaps because of this aspect, the major actors in the field have done virtually nothing during the past 25 years to improve vital statistics. The United Nations system is, in this writer's opinion, regrettably deficient in this area. UNFPA is putting some funds into the general field, although this writer considers that a questionable use of UNFPA funds. The Statistical Office of the United Nations states that the improvement of vital statistics is important, but it attaches very low priority to this activity. An independent effort is being made to organize an international vital statistics group to focus attention on this area and perhaps this is the best way to proceed. In any event, there is a very great need here and very little attention is being given to it;

(2) *Five-year censuses*. In the absence of good vital statistics, a programme that takes censuses every five years would be an important addition to the current knowledge about population and population change. Censuses are, of course, relatively expensive, but they have the virtue of requiring little or no foreign exchange if the countries involved feel a need for such data. It is suggested that a programme to encourage countries to take censuses every five years, particularly those with poor vital statistics, could result in a marked improvement in the knowledge about population change.

COMPONENT PROJECTION *VERSUS* OTHER TECHNIQUES FOR ASSESSING PROGRAMME ACHIEVEMENT TOWARDS A TARGETED FERTILITY REDUCTION

*Robert G. Potter**

A number of the countries with national family planning programmes have set demographic goals, usually expressed in terms of specified reductions of either the crude birth rate or the growth rate within a given target period. In this context, there is likely to be not only an interest in the impact on fertility of past programme effort but a need to estimate the required acceptances of programme methods necessary to realize the current demographic goal. A variety of techniques have been utilized,[1] conditioned heavily by the data and analytical skills locally available, but most of which represent varyingly complete realizations of what in the background paper on problems and issues (ESA/P/AC.7/1) is called the component projection approach (CPA). It is argued here that CPA does possess special advantages for target setting and that, on balance, none of the other approaches listed in the background paper are quite as appropriate.

Additional work on the component projection approach

Besides the applied work recapitulated by Ross, there has been extensive work systematizing CPA since the pioneering effort of Lee and Isbister. A project of the Economic and Social Commission for Asia and the Pacific (ESCAP) led to the development of a computerized component projection scheme (PROJ5), which was applied to Pakistan.[2] Meanwhile, Nortman and Bongaarts[3] were designing their first uncomputerized version of TABRAP, which for the first time offered a direct rather than a trial-and-error means of deriving acceptor targets. In a subsequent ESCAP project, the Multinational Study in Methodologies for Setting Family Planning Targets in the ESCAP Region, planned and at first directed by K. Srinivasan, TABRAP has been amplified into a nine-programme, documented computer package called the ESCAP Target-Setting System and applied to data from eight Asian countries.[4] Another study[5] presents a more theoretical treatment of CPA, based on stable population theory.

Basic factors

Three sets of basic factors have to be considered when bridging demographic goals and contraceptive targets. They comprise demographic setting, demographic goal and family planning programme factors:

(1) *Demographic setting:*

(*a*) Size of population;

(*b*) Initial age-sex-marital status distribution;

(*c*) Male and female age schedules of mortality and changes over the target period;

(*d*) Age-specific proportions of those married among women of reproductive age and their trend over the target period;

(*e*) Age schedule of marital fertility among non-users of contraception and changes over the target period;

(2) *Demographic goal:*

(*a*) Length of target period;

(*b*) Change in crude birth rate (or rate of natural increase) aspired to;

(3) *Programme factors:*

(*a*) Continuation and effectiveness of each method (programme or private sector) of contraception offered and the age distribution of the users of each method;

(*b*) Age of acceptors;

(*c*) Relative popularity of the different contraceptives, in each age class of acceptors;

---

* Professor, Department of Sociology and Anthropology, Brown University, Providence, Rhode Island, United States of America.

[1] John A. Ross, "Acceptor targets", in C. Chandrasekaran and Albert I. Hermalin, eds., *Measuring the Effect of Family Planning Programs on Fertility* (Liège, International Union for the Scientific Study of Population for the Development Centre of the Organisation for Economic Co-operation and Development, 1975), pp. 55–92.

[2] Robert G. Potter and S. L. N. Rao, *Some Techniques for Measuring the Impact of Contraception: An Aid to Target Setting*, Asian Population Studies Series, No. 18 (E/CN.11/1119) (Bangkok, 1974). See also *idem*, "Future family planning impact; method and data requirements", *Economic Bulletin for Asia and the Far East*, vol. XXIV, No. 1 (June 1973) (United Nations publication, Sales No. E.74.II.F.3), pp. 74–82.

[3] Dorothy Nortman and John Bongaarts, "Contraceptive practice required to meet a prescribed crude birth rate target: a proposed macromodel (TABRAP) and hypothetical illustrations", *Demography*, vol. 12 (August 1975), pp. 471–490.

[4] S. Kirmeyer, "The ESCAP target-setting system: a computerized component projection approach", background paper prepared for the Economic and Social Commission for Asia and the Pacific (ESCAP), September 1975; ESCAP, Population Division, "Report of the Study Directors' Meeting for the Multinational Study in Methodologies for Setting Family Planning Targets in the ESCAP Region", Bangkok, 24–30 September 1975; Robert G. Potter, "ESCAP target-setting system: rationale, strengths and limitations", background paper prepared for ESCAP, September 1975.

[5] Robert G. Potter and S. L. N. Rao, "Contraceptive impact over several generations", in T. N. E. Greville, ed., *Population Dynamics* (New York, Academic Press, 1972), pp. 137–166.

(*d*) Potential age-specific fertility of users, had they not accepted contraception.

Logic of the component projection approach

The logic of CPA may be illustrated by the two computer programs, TABRAP, which gauges acceptance requirements to realize a pre-designated crude birth rate path, and CONVERSE, which treats the converse problem of determining consequences of specified acceptance schedules on the crude birth and growth rates.

Both computer programs embody conventional component projection schemes that, utilizing five-year age classes and projection steps of one year, are elaborated to handle age schedules of acceptance and continuation of contraception. Only married women 15–44 years of age and not currently using a method are defined as eligible to accept.

In CONVERSE, births allowed by family planning are estimated in five main steps:

(1) Partition acceptance into private and public sectors and within each sector by age class and contraceptive method;

(2) Convert acceptance into woman-years of protection classified by age of user;

(3) Multiply woman-years of protection by age-specific potential fertility rates of users, which products, appropriately lagged and summed, yield annual births prevented;

(4) Subtract births prevented from "potential" births, i.e., births expected without family planning, to obtain births allowed.

Two projections are always computed: one based on allowed births, and the other based on potential births.

In TABRAP, the problem of deducing acceptance requirements to achieve a pre-designated crude birth rate path presupposes a number of constraints to make the solution direct rather than iterative. The constraints chosen for TABRAP relate to the relative popularity of programme and private sector methods within five-year age classes and to relative age-specific dispositions to accept among eligible women. Analysis proceeds year by year. The sequence of steps is as follows:

(1) For any programme year t, determine total births to be prevented in order to realize the crude birth rate targeted for that year;

(2) Determine births already averted by users who accepted during programme years $t - 2$ or earlier;

(3) Difference these two quantities as an estimate of births to be prevented in year t by new acceptors in year $t - 1$;

(4) Convert this last quantity into corresponding woman-years of use;

(5) Transform this protection time into required numbers of new acceptors classified by age and method.

Diversity of objectives

Assessing the impact of family planning on fertility for purposes of setting contraceptive targets properly entails a variety of evaluations. It matters greatly whether personnel from the family planning administration participate in the formulation of the demographic goal or whether it is decreed by a different agency, such as central planning. In the latter case, first priority may have to go to showing that a continuation of current acceptance trends, or even a considerable acceleration thereof, will still leave the decline of vital rates far short of the pronounced goal. In the former case, with a relatively more feasible goal, the primary task is likely to be one of converting that goal into a series of annual acceptance quotas. However, usually only a terminal crude birth or growth rate is given and the path of annual rates connecting initial and terminal rate is not specified. Multiple component projections may be required to decide on an optimal path that does not, for example, call for an excessively fast buildup of acceptance. If certain factors are believed to be under the control of the programme, such as influencing method preferences, improving continuation for certain contraceptives or attracting younger acceptors, then more projections are needed to investigate the pay-offs of these influences. If, in order to implement the projections at all, it has been necessary to borrow data from another country for one or more factors, then it behooves the evaluator to make additional projections in order to test the sensitivity of results to higher or lower estimates of those factors. Then too, participation in the formulation of the demographic goal may have meant making a preliminary set of projections to determine what is feasible. If service statistics are available for the programme, but not for the private sector, then inferences about the impact of the latter may be possible by comparing observed fertility with the fertility estimated on the basis of a projection assuming the sole operation of the programme. Lastly, if the computerized projection scheme carries a detailed output, then valuable insights may be gained as to why so much acceptance is required to register a given effect on vital rates (e.g., adverse age compositional effects, high acceptance-to-use ratios reflective of discontinuation, sterility or divorce).

Assessment of the component projection approach

In favourable circumstances of data and computerization, the strengths of CPA are its ability to encompass the factors enumerated above, its provision of detailed output and its being cheap enough in operation to run many projections, when desired. By recourse to the programming technique of overlaying, either TABRAP or CONVERSE can be fitted onto a 32K computer.[6]

[6] S. Kirmeyer, "A comparison of the three versions of TABRAP/CONVERSE—principal models of the ESCAP target-setting system", unpublished paper, March 1976.

Several weaknesses exist, of course. First, data requirements are heavy. Especially problematical are the estimation of continuation rates and the age-specific potential fertility rates of future users and married non-users, these latter two sets of quantities being incapable of direct observation. Secondly, CPA operates at the level of births, not pregnancies, making it awkward to incorporate induced abortion. In some manner each abortion has to be made equivalent to an appropriate amount of contraceptive protection time.

Thirdly, CPA shares with several of the other techniques listed in the background paper (ESA/P/AC.7/1) a lack of any systematic procedures for estimating indirect effects of family planning programme on private sector birth control or for assessing to what extent programme contraception is substituting for fertility control that would have occurred anyway. Fourthly, if data on the private sector are missing, then one is hard put to apportion roles for private sector and programme; and even if one succeeds in contriving an estimate of comparative contributions at the beginning of the target period, their evolution during the remainder of the target period remains necessarily speculative.

Comparative utility of other techniques for target setting

Most of the techniques listed in the Secretariat background paper (ESA/P/AC.7/1) are designed to evaluate the impact of family planning on fertility for purposes other than target setting.

For example, standardization is typically employed to resolve an observed past change of fertility into components attributable to concurrent changes in age structure, proportions married and marital fertility. The part played by family planning in the last-mentioned change is not rendered explicit.

Because trend analysis (the projection approach) is directed towards estimating "net" effects of programme activity, it has to be speculative; and the interactions of contraceptive acceptance, use and potential fertility within and between private and public sectors are left implicit. Though subject to problems of matching and contamination, the approach of experimental designs has the advantage of offering direct estimates of net programme impact. Nevertheless, a comparison of the fertility change of a non-random, self-selected experimental sample with that of a purportedly matched control sample is hardly the same as measuring the fertility consequences of a programme for an entire population or its acceptance requirements to attain a prescribed demographic goal.

What is called the analysis of the reproductive process constitutes a specialized and elaborate tool for assessing births averted per segment of intra-uterine devices (IUD) and, less conveniently, of other contraceptives whose practice is less continuous. The high data requirements of this methodology rule out application except in rare instances.

The couple-years of protection (CYP) approach represents a set of rules for converting the acceptances of a mixture of contraceptives into an absolute number of births averted, but the technique lacks age specificity and fails to locate births averted in time. Moreover, the conversion rules, derived originally for Pakistan, lack general applicability to other cultures and times.

If family planning service statistics are available but only limited demographic data exist, appeal may be made to linear regressions predicting the crude birth rate or growth rate as a function of the proportion of married women of reproductive age using contraception.[7] However, a regression crudely summarizing the experience of many countries cannot be expected to reflect the idiosyncracies of a particular country. A much more powerful regression analysis becomes possible when comparable family planning and demographic data are available for subareas of a country. The implicit interaction among acceptance, use, and potential fertility and apportionment between private and public sectors is an appropriate aspect, namely, that belonging to the immediate past of the country. How relevant the resultant regressions are for target setting depends upon the indexes available as independent and dependent variables. Unfortunately, few countries have the data to support this type of analysis.

Most simulation models of the family-building process are disqualified for target-setting relevance by virtue of representing cohorts instead of period populations. Exceptions are the micro-simulation model POPSIM[8] and the stochastic population model.[9] As family-building models operating at the pregnancy level, both models are better able to handle induced abortion than is CPA. With a monthly time unit rather than a year, they can be more precise about the timing of events. However, data requirements as well as development and operating costs are much higher. The latter are so high, in fact, as to be incompatible with the substantial number of runs typically necessitated by the several objectives outlined above. Nor does the extra precision really justify the added costs inasmuch as the basic weaknesses of CPA—estimation of continuation rates and of potential fertility along with apportioning shares to programme and private sector—are not being solved by the simulation models.

---

[7] R. G. Potter and S. L. N. Rao, *Some Techniques for Measuring the Impact of Contraception*; and for an updating, Dorothy Nortman and E. Hofstatter, *Population and Family Planning Programs: A Factbook*, Reports on Population/Family Planning, No. 2, 7th ed. (New York, The Population Council, 1975), p. 85.

[8] D. G. Horvitz and others, *POPSIM, A Demographic Microsimulation Model*, Carolina Population Center Monograph No. 12 (Chapel Hill, N.C., University of North Carolina, 1971).

[9] C. J. Mode, "Perspectives in stochastic models of human reproduction: a review and analysis", *Theoretical Population Biology*, vol. 8 (December 1975), pp. 247–291.

A NOTE ON TWO APPLICATIONS OF MICRO-SIMULATION MODELS TO THE PROBLEMS OF EVALUATING FAMILY PLANNING PROGRAMMES

*Jeanne Clare Ridley**

The purpose of this brief note is to report on two recent applications of micro-simulation models to methodological problems encountered in the evaluation of family planning programmes.[1] Not only do these applications illustrate the types of methodological questions that may be investigated but they are suggestive of a wide range of questions of a methodological nature that can be investigated with such models. Further, the results of these applications indicate that the more conventional fertility measures may be more adequate for assessing changes in fertility than is commonly assumed.

The two micro-simulation models, POPREP[2] and REPSIM-B, employed in the two applications summarized here, have been described in detail elsewhere.[3] Both models employ Monte Carlo methods. Briefly, POPREP is an elaboration of POPSIM[4] in that it expands considerably the fertility component of POPSIM by including a number of the basic biological determinants of human reproduction. POPREP generates the reproductive histories of a hypothetical female population of all ages. The model provides for the demographic factors of marriage, marital dissolution, remarriage and death. The biological factors included are fecundability; various outcomes of pregnancy, including induced abortion; the post-partum non-susceptible period; and sterility. In addition, it also provides for the use of contraception, surgical sterilization and induced abortion.

REPSIM-B is an extension and elaboration of REPSIM-A.[5] REPSIM-B simulates the detailed reproductive history of a hypothetical cohort of women. It provides for a woman marrying, dying, becoming sterile and becoming pregnant, and for varying outcomes of pregnancy including induced abortion. Provision is also made for the adoption of contraception or surgical sterilization.

The specific objective of the POPREP application summarized here[6] was to explore the effects of different patterns of contraceptive use on alternative measures of fertility. The measures studied were age-specific birth rates, the total and general fertility rate and various measures of birth intervals.

The initial population assumed in the simulation was derived from the census data for India in 1961. Assumptions regarding marriage, mortality and other parameters were assumed to be similar to those characterizing Indian women. For example, over half of the women were assumed to marry by age 16; life expectancy at birth was assumed to be approximately 50 years; the mean length of the post-partum period was assumed to be about 12 months; and fecundability was assumed to vary between women and by age with a mean maximum level of 0.20.

The initial population was first simulated for a period of five years. During that period, no use of contraception was assumed. Beginning in the sixth year, three different patterns of contraception were postulated. Acceptance of contraception was assumed to vary with age and with parity. The first pattern of contraceptive use assumed a moderate rate of acceptance of contraception with a contraceptive effectiveness of 100 per cent. The second pattern had higher acceptance rates and also had a contraceptive effectiveness of 100 per cent. The third pattern postulated the high acceptance rates with contraceptive effec-

---

* Professor of Sociology, Kennedy Institute, Center for Population Research, Georgetown University, Washington, D.C.

[1] For recent reviews of various applications of micro-simulation models as well as other types of models, see Jane Menken, "Simulation studies", in C. Chandrasekaran and Albert I. Hermalin, eds., *Measuring the Effect of Family Planning Programs on Fertility* (Liège, International Union for the Scientific Study of Population for the Development Centre of the Organization for Economic Co-operation and Development, 1975), pp. 351–380; and P. A. Lachenbruch and others, "Use of models in evaluation of fertility dynamics: a review", paper prepared for the World Health Organization Task Force on Family Planning Evaluation.

[2] A manual describing the computer program of POPREP has recently been completed; see *User's Manual for POPREP and INTRVL* (Chapel Hill, N.C., University of North Carolina, Department of Biostatistics, 1975).

[3] P. A. Lachenbruch, M. C. Sheps and A. M. Sorant, "Applications of POPREP, a modification of POPSIM"; and Alice S. Clague and Jeanne C. Ridley, "The assessment of three methods of estimating births averted", both in Bennett Dyke and Jean W. MacCluer, eds., *Computer Simulation in Human Population Studies* (New York, Academic Press, 1973), pp. 305–328 and 329–382, respectively.

[4] D. G. Horvitz and others, *POPSIM, A Demographic Microsimulation Model*, Carolina Population Center Monograph No. 12 (Chapel Hill, N.C., University of North Carolina Press, 1971).

[5] Jeanne Clare Ridley and Mindel C. Sheps, "An analytic simulation model of human reproduction with demographic and biological components", *Population Studies*, vol. XIX (March 1966), pp. 297–310.

[6] This summary based on C. M. Suchindran and others, *Sensitivity of Alternative Fertility Indices*, final report on contract No. NICHD-2187 (Chapel Hill, N.C., University of North Carolina, 1976).

tiveness assumed to be only 90 per cent effective. The simulation for the three patterns of contraceptive use were carried out for 10 years.

Not surprisingly, for the entire 10-year period of contraceptive use all the fertility measures reflected the predicted impact of the three different patterns of contraceptive use. As stated by Suchindran and his colleagues, however: "What is more surprising, and perhaps disquieting, is the fact that, when these indices are calculated at earlier time points, they fail to indicate the relative effectiveness of the three patterns, and, in some instances, fail to provide conclusive evidence of a change in fertility." Only after the third year of contraceptive use did any of the various measures begin to reflect accurately the various patterns of contraceptive use. The measures most sensitive to showing change at the earliest points were the age-specific fertility rates, general fertility rates and total fertility rates. The open birth interval showed little change in the early years of contraceptive use although it did show considerable change towards the end of the 10-year period of contraceptive use. Closed intervals were not expected to show any change when contraceptive effectiveness was 100 per cent; they did not, however, show much change when contraceptive effectiveness was only 90 per cent. Life-table estimates of birth intervals were striking in that they showed little or no change.

Supporting the above-mentioned results as to the usefulness of the more conventional fertility measures are results obtained with an application of REPSIM-B. The original purpose of the particular experiments by Ridley and Clague that are summarized here[7] was not to assess the usefulness of various fertility indexes but rather to assess various methods of estimating births averted. Using data from two cohorts simulated under conditions characterizing women in India and in Latin America, the methods of estimating births averted developed by Lee and Isbister, by Potter and by Wolfers were evaluated against criterion values derived directly from two simulated cohorts. In the first cohort of each experiment, no contraceptive use was postulated; and in the second cohort, contraceptive use was postulated for a period of five years. Acceptance of contraception was at a moderate level and depended upon desired family size; contraceptive effectiveness was assumed to be 95 per cent.

Although the results obtained are limited by the experimental conditions assumed, the simpler method first developed by Lee and Isbister, which depends upon estimating age-specific birth rates, reflected most accurately changes in fertility resulting from contraceptive practice. The results also indicated that the effects of contraceptive use in a particular year were not clearly reflected in any of the estimates. The estimates for the total five-year period were considerably better.

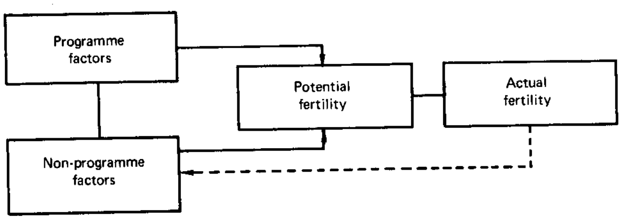
Although the results obtained in these two applications must be viewed most cautiously because of the restricted assumptions under which these investigations were carried out, they do suggest that future investigations might be profitably pointed in certain directions. First, more attention should be placed on studying the usefulness of more conventional fertility measures in evaluating fertility changes. The development of better data for estimating these measures, therefore, becomes of highest priority. Secondly, more investigations of the sensitivity and robustness of various fertility indexes are needed, as are studies of their sampling variability. Simulation models afford an important tool for carrying out such investigations.

[7] Jeanne Clare Ridley and Alice S. Clague, "Experiments with a microsimulation model of human reproduction" (forthcoming).

SOME ISSUES IN DETERMINING APPROPRIATE METHODS OF EVALUATING THE FERTILITY IMPACT OF FAMILY PLANNING PROGRAMMES

*Ismail Sirageldin**

The purpose of the exercise is to choose an index of family planning programme efforts; to choose an index of fertility; to measure changes in both indexes over a given period of time—usually, five or fewer years; to choose an appropriate theoretical frame and an appropriate methodology to relate family planning programme efforts to changes in estimated fertility after allowing for the effects of all other relevant factors, be they individual, group and/or societal; and, lastly, to make evaluative statements about the fertility impact of the family planning programme. This three-way relation is schematically illustrated below in its more general and gross way.



As is well known by now, behind this seemingly simple diagram are some of the most complex theoretical and methodological questions, many yet unresolved in the social and behavioural sciences. Each of the two boxes in the left-hand side of the figure includes a large set of factors for which there does not yet exist an established body of theory about their separate and/or combined effects on fertility, especially in the context of developing countries. Space does not allow a review of current knowledge. However, there are several recent surveys available.[1] This state of affairs undoubtedly increases the probability of introducing specification errors (not only of missing variables but of functional forms) when developing statistical models to estimate the basic structural parameters of the system. In the case of regression analysis, the exclusion of relevant variables from the regression may be a very serious error since the estimated coefficients may be seriously biased and the estimate of the residual variance could be biased upwards.[2] This problem of specification error is further compounded by the related problem of measurement errors in some of the key demographic and socio-economic variables, mainly because of the lack of adequate data base, and accordingly analysts resort to inadequate proxy variables. The relevant questions appear to be: whether it is possible to develop some general criteria to rank the various evaluation approaches; and whether useful evaluation can be made without adequate specification and given the inadequacy of available data.

In the useful background paper[3] prepared for this meeting, a survey of eight measurement methods was presented: standardization approach; trend analysis (fertility projection approach); experimental designs; couple-years of protection (CYP); component projection approach; analysis of the reproductive process; regression; and simulation models. Some of these measures vary widely in their theoretical foundation and specification details, in their levels of analytical sophistication and accordingly in the level of manpower skills required, their population coverage, their required data, their potential errors; and in the limitations imposed by their specifications on interpreting their derived conclusions and on their level of generalizations.

Ideally, in a cost-benefit framework, one should be able to say that for a given purpose method A is better (or worse) than method B if both methods produce the same service or "benefit" but the "cost" of A is lower (or higher) than that of B; or, alternatively, if both incur the same "cost" but A produces more (or less) of the same "benefit" than B. What one has, however, is a multiple-cost criterion and a multiple-benefit criterion that cannot be reduced to a common criterion without having a common yardstick for measuring the various cost and benefit elements. It is evident that in order to arrive at a rational criterion for choice among these alternative methods, one needs, first, to enumerate the various elements of these costs and benefits or at least the important ones; and, secondly, to have (or to develop) a system of weights assigned to

---

* Professor, Population Dynamics and Political Economy, The Johns Hopkins University, Baltimore, Maryland, United States of America.

[1] See, for example, Ansley J. Coale, ed., *Economic Factors in Population Growth,* Proceedings of a Conference held by the International Economic Association, Balescure, France (New York, John Wiley and Sons, 1976); H. Leibenstein, "An interpretation of the economic theory of fertility", *Journal of Economic Literature*, vol. XII (June 1974); N. K. Namboodiri, "Prologue", *Social Forces*, vol. 54 (September 1975).

[2] J. Johnston, *Econometric Methods* (New York, McGraw-Hill, 1972).

[3] "Methods of measuring the impact of family planning programmes on fertility: problems and issues" (ESA/P/AC.7/1); see part one of the present publication.

these various elements. It is argued below that the assignment of these weights, and to some extent the full enumeration of the cost and benefit elements, cannot be developed independently of the "purpose" of evaluation; and that that "purpose" will vary greatly, depending upon the type and the life cycle of the programme under consideration, i.e., upon the intended use of the evaluation findings. Lastly, an attempt is made, for purpose of stimulating discussion, to develop a simple and preliminary typology that relates programme objectives, evaluative needs and appropriate methods of evaluating the fertility impact of family planning programmes.

The "total cost" of a given method could be conceptualized as the market or imputed value of all inputs required to produce an "estimate" of a given precision where the cost elements include:

(*a*) The type of data required—whether they are already available (e.g., service statistics or census data) or must be produced independently;

(*b*) The time needed to produce such estimate (of a given precision);

(*c*) The level of technical skills required relative to the available pool;

(*d*) The alternative uses of the data and analyses produced—comparability with other estimates in terms of scope and coverage, its future use or its utilization by other governmental agencies. Such alternative uses might justify relatively large overhead costs.

Clearly, the type and magnitude of these cost elements will depend upon the type of "estimate" required and upon its desired level of precision.

There are various reasons for examining the relationship between family planning efforts and changes in fertility. For example, the focus might be on one or more of the following purposes:

(*a*) To examine the "aggregate" effect of the programme on "aggregate" fertility changes; this objective is usually applicable when evaluating the total investment of a national effort;

(*b*) To examine the "aggregate" effect of the programme on "differential" fertility, where the focus of the programme is on serving special target population groups, e.g., high parity, rural, poor or teen-agers;

(*c*) To examine the interaction between programme and non-programme factors on fertility in order to develop optimal population strategies; this emphasis becomes important especially where there are rapid social and structural changes (e.g., Venezuela, Iraq and Iran) or random shocks (e.g., Bangladesh) which have independent influences on the basic determinants of fertility. An important example would be the interaction between health, nutrition and fertility;[4]

(*d*) To examine the relative effects of individual programme components or strategies on fertility, as, for example, trying a new contraceptive method or a new component in the delivery system, or assessing whether the effectiveness of a given method(s) has changed over time.

Furthermore, some countries (e.g., Egypt) have broadened the scope of their family planning programme to include inputs that influence fertility indirectly, e.g., influencing age at marriage; female labour force participation, education and literacy; or rural cottage industries, which makes it harder to define programme factors and to measure their independent influence on fertility, mainly because of the lack of adequate understanding of many of these linkages. It becomes difficult to identify "programme acceptors", especially through service statistics.

At one time or another during the life cycle and development of a programme, its organizers or planners will need to have answers to these various questions.

At the early stages of a programme, information about the relative effectiveness of the various methods is needed both for initial planning or target setting and for ongoing evaluation of performance. Probably the CYP method or analysis of the reproductive process, both of which focus primarily on programme acceptors and are based on service statistics and some economical retention surveys, would be appropriate at the initial stages of programme operations and when programme acceptors are reaching the anticipated target. But for adequate specification, data are not usually immediately available.[5]

However, when acceptance rates are very low and continue to be low, the efficiency of these techniques appears to be reduced. In such cases (e.g., Bangladesh and Pakistan), what is needed is to understand why programme factors do not influence fertility. Information about "non-clients", what determines their fertility norms and behaviour, becomes of primary focus. For that purpose, some theoretical frame that includes the socio-economic and biological determinants of fertility should serve as a guide for this type of research. Multivariate analyses (e.g., regression analysis, including path analysis; automatic interaction detector technique) appear to be needed for this type of approach, and probably survey samples, if adequately

[4] See, for example, W. Butz, "Socioeconomic and biomedical analysis of nutrition, birth interval and infant development", Santa Monica, Calif., Rand Corporation, Institute of Central America and Panama, 1973 (mimeographed); Hector Correa, *Population, Health, Nutrition and Development* (Lexington, Mass., Lexington Books, 1975); W. H. Mosley and L. C. Chen, "Health and human reproduction in developing countries", submitted to *Science*, April 1976; R. V. Rider, "Research issues concerning the relationship between morbidity and fertility", paper prepared for the World Health Organization Scientific Group Meeting on the Relationship Between Morbidity and Population Trends, Geneva, 5–11 December 1972.

[5] Lee L. Bean and William Seltzer, "Couple years of protection and births prevented: a methodological examination", *Demography*, vol. 5, No. 2 (1968), pp. 947–959.

designed, could be an efficient way of collecting the relevant data.[6]

When the time period of evaluation is relatively long, say, 10 years, and accordingly the expected change in fertility is relatively large, then the standardization approach along with a multivariate analysis that controls for the main socio-demographic variables might be appropriate, especially when significant changes in these variables have also occurred.

To summarize the discussion thus far, appropriate methods will depend upon programme needs and development, or, more generally, upon the way in which the findings will be utilized. It appears that evaluating the fertility impact of a family planning programme should go beyond aggregate relations—it should identify pockets of high fertility in society and their response to programme factors. It should also examine whether short-run apparent changes are delayed pregnancies or are there to stay. These are questions that cannot be answered in the abstract but need specific reference-setting (e.g., teen-age pregnancies in Jamaica, the health-nutrition-fertility relationship in Bangladesh, the State/method differential in India).

When comparing the relative effectiveness and utility of a given method, the objectives of evaluation should be indicated, for example, whether its purpose is for short-run management use (1–4 years) or for long-run planning use (say 5 or more years); and also whether the focus is on the relative efficiency of programme input or on the effectiveness of programme operation. Furthermore, the purpose of evaluation should be specified with respect to the objectives of the programme in terms of family planning only or other objectives, such as family health and welfare. A schematic illustration of a possible classification is given below.

The difficulty of choosing an evaluation method is not unique in demographic analysis; as Jaffe noted "Our articles of faith—and the public and private policies and programs which express them—remain mostly unevaluated in any formal sense".[7] The case

[6] Ismail Sirageldin, "The survey method in family planning research and evaluation", in J. F. Kantner and L. McCaffrey, eds., *Population and Development in South East Asia* (Lexington, Mass., Lexington Books, 1975).

[7] F. S. Jaffe, "Issues in the demographic evaluation of domestic family planning programs", in J. R. Udry and E. E. Hyuck, eds., *The Demographic Evaluation of Domestic Family Planning Programs* (Cambridge, Mass., Ballinger Publishing Company, 1975).

A POSSIBLE CLASSIFICATION OF EVALUATION OBJECTIVES, PROGRAMME OBJECTIVES AND CRITERIA FOR COMPARING METHODS OF MEASURING THE FERTILITY INPUT OF FAMILY PLANNING PROGRAMMES

	Objective of evaluation																			
	Short-run/management (1-4 years)										*Long-run/planning (5 + years)*									
Objective of programme and methods of measurement	*Input focus; efficiency (e.g., method specific, clients' age characteristics)*					*Population focus: effectiveness (pregnant/non-pregnant acceptors)*					*Population focus*					*Alternative investment (e.g., cost/benefit)*				
	Some selected criteria[a]																			
	Fertility variables	*Data base*	*Independence*	*Validity reliability*	*Cost*															
	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
A. Family planning only																				
1. Standardization																				
2. Trend analysis																				
3. Couple-years of protection																				
4. Components projection																				
5. Analysis of reproductive process. . . .																				
6. Simulation models																				
7. Regression analysis																				
(*a*) Area																				
(*b*) Other																				
8. Experimental designs																				
B. Family planning and other																				
1. . . 8. }																				

Note: Other considerations: 1. Age of the programme in relation to the level of participation;
2. Rate of structural change: e.g., industrialization, urbanization, health/mortality.

[a] Possible criteria to compare and evaluate the various methods of measuring the fertility input of Family Planning Programmes.

can be made, however, for population programme evaluation, but a balance should be made between perfecting a given technique and its utility. There is no question that the development of highly refined models for intra-uterine device (IUD) decay are useful if the IUD is highly accepted and could correct planners' conception about what parameters to use in their CYP, as was the case in Pakistan. In the latter case, however, it was only dealing with less than 1 per cent of the fertile women and the method acceptability was losing popularity. Analysis of the reproductive process, on the other hand, appears to have made an important contribution towards understanding changes in fertility, especially short-run changes and especially in interpreting socio-economic theories of fertility. Further empirical testing of this theoretical framework is needed (current attempts are being made in the Matlab region of Bangladesh and in Guatemala).

INTERACTION OF SOCIO-ECONOMIC CHANGES WITH FAMILY PLANNING PROGRAMMES: AN ASSESSMENT MODEL

*K. Srinivasan**

It has been a frustrating experience for family planning administrators, demographers and social scientists alike not to be able to assess, with the desired level of precision, the impact of family planning programmes on fertility, especially in the developing countries where there is a pronounced need for such an assessment. The frustration arises in part from the inadequacy of the currently available models of evaluation for dealing with the combined effects of social change and family planning programmes on fertility, and in part, the paucity of reliable data in developing countries even for the application of simplistic methods of evaluation. In recent years, there has been a spate of publications on the methodological aspects, analytical problems, case study applications and problematical issues pertaining to this topic. One comprehensive review of the current state of the art in this field, with an identification of gaps in knowledge, has recently been published under the auspices of the International Union for the Scientific Study of Population (IUSSP).[1] The major problems and issues in the evaluation of impact are also succinctly reviewed and discussed in the background paper (ESA/P/AC.7/1) prepared by the United Nations Secretariat for the meeting.

In the present paper, an attempt is made to identify and discuss some special methodological and data problems faced in the developing countries in the evaluation of the impact of family planning programmes in the context of rapid socio-economic changes and to develop an assessment model in order to cope with the problem.

As has been well recognized by now, the plethora of social, economic and communication development programmes undertaken in developing countries constantly interact with family planning programmes and lead to a variety of changes in knowledge, attitudes and behaviour related to fertility modification. Paucity of necessary data makes it difficult to quantify the changes as they occur even in the fertility levels of the population and almost impossible to isolate the net effects of family planning programmes after partialling out the effects of social and economic changes on fertility. The question arises whether even at a conceptual level, there exists what can be termed a family planning programme which can be identified as a separate entity from other developmental programmes, with separate effects on fertility that could be measured. This note attempts to highlight some crucial issues in this area of interaction, to develop a model for evaluation and to indicate a plausible method of compilation of necessary data for the application of the model.

Conceptual scheme

The demand for family planning services in a developing country is a function of a number of interdependent factors. These factors can be conveniently categorized under three major groups, recognizing that the groups are not necessarily mutually independent. The first category includes the family planning programme variables, such as inputs of money, men, materials and methods into the programmes, activities and systems of contraceptive delivery services; the second group includes attitudinal variables on people's perception of the programme, which is dependent upon the socio-cultural background of the population as well as upon the type of personnel employed in the family planning programme and the nature of facilities made available; and the third category includes the various socio-cultural and political variables in the population.

An operational definition of the family planning demand structure, which will facilitate quantification, can be made as "the proportion of eligible couples in the population at different stages of readiness to adopt family planning". From past experience with family planning programmes in developing countries, it has been found that in every community a certain proportion of eligible couples will adopt a modern method of family planning as soon as the services are offered to them; a certain proportion will do it when the programme is supported by their community leaders; a certain proportion when it is backed by a scheme of incentives and disincentives and a certain proportion only when there are changes in their socio-economic conditions. An ordered scale of readiness to accept family planning under different levels of stimuli can be considered and would form the basis for measuring the structure of demand for family planning. The three sets of factors mentioned above, which are not mutually

*Director, Population Centre, Bangalore, India.

[1] C. Chandrasekaran and Albert I. Hermalin, eds., *Measuring the Effect of Family Planning Programs on Fertility* (Liège, International Union for the Scientific Study of Population for the Development Centre of the Organization for Economic Co-operation and Development, 1975).

independent, influence the demand structure which in turn affects the number of couples accepting family planning methods, their continuation rates and fertility change. A schematic diagram of the conceptual model of structure of demand for family planning services is given below.

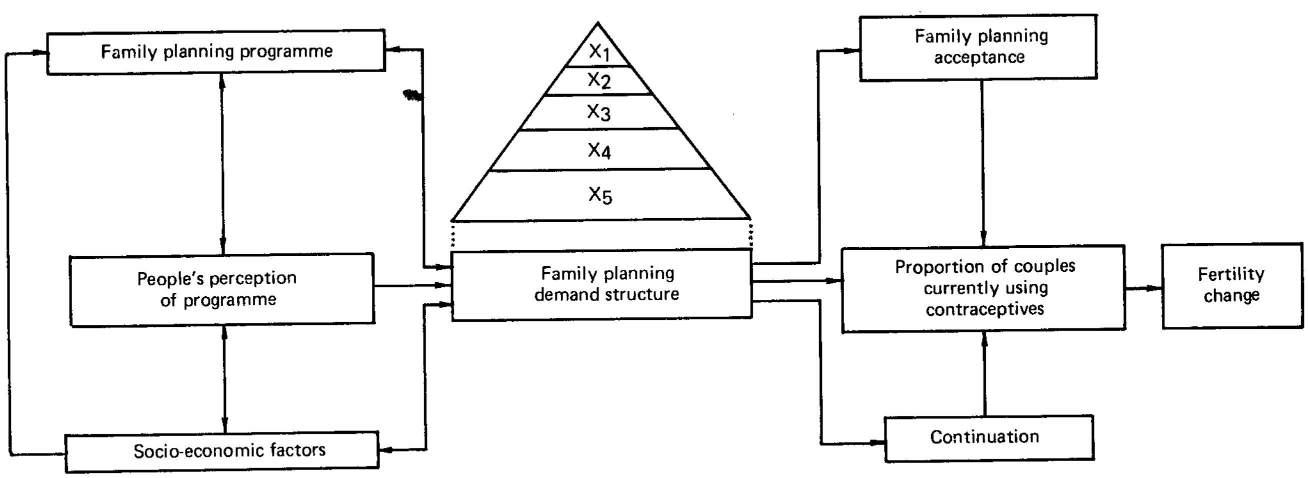
Adopting the conceptual framework mentioned above, it follows that the demand for family planning in any society changes continuously, depending upon the interaction of the family planning programme variables and the socio-economic structural variables. Further, it is obvious that the demand at any time is also a function of the demand and contraceptive practices at earlier times because the potential couples available for fresh recruitment to contraception depends not only upon the total number but upon the proportion who have been recruited earlier. In such a framework, evaluation of the impact of family planning programmes on fertility should be stochastic in nature and cannot be deterministic wherein analysis is restricted to the situational position of the variables at any one point of time or at most some variables are measured as changes during a given period of time.

There are continuous changes both in the social structure and in the programme variables which have to be taken into account in any analysis of the impact of the programme. For example, in India, many traditional customs and practices have long operated as checks on fertility. Since the distant past, the taboos on sexual intercourse on selected days for religious or social reasons, the practice of the pregnant woman going to her mother's home for delivery and staying there for a number of months after delivery, the long periods of breast-feeding of children, and the social taboos on widow remarriage have all contributed, though indirectly, to keeping down crude birth rates to about 40, which is much below the biological potential. There has been a considerable variation even in the natural fertility, fertility in the absence of wilful contraceptive practices among different groups. During the past two decades, however, owing to the processes of modernization unleashed as a part of development strategy, many of the traditional checks have begun to be released and the potential fertility in the absence of contraception can be expected to increase in many groups.

The existing techniques for evaluation of the impact of family planning programmes do not adequately deal with this problem conceptually or analytically, though they recognize the existence of such interaction. For example, the standardization approach assumes that whatever differences in fertility level between two points of time remain as a residue, after standardization for certain structural factors, such as age-sex-

Conceptual model of structure of demand for family planning services

Tentative scale of demand structure



Note:

$X_1 + X_2 + X_3 + X_4 + X_5$ = total eligible couples;

X_1 = currently contracepting couples;

X_2 = couples likely to adopt family planning with modifications in the programme but without additional inputs;

X_3 = couples likely to adopt family planning with additional inputs into the programme;

X_4 = couples likely to adopt family planning with modifications in the programme and additional inputs into it;

X_5 = never-adopters.

marital status distribution and possibly rural-urban differences, is due to family planning programmes. Whatever change in fertility that cannot be accounted for, after consideration of a few factors for which data are available, is considered to be due to family planning. Similarly, in the fertility projection approach, the fertility in the absence of family planning programmes is projected on the basis of pre-programme levels or of some previous trend extrapolated. In the component projection or analysis of reproductive process methods, serious assumptions are involved regarding the potential fertility of acceptors, fertility that would have prevailed had they not come into the fold of the programme or into family planning. Such assumptions are generally made on the basis of the pre-acceptance fertility of the acceptors or a comparison of the fertility of acceptors with non-acceptors assuming that these differences would persist in future also. The method of computing the couple-years of protection (CYP) and converting to births averted, on the basis of the formula 1 CYP = C births averted, again ignores the socio-economic variables and their effect on potential fertility. Among the existing evaluation methods, the two which explicitly take into consideration socio-economic variables and their interaction with family planning programmes are experimental designs and regression analysis. In the experimental-design approach, two populations matched for a number of social and economic variables, including fertility and mortality levels, are selected, and one is exposed to the family planning programme and the other kept as a control. Here also, there is a tacit assumption that the future trends in socio-economic factors in both the populations are nearly the same and what accounts for the differential change in acceptance rates or the fertility levels during the period of study is the programme. In regression analysis, fertility at any point of time or fertility change between two points of time in different areal units is considered to be a dependent variable and regressed on a host of exogenous variables, which include such socio-economic factors as literacy rate, rates of employment of women and mortality levels; such programme variables as family planning; and such input variables as family planning personnel per 1,000 population, number of clinics, money spent on family planning in each of the areal units etc. Areal analysis takes a geographical area, such as a county, block, district or state, as a unit for carrying out the regression or path analysis. The variance or differences that exist in the exogenous and programme factors within each areal unit and their association with fertility is ignored in the analysis. It has to be recognized that such fertility differentials on the basis of socio-economic groups are of fundamental importance in fertility change and averaging these differentials on the basis of areal units may vitiate the pattern of relationship. This possibility has been recognized in the earlier works on areal analysis.[2] In the following section an attempt is made to develop a research design which is focused on the measurement of interaction effects of social change and family planning programmes.

An assessment model

In any analysis of the effects of interaction of family planning programmes with socio-economic variables, the basic information that is needed is the probability of an eligible couple belonging to a particular socio-economic class, accepting a family planning method under a given régime of family planning programme. The variance in these probabilities has then to be divided into three portions, one due to socio-economic class, the second due to the family planning programme and the third due to the interaction between the two. A simplistic model of this procedure is presented below for the purposes of illustration.

Suppose that in any population at time t, N_t represents the total number of couples in reproductive age group and that they can be classified into k socio-economic groups, which can be placed in an ascending order from 1 to k. Let the number of couples in the ith socio-economic group be $E_i(t)$. It is further assumed, for the sake of simplicity, that the population of couples in the reproductive ages is stationery, i.e., N_t is independent of t and equals N. Let $p_i(t)$ be the probability that an eligible woman in class i will accept a family planning method at time t, and $A(t)$ the total number of acceptors at time t. It is further assumed that the proportion of a cohort of acceptors continuing the method at time t after acceptance can be defined by the function e^{-rt}.

It has to be recognized that in the model, though the number of couples with wives in age group 15–44 remain the same over time, N, the proportion in different social classes may keep changing because of development and social mobility. Let the matrix $M = (m_{ij})$ denote the probability of transition from class i to j in any unit of time. Then, the following set of equations can be formulated to express the interaction.

$$A\ (t) = \sum_i E_i^*\ (t)\, p_i\ (t) \tag{1}$$

where $E_i^*\ (t)$ denotes the number of couples at time t who are available for family planning acceptance because they had not accepted any method before or they had accepted earlier but discontinued in year t and are available for reacceptance in the same year. If data are available on the number of acceptors, $A_i\ (t)$ in any year from any social class i, then:

$$\hat{p}_i\ (t) = \frac{A_i\ (t)}{E_i^*\ (t)} \tag{2}$$

and $p_i\ (t)$ can thus be estimated for different i and different points of time or different areal units, and a matrix P of $(p_i\ (t))$ could be formed.

[2] Albert I. Hermalin, "Regression analysis of areal data", in C. Chandrasekaran and Albert I. Hermalin, eds., *Measuring the Effect of Family Planning Programs on Fertility*, pp. 245–300.

An analysis of variance could then be performed in this P-matrix to determine what proportion of the variance in p_i (t) could be attributed to the differences in social classes, what proportion is due to changes in programme inputs and the balance due to interaction. Such an analysis would be useful in determining the impact of interaction on family planning acceptance. On the other hand, if data are not available on A_i (t) but only on the total number of acceptors A (t), then p_i (t) can be estimated indirectly adopting the following procedure.

Let the row vector E^* (t) denote $(E^*(t), E_2^*(t) \ldots E_k^*$ $(t))$ giving the number of eligible couples in different classes in year t. Then it can be shown that:

$$E^* \ (t) = E^*_q \ (t-1) \ M + e^{-rt} \ (e^r - 1) \sum_{s=1}^{t=1} e^{rs} \ A \ (s) \ (M)^{t-s} \qquad (3)$$

where E_q^* $(t-1)$ denotes a row vector $(1 \times k)$ of eligible couples not accepting any family planning method in time $(t-1)$ and E_{qi}^* $(t-1)$ equals E_i^* $(t-1)(1-p_i$ $(t-1))$ and A (s) is the row vector $(1 \times k)$,

$$(A_1 \ (s), A_2 \ (s) \ldots\ldots A_k \ (s))$$

where A_i $(s) = E_i^*$ (s) p_i (s).

The first term in the expression for E^* (t) is the number of couples who were eligible at time $(t-1)$ but did not accept the programme in $(t-1)$ and thus were eligible in time (t); and the second term gives the number of couples who had accepted the programme earlier but discontinued during year t, making themselves eligible again. Now if r, M and E (t) are available and $A(t)$ is known for $t = t_1, t_2, \ldots\ldots$, equation (1) together with equation (3) becomes a system of non-linear equations in $(p_i$ $(t))$. These could be solved and the matrix (P) obtained. This matrix can be used for an analysis of variance to determine the effect of interaction.

The method given above is capable of being developed further, relaxing the assumptions of stationariness of the population. A preliminary application of the model is currently being attempted from the data collected through a survey in Karnataka State, India.

NOTES ON CAUSAL RELATIONSHIPS IN MEASURING FERTILITY CHANGE

*H. Bradley Wells**

Identification and quantification of the effects of family planning programmes on fertility are difficult because it is usually impossible to take all related variables into account. In this respect, the problem of measuring the impact of family planning (or any other variables, e.g., economic developments) on fertility is analogous to measuring the impact of a health programme on mortality. In either case, one usually has only data from observational studies, such as census, vital statistics and surveys, rather than results of *bona fide* experimental studies in which subjects are randomly allocated to treatments.

Family planning programmes, at best, work in conjunction with numerous other factors to reduce and/or to increase fertility. In this paper, an attempt is made to illustrate the difficulties inherent in choosing models and to point out the need for improvement of the quality of data. Admittedly, the discussion cannot be presented in complete detail in a brief statement. Perhaps it will be sufficient to stimulate further discussion on the problems of decision making in the face of considerable uncertainty.

MODELS OF CAUSAL RELATIONSHIPS

Even if reliable measures of time trends in fertility levels and in the many variables that affect the fertility levels of a society or of a stratum or category within that society[1] were readily available, demographers and other data analysts would have much difficulty in agreeing on models showing causal relationships between all of these variables and fertility. These difficulties exist whether the model relates to individual women or to groups, e.g., all women in a county. First, there would be some disagreement on how many variables and the level of detail at which they should be included in any model; and, secondly, there would be disagreement regarding the functional form(s) of relationships. (Note that the foregoing statement implicitly assumes that problems of definition and measurement do not exist although in the real world they are major deterrents to determining causes of fertility change. They are considered below in the section on data availability.)

The "intermediate variables" of Davis and Blake,[2] which Freedman included under his classification one, may be used to illustrate the difficulties in specifying the level of detail required for independent variables in a model. This list of "intermediate variables" would include programmes as well as other means of limiting fertility. The intermediate variables may be classified into 11 groups as follows (by taking some liberties with the original):

A. *Time spent in sexual unions by women:*

1. Age at entry into sexual union;

2. Proportion never entering sexual union;

3. Time spent outside union due to death of husband, divorce, separation or desertion;

B. *Exposure to intercourse within union:*

4. Voluntary abstinence;

5. Involuntary abstinence;

6. Coital frequency (excluding periods of abstinence);

C. *Conception variables:*

7. Fecundity or infecundity as affected by involuntary causes;

8. Use or non-use of contraception;

9. Voluntary fecundity or infecundity (sterilization etc.);

D. *Gestation variables:*

10. Foetal mortality from involuntary causes;

11. Foetal mortality from voluntary causes (abortion).

Continuing with the present assumption that reliable data are available, one should consider what level of detail one would use in an analysis. This issue would, of course, be related to the functional model postulated. Assume further that one has been asked to determine the magnitude of the effect of variable C (8) on fertility in country X at time t.

Since the above-mentioned data are available over time and in the detail required, it remains only to make three other general decisions regarding:

(*a*) The level of detail in the response or dependent

---

*Professor of Biostatistics, University of North Carolina, Chapel Hill, North Carolina, United States of America.

[1] Ronald Freedman, "The sociology of human fertility: a trend report and bibliography", *Current Sociology*, vol. X-XI (1961–1962).

[2] Kingsley Davis and Judith Blake, "Social structure and fertility: an analytic framework", *Economic Development and Social Change*, vol. IV (April 1956), pp. 211–235.

variable(s), i.e., what measure(s) of fertility should be analyzed; whether it is appropriate to analyze the total fertility rate or should age-specific fertility rates be analyzed; whether age-parity specific rates should be analyzed instead; and how fertility measures from previous years should be handled, i.e., whether they are response variables or intervening variables. (The present writer thinks they are response variables, as is explained below.) Obviously, multivariate response variables are more difficult to analyze than a univariate summary measure, such as the total fertility rate;

(*b*) The level of detail in the independent variables listed above as well as other independent variables, including: social norms about the intermediate variables listed above; family size preferences; societal rewards and punishments which support or negate social norms regarding family size and the intermediate variables; mortality levels; net migration levels; and environmental factors.[3] Clearly, this list will be longer than the list of intermediate variables—how much longer will depend upon the analyst and the model. Family planning programme variables other than contraceptive use, e.g., education and propaganda, can be included here.

(*c*) The model(s) must be specified showing how the independent variables relate to the dependent variables. Models for individual women would be somewhat different than models using areal aggregate measures. A general model can be specified as follows:

$$R = F(A,B,C,D,AB,AC,AD,BC,BD,CD,\\ ABC, ABD,BCD,ABCD)$$

where R = a matrix of specific fertility rates in the required detail by age and parity and by calendar year;

F = an indicator that R is a function of the variables in parentheses which are defined as follows:

A = a time-series matrix of intermediate variables, excluding use or non-use of contraception, and their interactions with one another;

B = a time-series matrix of measures of use and non-use of different types of contraception. These types should be further subdivided into contraception provided through the programme and those obtained from other sources if the effect of programme contraception is to be measured;

C = a time-series matrix of the other dependent variables, excluding any non-contraceptive programme input variables;

D = a time-series matrix of the non-contraceptive family planning programme input variables, such as levels of staffing, educational efforts to change social norms regarding acceptance and use of contraception. It is necessary to include these variables if the model is to provide a measure of programme effects on non-programme use of contraception;

AB and other two-factor terms refer to first-order interactions between the four sets of variables, ABC and other three-factor terms refer to second-order interactions and $ABCD$ refers to third-order interactions.

Even if one restricts the model to a single response variable, say, the total fertility rate, and the other variables to one per set and a single year, perhaps year $t-1$, there still are 14 possible effects to be estimated after specifying the analytical model. Ten of these effects are, of course, the interaction effects which in practice are often ignored. Thus, a major problem in any analysis is deciding not only what dependent variables to include but what interaction effects can be included (or eliminated).

Nevertheless, it is useful and necessary to utilize less than a full model in most situations. It should be re-emphasized that careful consideration of all factors in the full model at the design and analysis phases of study is required regardless of whether they are included in the final model(s). One must not overlook the possibility that excluded variables may influence results.

Replication

The most acceptable method of determining cause and effect requires randomized assignment of individuals or groups of individuals in controlled experiments. This type of experiment is difficult to do in measuring programme impact on fertility itself. Hence, even if reliable observation, as opposed to experimental, data are available in the detail required, and one agrees on the model(s) relating independent variables to fertility, the conclusions must be limited to the detection of relationships and, hence, to developing additional hypotheses to be tested rather than estimating with specified levels of confidence how much the programme reduced fertility. This limitation applies to any statistical model at this point in time.

In observational studies, replication of different tests of the same hypothesis must be repeated using different models and different sets of data in different circumstances in order to systematically compile evidence for the effects of family planning programmes with different levels of the other covariables. Alternative hypotheses which could explain observed results through excluded variables should be tested to rule out the possibility of hidden effects of interactions with included variables.

Availability and quality of data

The foregoing discussion was intended to illustrate difficulties faced when reliable data are available. In fact, the data required for a desired model are seldom

[3] R. Freedman, *loc. cit.*

all available at a given time and place. Hence, one bravely makes do with what is available. This is acceptable as an interim approach pending more systematic planning and collection of data and analysis in a forward-looking time frame in order to examine the question of causal relationships from as many angles as possible. Even so, the number of dependent variables that can be included in a specific analysis is usually severely limited.

In the "make-do" mode, in the "forward-time frame" mode and even in the experimental-design mode, the magnitude of measurement errors and biases in the data are usually unknown. This writer strongly suspects that the inconsistency of results from different times, places and/or analysts may well be due to different levels of measurement errors and/or biases for the different variables.

It is becoming fashionable to take account of sampling errors, and procedures for doing this are fairly straightforward; indeed, the only justification of tests of significance in observational studies is to determine the probability that observed results could have occurred by chance in an experimental design with the same number of observations. Techniques for detection of and adjustment for measurement errors are not as well developed as for sampling errors.

There is much to be done in determining how measurement errors influence demographic analysis and decisions under different levels of fertility trends and different levels of independent variables. In the meantime, analysts would be well advised to consider whether measurement errors of included variables might also have "explanatory" effects in a particular model.

كيفية الحصول على منشورات الامم المتحدة

يمكن الحصول على منشورات الامم المتحدة من المكتبات ودور التوزيع في جميع انحاء العالم · استعلم عنها من المكتبة التي تتعامل معها أو اكتب الى : الامم المتحدة ،قسم البيع في نيويورك او في جنيف ·

如何购取联合国出版物

联合国出版物在全世界各地的书店和经售处均有发售。请向书店询问或写信到纽约或日内瓦的联合国销售组。

HOW TO OBTAIN UNITED NATIONS PUBLICATIONS

United Nations publications may be obtained from bookstores and distributors throughout the world. Consult your bookstore or write to: United Nations, Sales Section, New York or Geneva.

COMMENT SE PROCURER LES PUBLICATIONS DES NATIONS UNIES

Les publications des Nations Unies sont en vente dans les librairies et les agences dépositaires du monde entier. Informez-vous auprès de votre libraire ou adressez-vous à : Nations Unies, Section des ventes, New York ou Genève.

КАК ПОЛУЧИТЬ ИЗДАНИЯ ОРГАНИЗАЦИИ ОБЪЕДИНЕННЫХ НАЦИЙ

Издания Организации Объединенных Наций можно купить в книжных магазинах и агентствах во всех районах мира. Наводите справки об изданиях в вашем книжном магазине или пишите по адресу: Организация Объединенных Наций, Секция по продаже изданий, Нью-Йорк или Женева.

COMO CONSEGUIR PUBLICACIONES DE LAS NACIONES UNIDAS

Las publicaciones de las Naciones Unidas están en venta en librerías y casas distribuidoras en todas partes del mundo. Consulte a su librero o diríjase a: Naciones Unidas, Sección de Ventas, Nueva York o Ginebra.

Printed in U.S.A.
36522—July 1978—5,400

Price: $U.S. 11.00
(or equivalent in other currencies)

United Nations publication
Sales No. E.78.XIII.2
ST/ESA/SER.A/61