Statistics in Fertility Research: Value and Limitations (1983)

Pages
40

Size
6 x 9

ISBN
0309326494

Brillinger, David R.; Panel on Fertility Determinants; Committee on Population and Demography; Commission on Behavioral and Social Sciences and Education; National Research Council

COMMITTEE ON
POPULATION AND DEMOGRAPHY

Report No. 19

Statistics in Fertility Research

Value and Limitations

David R. Brillinger

Panel on Fertility Determinants
Committee on Population and Demography
Commission on Behavioral and Social Sciences
and Education
National Research Council

NATIONAL ACADEMY PRESS
Washington, D.C. 1983

NOTICE: The project that is the subject of this report was approved by the Governing Board of the National Research Council, whose members are drawn from the councils of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine. The members of the committee responsible for the report were chosen for their special competences and with regard for appropriate balance.

This report has been reviewed by a group other than the authors according to procedures approved by a Report Review Committee consisting of members of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine.

The National Research Council was established by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and of advising the federal government. The Council operates in accordance with general policies determined by the Academy under the authority of its congressional charter of 1863, which establishes the Academy as a private, nonprofit, self-governing membership corporation. The Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in the conduct of their services to the government, the public, and the scientific and engineering communities. It is administered jointly by both Academies and the Institute of Medicine. The National Academy of Engineering and the Institute of Medicine were established in 1964 and 1970, respectively, under the charter of the National Academy of Sciences.

Available from

AV NATIONAL ACADEMY PRESS (Publication-on-Demand)
2101 Constitution Avenue, N.W.
Washington, D.C. 20418

Printed in the United States of America

PANEL ON FERTILITY DETERMINANTS

CH W. PARKER MAULDIN (Chair), The Rockefeller Foundation, New York
ELZA BERQUO, Centro Brasileiro de Analise e Planejamento, Sao Paulo, Brazil
WILLIAM BRASS, Centre for Population Studies, London School of Hygiene and Tropical Medicine
DAVID R. BRILLINGER, Department of Statistics, University of California, Berkeley
V.C. CHIDAMBARAM, World Fertility Survey, London
JULIE DAVANZO, Rand Corporation, Santa Monica
RICHARD A. EASTERLIN, Department of Economics, University of Southern California, Los Angeles
JAMES T. FAWCETT, East-West Population Institute, East-West Center, Honolulu
RONALD FREEDMAN, Population Studies Center, University of Michigan
DAVID GOLDBERG, Population Studies Center, University of Michigan
RONALD GRAY, School of Hygiene and Public Health, The Johns Hopkins University, Baltimore
PAULA E. HOLLERBACH, Center for Policy Studies, The Population Council, New York
RONALD LEE, Graduate Group in Demography, University of California, Berkeley
ROBERT A. LEVINE, Graduate School of Education, Harvard University
SUSAN C.M. SCRIMSHAW, School of Public Health, University of California, Los Angeles
ROBERT WILLIS, Department of Economics, State University of New York, Stony Brook

ROBERT J. LAPHAM, Study Director

COMMITTEE ON POPULATION AND DEMOGRAPHY

ANSLEY J. COALE (Chair), Office of Population Research, Princeton University
WILLIAM BRASS, Centre for Population Studies, London School of Hygiene and Tropical Medicine
LEE-JAY CHO, East-West Population Institute, East-West Center, Honolulu
RONALD FREEDMAN, Population Studies Center, University of Michigan
NATHAN KEYFITZ, Department of Sociology, Harvard University
LESLIE KISH, Institute for Social Research, University of Michigan
W. PARKER MAULDIN, Population Division, The Rockefeller Foundation, New York
JANE MENKEN, Office of Population Research, Princeton University
SAMUEL PRESTON, Population Studies Center, University of Pennsylvania
WILLIAM SELTZER, Statistical Office, United Nations
CONRAD TAEUBER, Kennedy Institute, Center for Population Research, Georgetown University
ETIENNE VAN DE WALLE, Population Studies Center, University of Pennsylvania

ROBERT J. LAPHAM, Study Director

NOTE: Members of the Committee and its panels and working groups participated in this project in their individual capacities; the listing of their organizational affiliation is for identification purposes only, and the views and designations used in this report are not necessarily those of the organizations mentioned.

CONTENTS

PREFACE

This report discusses the value and the limitations of statistics in fertility research. Prepared by David R. Brillinger, it is designed to complement the other methodological and substantive reports of the Panel on Fertility Determinants and to focus attention on issues involving the use of statistics that too often are brushed aside in fertility research, as in other areas of social science research.

With the rapid expansion in world population, fertility and its determinants have been urgent topics for research in recent decades. Attempts to control population growth have focused on reducing fertility, with some apparent effect. The peak rate of growth in the world's population has now been passed, although growth is still at a high level in almost all the developing countries. In absolute numbers, the increase in the world's population continues to rise; according to United Nations medium projections, more people will be added each year for the next 50 years than were added in 1980.

The Panel on Fertility Determinants was created by the Committee on Population and Demography and the Commission on Behavioral and Social Sciences and Education of the National Research Council in response to a request from the Agency for International Development. In addition to this report, the panel has prepared studies of several developing countries, a few illustrative cross-national analyses, and two volumes that review the research evidence about determinants of fertility differentials and fertility change in the developing countries. To encompass such research, the panel was of necessity a heterogeneous group, including scholars from several disciplines: anthropology, demography, economics, epidemiology, psychology, sociology, and statistics.

Part of the background for the panel reports was provided by previous work of the Committee on Population and Demography and its other panels to try to determine actual fertility levels and trends in selected developing countries: that work, also supported by the Agency for International Development, is detailed in a series of country reports from the National Academy Press; the demographic methodology developed for the work is presented in a volume issued by the United Nations and several reports on data collection, all prepared by the committee.

During its deliberations, the Panel on Fertility Determinants wanted a critical review of the uses of statistical techniques in fertility research. In response, this report examines some of the common statistical methods used by fertility researchers, focusing on their limitations. To guide the review, the panel provided a selected bibliography of recent papers on the topic; that bibliography is reproduced in the Appendix with annotations by the author of this report.

It became apparent early in the preparation of this review that there are few, if any, statistical problems unique to fertility studies; rather, the focus must be on the values and limitations of the statistical techniques used throughout the social sciences. Because this report is limited in length, it is necessarily somewhat incomplete, but it points out some problems that fertility determinants researchers need to consider and potential solutions to those problems, thus helping to make their work more convincing. Although much of the report's content will be well known to statistically inclined demographers, the newer references may nevertheless be helpful.

Some might view this report as an attack on the use of formal analytic methods in fertility research. They should not. Although the discussion here describes many limitations to the analytical approach, the purely qualitative descriptive approach obscures critical issues; it either does not go far enough (if it consists of merely a record of observations) or is overly subjective (if it goes too far with imprecise concepts). Tight argument is necessary if the conclusion of research are to be convincing to properly skeptical scientists and policy makers.

The panel expresses sincere appreciation to several persons for contributions to this report. In addition to

INTRODUCTION

Variations in human population numbers and composition are of concern to all, especially to those involved in planning and policy. Fertility, a major element of such variations, differs greatly with age, socioeconomic class, time, location, and a number of other factors. In attempting to explain the "causes" of these differences, fertility researchers try to provide reliable information about the current state of fertility, as well as careful projections of trends and of the effects of changes in exogenous variates. Knowledge in this field is a balance of observation and theory; some is largely descriptive, while some is based on high-powered analytic procedures. Workers in the field often use statistics to justify their methods and conclusions, addressing questions of interest by the gathering and analysis of data. In drawing conclusions from the data collected, they typically use the arguments of statistical inference. If statistics is to be of value to these researchers, its limitations must be recognized and statistical techniques themselves critically evaluated. This is the objective of the present paper.

First, the statistical approach is described--its variations and its different roles in research. Next, some general statistical tools are examined, including models and various data collection techniques. This is followed by brief descriptions of a number of specific statistical methods. The paper concludes with a summary discussion and some recommendations for the field of fertility research. The Appendix presents an annotated bibliography of recent fertility research literature items used in the preparation of this paper.

THE STATISTICAL APPROACH

The field of statistics plays several roles in the empirical sciences. The traditional one is that of efficiently summarizing large and complex sets of data of the sort frequently produced by fertility studies. The summarization involves the computation of characteristics (for example, sample means and variances) whose properties are fairly well understood, followed by a search for patterns. In this process, existing knowledge and theories are taken into account, as are the needs of the particular situation.

A more important role for statistics is that of guiding the whole flow of an empirical study: the initial formulation of goals and hypotheses in the light of existing knowledge; the effective collection of information; the formulation of strategies for analysis; the choice of tools for insightful analysis; and the selection of indicators to be used to draw conclusions and generate new hypotheses leading to the reanalysis of existing data and the design of new studies. Tukey (1980) has specified two basic modes in which statistics operates--exploratory and confirmatory--with the strength and character of any inferences made at the end of a study depending strongly on the mode adopted. In exploratory data analysis, as the name suggests, the researcher allows the data to generate suggestions, hypotheses, and statistical (mathematical) models; a flexible attitude is adopted while employing all the available analytic and computing skills (see Tukey, 1977; McNeil, 1977; Velleman and Hoaglin, 1981). The intention is to make discoveries among the data, recognizing that artifacts may be turned up as well. The techniques used are meant to be robust (not strongly dependent on specific distributional assumptions) and resistant (not overly affected by unusual observations), and to summarize the data no more than is necessary. (For an example of an exploratory analysis of fertility data, see the section by Sykes in Mauldin and Berelson, 1978.) Confirmatory data analysis approaches the data with specific hypotheses and a specific stochastic model (enforced by the experimental design, whenever possible). In contrast with exploratory analysis, in the confirmatory case the model is inflexible, rather than the data. Generally, confirmatory analysis is followed by exploratory analyses, some of which may lead to later confirmatory studies; this is of course the usual iterative pattern of scientific research. In the case of empirical research on fertility, the mode most frequently used is exploratory: models are constructed and assumptions are checked with the data at hand. This generates a concern that the conclusions are (mentally) fragile.

Another role of statistics is the formal analysis and description of the variations present in data sets; this includes the realistic assessment of the stability of meaningful quantities extracted from the data. To this end, statisticians have developed specific variance estimates for particular models and data quantities. They have also developed general procedures useful in a broad class of situations. Foremost among these is the "jackknife," which involves carrying through the computations of interest on representative subsets of the data and then combining the results of the separate computations in a particular fashion (see, e.g., Mosteller and Tukey, 1977:Ch. 8). Fertility researchers usually try to provide measures of the statistical uncertainty of their computations, especially in the case of sample survey-based quantities; in fact, for real use to be made of any estimate, such a measure must always be provided.

Perhaps the key role of statistics in the scientific process is to restrain researchers from viewing their results with unrealistic enthusiasm and from drawing improperly qualified conclusions. Assumptions and procedures must be evaluated critically before conclusions can be drawn; subjective elements must be distinguished from objective ones; anecdotes and specific cases must be treated as such. There is a need to counter the great temptation in fertility research to describe factors as causative that have really only been established as correlative (being based on observational studies alone). Statistical criticism of fertility studies is not meant to be destructive. Rather, its call for the use of formal procedures for drawing conclusions and the making of inferences is meant to strengthen fertility research.

It should be noted that fields other than statistics have much to offer fertility researchers in the way of quantitative analysis and the logic of inference. For example, econometrics is concerned with complex multivariate (economic) systems and has developed and investigated powerful techniques for dealing with systems of simultaneous equations and errors of measurement. The economist's supply/demand approach also appears pertinent to many fertility discussions. The field of biometrics, concerned with the description and analysis of biological phenomena, uses procedures for involving theoretical background (e.g., genetics) in empirical analyses that are worth noting. Finally, epidemiology has had to wrestle extensively with the problems of drawing conclusions from observational (versus experimental) studies and of causation versus correlation.

SOME GENERAL STATISTICAL TOOLS

MODELS

Models play an essential role in contemporary fertility research. The models employed vary widely in complexity and subtlety and have been put to diverse uses. This section reviews some of the meanings, uses, and limitations of modeling.

Scientific workers attach a broad array of meanings to the words "model" and "modeling." (Indeed, Suppes in 1960 reviewed 16 different definitions of the scientific notion of model.) To some, a model is simply a framework of words used to organize thoughts and facts. To others, it is a highly complicated computer program for simulating a situation of interest and may have taken years to develop. It is meant to represent a real process, but to prove useful it must simplify that process; it must resemble reality, yet be more easily handled. Some would view a flow chart or block diagram as a model, while others would demand a total description by precise mathematical relationships. In the former case, conclusions are reached by thinking through the consequences of the steps of the flow chart; in the latter case, conclusions are drawn based on the values of numerical quantities derived from the model. Models may be deterministic or stochastic, with their estimation and use depsnding strongly on which sort is at hand. A model may be generated by theoretical reasoning or by empirical analyses, although the most common are a combination of these two. A model may be static (cross-sectional), referring to a single point in time, or dynamic (longitudinal), describing the evolution of the variates of interest over time. The types of conclusions produced by these two forms vary greatly, as do the techniques used in their development.

Fertility researchers have provided descriptions in block diagram form of the essential features of human fertility and the components that affect it (see, e.g., Freedman, 1967; WFS Central Staff, 1977; Bulatao et al., 1983). Such block diagrams help suggest data to be collected, experiments to be organized, and analyses to be carried out, as well as the authors' opinions of the important causal directions and relationships involved; they are meant to provide some qualitative understanding. On other occasions, fertility researchers have provided complicated analytical models of the fertility process through symbols and equations. These require statistical implementation for estimation, testing, use in forecasting, and so on (see, e.g., Heckman and Willis, 1976; Mode, 1975; Hermalin and Mason, 1980).

There are a number of formal statistical models and associated computer packages that are of great use in fitting empirical models to data. These include multiple regression (see, e.g., Mosteller and Tukey, 1977), the log linear for contingency tables (see Bishop et al., 1975), the generalized linear interactive model (GLIM) (see Baker and Nelder, 1978), and proportional hazards (see Kalbfleisch and Prentice, 1980). (GLIM is available from the Royal Statistical Society; the others are in BMDP.)

Models and their fitting, or identification, have inherent limitations. First, because the assumptions on which the models are based cannot be expected to be exactly true, conclusions drawn from work with models cannot be expected to be exactly true. The model cannot be expected to be better than the data on which it is based. Effort spent auditing fertility data is well worthwhile since, even if the data are left unaltered, conclusions can be drawn with greater confidence. Residual plots are one elementary way to check a variety of assumptions and are part of modern regression packages. It may be noted that much contemporary statistical research is devoted to developing procedures that are insensitive to moderate departures from assumptions. Use of such techniques as robust regression is clearly called for in fertility work (see Mosteller and Tukey, 1977: Section 10F). Such nonsubjective procedures are needed, for if one simply rejects extreme data points, one will make the apparent inherent errors seem too small. Also, if the data set is large, as is the case with much fertility research, one cannot realistically scan the data for strange values.

The next difficulty is with fitting a model: there may be insufficient data; the data may be autocorrelated, truncated, or missing (see Dempster et al., 1977, for a general method useful in addressing this last problem);

the data may be subject to measurement error; feedback may be present; the model may be so irregular and relationships so nonlinear that no reasonable fitting procedure suggests itself; or the model may be too complicated for the computing facilities at hand. When difficulties such as these are encountered, the model is often simplified, with the consequence that the assumptions are then farther from being true.

Once a model has been fit, difficulties frequently arise that complicate its use. For example, values of exogenous variates may have to be assumed before forecasts can be constructed. Provision of adequate uncertainty measures for derived values may be analytically intractable. The researcher might wish to manipulate the equations in a logical fashion, e.g., interchanging Xs and Ys; such manipulations may be inconsistent with the fitting procedures employed.

Finally, one all too common occurrence with data sets is that several models appear to fit the data equally well. This serves as a clear warning to those trying to interpret estimates of specific parameters of an analytic model. These difficulties as they relate to the interpretation of regression parameters are discussed further below.

On the positive side, it may be remarked that there is now a growing collection of results and procedures for use in the validation of models (see, e.g., Gass, 1980). Many of these procedures appear applicable to analytic fertility models.

DATA COLLECTION TECHNIQUES

There are two principal types of studies through which fertility data are collected--observational and experimental. These will be described shortly, but briefly in the former, the assignment of treatments to units is beyond the researcher's control. This causes substantial difficulties of interpretation. In the latter, the investigator controls treatment assignment.

The subsections below briefly describe several types of study, emphasizing some cautions that must be exercised in their application. In this connection, it is worth repeating that most of the difficulties mentioned below and found in fertility research generally are common features of research in the social and human sciences.

Observational Studies

If the results of an observational study are treated as purely descriptive, no controversies arise. However, researchers generally try to get beyond simple description in their collection and analysis of data. Their objective may be to explain some response in terms of predictor variables, perhaps providing measures of the importance of the various predictors. More often, they seek understandings of basic mechanisms, in the present context, the mechanisms of fertility. These may be used for forecasting, for regulation, or for policy recommendations. It is important to clarify at the end of a study just which inferences may validly be drawn, which are plausible in light of the data, and which are pure speculation. In observational studies, difficulties in reaching strong conclusions result from causal factors going unmeasured, key factors being outside the researcher's control, and "observed" factors being subject to large measurement errors. In fertility research, these difficulties appear to arise often.

In the case of human fertility, the three key sources of data are censuses, civil registrations, and sample surveys (see National Research Council, 1981). All of these have substantial limitations for the worker seeking to draw causal conclusions. Censuses may miss individuals in a biased fashion and may not record pertinent variates; individuals may provide incorrect information. Data obtained by civil registration suffer from the same difficulties and can often generate conflicting interpretations. Well-designed sample surveys are no panacea. An effect may be due to a factor of relevance, or it may be due to the choice of frame. Moreover, one must always keep in mind the distinction between the target population (i.e., the population to which one wishes to extend inferences) and the sample population; these are rarely the same. Another important difficulty with sample surveys is that the results are subject to sampling fluctuations, and some measure of the size of these must be computed before conclusions can properly be drawn. If the survey is complex, confidence intervals and hypothesis tests may not be easily constructed (though some useful general construction procedures are discussed below).

Observational studies may be cross-sectional or time series. They may also be retrospective or prospective. The literature of epidemiology (see MacMahon and Pugh, 1970, for example) contains some discussion of the benefits and drawbacks of these various forms.

In their use of observational studies, fertility researchers want to provide suitably qualified conclusions couched in careful language; they want to speak clearly and truthfully. Observational studies play an essential role in suggesting causal factors. In some cases, it will be appropriate to act as if a factor is causal even when this has not been established rigidly yet (such as a mother's smoking reducing the birthweight of a baby); however, the researcher needs to delineate these cases carefully.

Experimental Studies

In an experimental study, explanatory variables are chosen by the researcher, and data are generated by applying these treatments to units. The pertinent variable, or variables, may be changed at will by the investigator. The need for experiments in the fertility field, where one so often has questions concerning causal issues, is well illustrated by the following remark (Box, 1966:629): "To find out what happens to a system when you interfere with it you have to interfere with it (not just passively observe it)." One branch of fertility research, namely family planning studies, has made extensive use of experiments and reached definitive conclusions in a number of cases.

The primary difficulty with observational studies is that changes may be resulting from some outside factor that is also causing corresponding changes in a proposed explanatory variable. By changing the explanatory variable independently, one can break that connection.

Another potential problem is that naive experiments in which treatments are not assigned randomly have a high risk of being invalid if bias or self-selection enters into the assignment of treatments to units. A classic example of this is provided by the Lanarkshire milk experiment (Student, 1931), in which the value of giving milk to children at school was studied. In the course of the experiment, the teachers apparently tended to give the less robust children the milk; the value of the milk treatment hence remained in doubt. If the units to which the treatments were applied were identical, randomization would not be needed; however this is seldom if ever the case. By applying treatments to units in a formal random fashion, biases are avoided on average. An investigator may then be willing to infer causation because many individual factors have been controlled, and because randomization has eliminated systematic effects due to uncontrolled factors. Gilbert et al. (1975) describe many medical and social experiments, some randomized, some not. They find that even well-executed nonrandomized studies often have conflicting interpretations, even after large, expensive, time-consuming evaluations. On this basis, they argue forcefully for the practicality and necessity of randomized studies.

Randomization alone is not a cure-all. Other aspects of a study must be carefully controlled. For example, the researcher must avoid the Hawthorne effect, in which people who know they are being treated differently do "better" for that reason. Studies often need to be blind, or double-blind, if valid conclusions are to be drawn. A substantial issue related to random experiments with humans is whether it is ethical to give an individual a treatment (for example a poorer diet) when it is virtually certain that a better treatment is available (see Gilbert et al., 1975). "Unethical" experiments may be applied to animal and insect populations; however, the researcher then has the problem of the extent to which inferences drawn may be extended to the human case.

On balance, it appears that fertility research could benefit from many more randomized controlled studies. Such studies appear practical in a wide variety of circumstances (more often than might be expected initially). The results are bound to be clearer than the findings from observational studies.

Comparative Studies

Comparative studies are parallel investigations of somewhat similar populations. They are both common and valuable in fertility research. The World Fertility Survey is a primary example. A comparative study may be either observational or experimental, randomized or not. The issues discussed above may all arise (see Freedman, 1979; Berquo, 1981).

The data used for comparative analysis are generally quantities computed for the units (e.g., countries) being compared (i.e., through higher-order analysis). The researcher must therefore adjust for systematic differences in background variables (since the units are not randomly constituted.) Comparative analyses are also often secondary, that is, making use of the data and results of others. The primary data collector and analyzer thus has the responsibility of specifying the limitations of the results (for example, providing sampling errors), while the secondary researcher has the responsibility of not forgetting those limitations.

Other Studies

Mosteller (1977) lists a number of other types of studies and methods beyond those considered above, most, if not all, of which have already played a role in fertility research. These include theory, simulation, sample surveys, regression models, quasi-experiments, management information systems, guessing, data banks, and cost-benefit analyses. Mosteller (1977:13) comments that these methods "have the weakness that they compare different situations as they stand but do not actually make changes in treatment in the field and observe their effect." In each case, recognition and statement of the limitations of the study technique employed are necessary and proper.

SOME SPECIFIC STATISTICAL METHODS

A number of specific statistical techniques and procedures are especially important in fertility research. These are described in the subsections below.

REGRESSION ANALYSIS

The importance of linear regression analysis in observational and experimental studies cannot be overstated. Most commonly, the explanatory variables are assumed to be measured exactly and given in the matrix X. The values of the dependent variable are assumed given by $y = Xb + e$, with b an unknown parameter, and with the errors of mean 0, constant variance and given by e. The entries of e are assumed independent of each other and of X. The value of b is estimated by ordinary least squares. Expressions exist for estimating the variance of b and for constructing confidence intervals and testing hypotheses. When these assumptions are satisfied and the matrix $X'X$ is not near singular, serious difficulties of estimation and interpretation do not arise. However, difficulties do arise when one moves outside this framework.

In fertility analysis and the social sciences generally, it is traditional to try to interpret the values of the estimates of the individual coefficients, and if not the values, their signs. The problem is that the meaning of a coefficient is strongly dependent on which other explanatory variables are considered simultaneously. The most one can say about a particular coefficient is that it tells about the apparent effect of its corresponding variable in the presence of the particular other explanatory variables that have been used. The practical difficulty is that unused, even unmeasured (lurking), variables often play a role in setting the value of y. This leads to numerous paradoxical--and incorrect--results (see Mosteller and Tukey, 1977:Ch. 13; Box, 1966; Joiner, 1981; Yates, 1981:Ch. 8). If the results of the regression are used to predict values of the dependent variate, the problem is not too great; it is when the results are used for explanation and interpretation that the most trouble arises.

Further complications come when the X's are subject to measurement error (being proxies for some key variable) and when X and e are correlated. Simultaneous equation methods (of modern econometrics) exist for dealing with such problems, in part. It must be remembered, however, that the justifications of these methods are asymptotic (based on large samples). The estimates are not ordinary least squares; they can differ from the ordinary least squares estimates substantially, as can the appropriate procedures for deriving measures of their uncertainty.

There are other difficulties as well: missing values, bad values (outliers), autocorrelation of the errors, and nonlinearity. Fortunately, a number of procedures and computer programs have been developed recently to help address such problems (see Mosteller and Tukey, 1977).

For the fertility investigator, the key point is that, although regression is a most useful technique, its justification is based on a number of critical assumptions. The plausibility of these assumptions for the situation at hand must be examined before the investigator can draw conclusions with any real confidence. Regression is most useful for approximating a "nice" function in a region covered by data values. Using regression to understand mechanisms and phenomena is quite another matter, however. In particular, it seems that hardly any variable can be taken at face value in fertility research. Further, it seems that many of the explanatory variates are highly intercorrelated, i.e., X'X is near singular. The method of ridge or damped regression (Hoerl and Kennard, 1970) might prove useful here.

Path analysis adds causal assumptions (a path diagram) to a regression analysis and thereby seeks to extract causal information from the data. The technique is discussed for fertility data by Kendall and O'Muircheartaigh (1977) (see also Little, 1980; Kendall, 1976; Hermalin and Mason, 1980). The difficulties of reaching causal conclusions with observational data, described elsewhere in this paper, are paramount, however.

INDEXES

Special indexes have long been used in demography generally and in fertility research in particular. Index numbers are meant to measure the effects of variables that cannot be observed directly, but are felt to have a definite influence on other variates that can be observed. They are used to make complex situations more understandable. They do so by quantification. Indexes may be used to measure change or to compare groups among other things (see Cox, 1950).

An obvious weakness of any idex is that, being an artificial construct, it may simply not be measuring the intended effect, but merely covarying with that effect for the data at hand.

In fertility research, indices have been constructed to avoid separate analyses for each age group, the index being a weighted average of age-specific rates of interest (the problem of which weights is that of standardization, discussed below). Because of this averaging, information will be lost, and on occasion, important differences will be obscured. In other words, the phenomenon of interest may be so complex that there is no useful way to express it through a numerically defined index.

Other less important but still significant problems arise with indexes. These include how to handle missing values and how to attach measures of uncertainty (see Kish, 1968, on the latter).

STANDARDIZATION

Users of indexes have found that difficulties resulting from reducing multi-dimensional data to single-dimensional, as is done in formulating an index, could be alleviated somewhat by standardization. This is an analytic procedure, taking several forms, designed to improve the understandability of quantities computed from data with multiple classes, especially when several quantities are to be compared or when change with time is to be examined. It is a crucial procedure in demography. The need for standardization results, in part, from the investigator's inability to impose experimental controls, and from the desire to address differences in background variables with respect to what is being compared, to correct for imbalances, and to reduce bias. (The background variables are assumed here not to be of interest themselves, but to be obscuring the relationship of interest.) A number of standardization techniques are available, some formal, some ad hoc, including direct, indirect, matching, analysis of covariance, and borrowing (see Mosteller and Tukey, 1977:Ch. 11; Maxwell and Jones, 1976; Pullum, 1978).

The standardization technique is not without limitations. Since several forms are available, the investigator must decide which to use in a particular situation. The assumptions on which each form is based are not clear, though some comments are made in the references listed above. Moreover, the technique is sometimes based on a "standard" population, the choice of which can be somewhat arbitrary. Also, the attachment of a measure of uncertainty to an end result can be complicated (Mosteller and Tukey, 1977, do suggest some procedures). Adjusting for some background variables may systematically unmatch others. Apparent changes with time may be due simply to changes in proportions in the classes involved, not to fundamental changes within classes. Mosteller and Tukey (1977:238) present an example taken from Woodward showing that standardization can reverse the order of rates in an unsettling way. Thus the technique, though clearly potent, must be handled carefully.

CONTINGENCY TABLES

The most common data structure is perhaps the table. Quite possibly, demographers initiated the field of contingency table analysis. In the simplest case, a table is rectangular, with the rows corresponding to one factor and the columns to a second; although multi-level tables have recently become common, difficulties in printing arise when more than two variables at a time are used.

(Although the statistical package S [Becker and Chambers, 1981] offers a useful method.) The factors may be qualitative or quantitative, ordinal or cardinal. The data may be observational or experimental. The table may have been formed by cross-tabulation. The tables may be intended to convey information or to suggest patterns of relationship.

The analysis of contingency table data may consist simply of forming and scanning the table, or it may involve fitting complicated and subtle models. The loglinear is a current popular and useful model (see Plackett, 1974; Bishop et al., 1975; Fienberg, 1977). Computer programs are commonly available (e.g., in BMDP and GLIM).

Difficulties in modeling arise because the data values are counts, rather than continuous. Difficulties in interpretation and even paradoxes arise if the populations involved are not homogeneous. Wagner (1982) presents two elementary examples (magazine renewal rates and income tax rates) of Simpson's paradox: when two populations are separated in parallel into a set of descriptive categories, the population with the higher overall incidence of some characteristic may exhibit a lower incidence within each table.

Finally, there are two complementary operations on tables. The first is collapsing or marginalization, which can lead to paradoxes like that described above. The second is disaggregation (also known as subclassification and poststratification).

TIME SERIES ANALYSIS AND FORECASTING

The situation of interest is often dynamic, rather than static. This leads to possibilities, difficulties, and questions. How should change be measured? How should one test for structural change? Is an apparent change due to shifts in relationships over time or in the relative sizes of subgroups of the population? How should time lags of effect be dealt with? If projections are computed, how should their uncertainties be indicated? How are the data to be handled if program goals are altered in the course of the data collection? (Indeed, if anything should alert fertility researchers to the limitations involved in applying statistical techniques to demographic data, it should be the succession of population-level projections that have proved badly in error with the passing of time. Cohen [1979] presents graphic examples of this for the official projections of births in England and Wales.) Nor can time effects be ignored totally in most static situations. Because causation has a temporal implication, causal discussions are not immune. It may be that changes in the basic variates, not their levels, are what is affecting fertility. Cross-sectional data cannot be used or interpreted as longitudinal without a strong assumption that the situation is stationary (that is, time invariant) or evolving in an understood fashion. Further, the path of a process developing in time in the presence of delays and even simple nonlinearities (due to human anticipation and behavior) can be exceedingly complex (see May, 1974; Brillinger, 1981). Modeling such behavior is fraught with difficulties, even when high-quality data are available at many time points and the situation is stationary. These last conditions rarely exist for fertility data.

Time series problems are critical to fertility research. How is one to detect, measure, and understand change? How is one to prepare forecasts? Modern time series analysis does have some techniques designed to address these questions. However, because these techniques are generally "correlational" rather than "causational," the difficulties remain; so, too, do the problems of adjustment, missing values, incorrect values, insufficient data, and the like. A further difficulty emerges from a time series view of a situation. The usual statistical procedures, especially those for estimating the level of uncertainty, assume the statistical independence of the basic data. If the data are serially correlated, the construction of tests and confidence intervals is greatly complicated, and the blind use of the procedures of the independent case is bound to deceive.

For those researchers who engage in forecasting, it is well worth repeating Lincoln Moses' admonition, "There are no facts about the future" (from Energy Information Administration, 1978:iii).

HYPOTHESIS TESTING

Two very common features of fertility research are the development of hypotheses (for parameter values or for the validity of concepts) and the testing of those hypotheses. This is, of course, the natural way in which science progresses.

Besides the difficulty of deciding which formal test procedure to use in these efforts, other problems arise. If a (null) hypothesis is rejected, what does one do next? How exactly is one to take note of the fact that typically, many tests will be made on the same set of data? (If tests are made at the 5 percent level of significance, then even if no effects are present, 5 percent of the test statistics may be expected to be significant.) The distinction between the exploratory and confirmatory modes of statistics, emphasized earlier in this report, is pertinent here.

If the data are observational, then the justification of most tests is very tenuous and the substantial advantages of having run a randomized experiment are clear once again. In practice, careful interpretation of test results (both significant and insignificant) is required in fertility situations. Variables are typically confounded, and a variable other than the obvious one may be leading to an apparent association.

THE MEASUREMENT OF ASSOCIATION

Many of the problems of fertility research come down to measuring and modeling the strengths of association among variates of interest. For example, one question is how much of an apparent decline in fertility is associated with various socioeconomic variables such as health, education, economic status, and urbanization? Sometimes the researcher measures association with a factor totally outside his or her control, such as age; sometimes the researcher uses a factor whose values can be regulated. As mentioned previously in this report, a key distinction is between causal and statistical association, with the former being more important. Addressing causal relationships in the fertility situation is extremely difficult because of multiple causes, complex connections, and the fact that some things can be regarded as both causes and consequences (for example, the decline in fertility and changes in marriage and the family). As indicated earlier, the most satisfying way to establish a causal relationship is by means of a randomized experiment. As noted above, however, in the absence of complete knowledge, in many circumstances, it may be wise to act on the assumption that an association is causal. (For some general discussion of the association/causation issue, see Mosteller and Tukey, 1977:Ch. 12; MacMahon and Pugh, 1970; U.S. Surgeon General's Advisory Committee on Smoking and Health, 1964; Wold, 1956.)

THE FORMAL ANALYSIS OF ERROR

A principal advantage of using formal statistical techniques is that a variety of procedures are available for estimating and describing the errors in results due to sampling fluctuations. Such quantification of the uncertainty of one's results is always important and sometimes essential. Statistical fluctuations in the variates analyzed have several sources: sampling error, measurement error, and model disturbance. To begin, the data should be audited prior to analysis to reduce measurement errors such as outliers. At the next stage, there is currently an extensive literature concerned with estimation of the variability of simple statistics computed for sample survey data. However, there is also increasingly extensive computation on analytically complex quantities, so that alternate procedures have had to be developed. These include replicated sampling (Kish, 1965:127; Kish and Frankel, 1974); the jackknife (Miller, 1974; Brillinger, 1976; Mosteller and Tukey, 1977); linearization (Woodruff and Causey, 1976); and the bootstrap (Efron, 1979). Once standard error estimates are available, one can go on to construct confidence intervals, compute "prob-values," and carry out hypothesis tests. In complicated situations (such as those that are highly structured or involve time series data), standard errors typically have to be estimated in an ad hoc fashion, unfortunately.

GRAPHIC METHODS

Graphic techniques have long played an important role in population research. Early procedures include Lexis diagrams and age pyramids. Modern computing and display equipment makes their use elementary. Indeed, graphs and more complicated displays lie at the heart of modern statistical analysis: they are essential for checking assumptions (e.g., the use of residual plots in regression analysis), for discovering unexpected phenomena, and for communicating the results of studies. The display may be static or dynamic. In the case of multidimensional data, satisfying nonlinear relationships, the one hope for insightful analysis seems to be dynamic displays such as PRIM-9 (Tukey et al., 1975; see also Donoho et al., 1981; Kolata, 1982). The statistical language "S" is particularly convenient for preparing elementary yet novel graphic displays (Becker and Chambers, 1981). One learns from an early age, though, that graphic methods have limitations; misleading graphs seem always to be included in works on how to "lie" with statistics.

COMPUTING

Computers are essential to modern fertility researchers who deal with large data sets and complicated models. The computer's impact, already great, can only increase. Computers provide numerous opportunities: simulations may be run, parallel analyses may be carried out easily and completely, complicated quantities may be evaluated, and sensitivity studies may be done. Large-scale (interactive) computer packages of statistical routines are now available, including SPSS, BMD, and ISP. Zlotnik (1981) has prepared a report on programs specifically intended for demographic estimation, and programs are now publicly available for the vast majority of the statistical techniques mentioned in this paper. The user should not forget, however, to inquire into the numerical accuracy of such programs as implemented on the computer being employed. Because of the fact that computers work with a finite number of digits, round-off error can occasionally make the computed results wildly incorrect. It may also be noted that with the opportunities arising from the existence of modern computing facilities, there also arise new concerns: it has never been simpler or less costly to carry out inappropriate analyses.

OTHER METHODS

The preceding sections have described some of the statistical methods that seem especially useful in fertility research. A few others of possible use might also be mentioned: factor analysis, Markov processes (to model, for example, the stages of fertility), spatial statistics, nonparametric procedures, general robust/resistant techniques (see Mallows, 1979), the Cox model of proportional hazards (see Kalbfleisch and Prentice, 1980), and the direct modeling of probabilities used in modern point process research and in risk analysis (see, for example, Brillinger, 1981).

SOME ISSUES AND DIFFICULTIES

A number of basic issues arise out of the preceding discussion: the choice of the unit of observation, the basic measurement, and the explanatory variables; the selection of a design for collection and analysis; the specification of quantities to be computed and graphed; and the assessment and analysis of variability. Several questions also arise: Is analysis to be micro or macro? What is the sensible level of sophistication of technique and analysis? What are valid and proper inferences in the light of the data and analyses? How and what are the results to be released? Basic difficulties have emerged: non-normal data; missing values (nonresponse); outliers; high intercorrelation of explanatory variates; communication of results to policy makers; measurement of change; data release; incorporation of external information; data reliability; incorrect models; systematic biases; degree of generalizability; handling of proxy variates; nonadditivity; incorrect assumptions; control of background variates; allowance for variates remaining constant throughout the course of the study, but probably causal in nature; competing risks; expense; evaluation in the absence of experimentation; misclassification; multiple sources of variation; and the question of how far extrapolations can be pushed.

DISCUSSION

A review of the recent fertility literature has shown that a large number of statistical techniques are used in that field, many of which are mentioned in this paper. The review also found that those techniques are sometimes misused and, more importantly, that unjustified inferences are made based on the data collected. Several specific limitations of the use of statistics in fertility research were noted; these include the following:

- An effect found by a correlational analysis to be highly associated with a variate of interest is often described as having a causal effect. Although this may be true, it cannot be demonstrated by correlational analyses; any causality claims must be justified and qualified appropriately.

- This review further noted that in the use of statistical techniques, it is rare for investigators to examine the assumptions that serve as a basis for the techniques being used. This is especially distressing for techniques (e.g., multiple regression) where there is currently a substantial collection of procedures available for checking those assumptions.
- Not many of the papers that construct models comment critically on the general limitations of the modeling approach; some authors may even have preferred to reason with words, rather than equations, so as not to raise doubts about the validity of their analytic procedures.
- Typically, there is no critical evaluation of the data employed in analyses, even though the authors did not collect that data themselves.
- Finally, all too seldom is enough information presented about computations so that others can assess or duplicate the results presented.

SOME RECOMMENDATIONS

A number of suggestions for the field of fertility research can be made based on the review conducted for this paper:

- More resources should be devoted to auditing data.
- There should be reporting standards for the results of statistical studies.
- Measures of sampling uncertainty should always be provided.
- Assumptions and conclusions should be critically evaluated.
- There should be more exploratory analyses.
- There should be more confirmatory studies and randomized experiments.
- Basic data should be released for independent study.
- Substantial use should be made of robust/resistant techniques.
- There should be validation of models constructed.

REFERENCES

Baker, R. J., and J. A. Nelder (1978) The GLIM System, Release 3. Oxford Numerical Algorithms Group, 256 Banbury Road, Oxford OX27DE.

Becker, R. A., and J. M. Chambers (1981) S: A Language and System for Data Analysis. Murray Hill, N.J.: Bell Laboratories.

Berquo, E. (1981) The comparative analysis of WFS data. Unpublished manuscript.

Bishop, Y. M., S. E. Fienberg, and P. Holland (1975) Discrete Multivariate Analysis. Cambridge, Mass.: MIT Press.

Box, G. E. P. (1966) Use and abuse of regression. Technometrics 8:625-629.

Brillinger, D. R. (1976) Approximate estimation of the standard errors of complex statistics based on sample surveys. New Zealand Statistician 11:35-41.

Brillinger, D. R. (1981) Some aspects of modern population mathematics. Canadian Journal of Statistics 9:173-194.

Bulatao, R. A., and R. D. Lee with P. E. Hollerbach and J. Bongaarts, eds. (1983) Determinants of Fertility in Developing Countries. New York: Academic Press.

Cohen, J. E. (1979) Ergodic theorems in demography. Bulletin of the American Mathematical Society 1:275-295.

Cox, P. R. (1950) Demography. Cambridge, Mass.: Cambridge University Press.

Dempster, A. P., N. M. Laird, and D. B. Rubin (1977) Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society, Series B 39:1-38.

Donoho, D., P. J. Huber, and H.-M. Thoma (1981) The use of kinematic displays to represent high dimensional data. Pp. 274-278 in Computer Science and Statistics: Proceedings of the 13th Symposium on the Interface. New York: Springer-Verlag.

Efron, B. (1979) Bootstrap methods: Another look at the jackknife. Annals of Statistics 7:1-26.

Energy Information Administration (1978) Projections of Energy Supply and Demand and Their Impacts. Annual Report to Congress, Vol. II 1977. U.S. Department of Energy DOE/EIA-0036/2.

Fienberg, S. E. (1977) The Analysis of Cross-Classified Categorical Data. Cambridge, Mass.: MIT Press.

Freedman, R. (1967) Applications of the behavioral sciences to family planning programs. Studies in Family Planning 23:5-9.

Freedman, R. (1979) Issues in the Comparative Analysis of World Fertility Survey Data. Paper No. 62. Honolulu: East-West Population Institute.

Gass, S. I. (1980) Validation and Assessment Issues of Energy Models. National Bureau of Standards Special Publication 569. Washington, D.C.: U.S. Government Printing Office.

Gilbert, J. P., R. J. Light, and F. Mosteller (1975) Assessing social innovations: An empirical base for policy. Pp. 39-193 in C. A. Bennett and A. A. Lumsdaine, eds., Evaluation and Experiment: Some Critical Issues in Assessing Social Programs. New York: Academic Press.

Heckman, J., and R. J. Willis (1976) Estimation of a stochastic model of reproduction: An econometric approach. In N. E. Terleckyj, ed., Household Production and Consumption Studies in Income and Wealth, Vol. 40. New York: National Bureau of Economic Research.

Hermalin, A. I., and W. M. Mason (1980) A Strategy for Comparative Analysis of WFS Data, with Illustrative Examples. United Nations Program for Comparative Analysis of World Fertility Survey Data. New York: United Nations Fund for Population Activities.

Hoerl, A. E., and Kennard, R. W. (1970) Ridge regression. Biased estimation for non-orthogonal problems. Technometrics 12:55-67.

Joiner, B. L. (1981) Lurking variables: Some examples. American Statistician 35:227-233.

Kalbfleisch, J. D., and R. L. Prentice (1980) The Statistical Analysis of Failure Time Data. New York: John Wiley.

Kendall, M. G. (1976) Some Notes on Statistical Problems Likely to Arise in the Analysis of WFS Surveys. Tech. 441. World Fertility Survey, London.

Kendall, M. G., and C. A. O'Muircheartaigh (1977) Path Analysis and Model Building. Tech. 414. World Fertility Survey, London.

Kish, L. (1965) Survey Sampling. New York: Wiley.

Kish, L. (1968) Standard errors for indexes from complex samples. Journal of the American Statistical Association 63:512-529.

Kish, L., and M. Frankel (1974) Inference from complex samples. Journal of the Royal Statistical Society, Series B 36:1-37.

Kolata, G. (1982) Computer graphics comes to statistics. Science 217:919-920.

Little, R. J. A. (1980) Linear Models for WFS Data. Tech. 1282P. World Fertility Survey, London.

MacMahon, B., and T. F. Pugh (1970) Epidemiology. Boston: Little, Brown.

Mallows, C. M. (1979) Robust methods--some examples of their use. American Statistician 33:179-184.

Mauldin, W. P., and B. Berelson (1978) Conditions of fertility decline in developing countries, 1965-75. Studies in Family Planning 5:89-148.

Maxwell, S. E., and L. V. Jones (1976) Female and male admission to graduate school: An illustrative inquiry. Journal of Educational Statistics 1:1-37.

May, R. M. (1974) Model Ecosystems. Princeton, N.J.: Princeton University Press.

McNeil, D. R. (1977) Interactive Data Analysis. New York: J. Wiley.

Miller, R. G. (1974) The jackknife--a review. Biometrika 61:1-16.

Mode, C. J. (1975) Perspectives in stochastic models of human reproduction: A review and analysis. Theoretical Population Biology 8:247-291.

Mosteller, F. (1977) Experimentation and innovations. Bulletin of the International Statistical Institute 47:559-572.

Mosteller, F., and J. W. Tukey (1977) Data Analysis and Regression. Reading, Mass.: Addison-Wesley.

National Research Council (1981) Collecting Data for the Estimation of Fertility and Mortality. Committee on Population and Demography Report No. 6. Washington, D.C.: National Academy Press.

Plackett, R. L. (1974) The Analysis of Categorical Data. London: Griffin.

Pullum, T. W. (1978) Standardization. Tech. 597. World Fertility Survey, London.

Student (1931) The Lanarkshire milk experiment. Biometrika 23:389-407.

Tukey, J. W. (1977) Exploratory Data Analysis. Reading, Mass.: Addison-Wesley.

Tukey, J. W. (1980) We need both exploratory and confirmatory. American Statistician 34:23-25.

Tukey, J. W., J. H. Friedman, and M. A. Fisherkeller (1975) PRIM-9, an interactive multidimensional data display and analysis system. Proceedings of the 4th International Congress for Stereology, Gaithersburg, Md.

U.S. Surgeon General's Advisory Committee on Smoking and Health (1964) Smoking and Health. Washington, D.C.: U.S. Government Printing Office.

Velleman, P. F., and D. C. Hoaglin (1981) Applications, Basics, and Computing of Exploratory Data Analysis. Boston: Duxbury.

Wagner, C. H. (1982) Simpson's paradox in real life. American Statistician 36:46-48.

Wold, H. (1956) Causal inference from observation data: A review of ends and means. Journal of the Royal Statistical Society, Series A 119:28-61.

WFS Central Staff (1977) Strategies for the Analysis of WFS Data. World Fertility Survey, London.

Woodruff, R. S., and B. D. Causey (1976) Computerized method for approximating the variance of a complicated estimator. Journal of the American Statistical Association 71:315-321.

Yates, F. (1981) Sampling Methods for Censuses and Surveys, 4th ed. New York: Macmillan.

Zlotnik, H. (1981) Computer Programs for Demographic Estimation: A User's Guide. Committee on Population and Demography Report No. 11. Washington, D.C.: National Academy Press.

APPENDIX: ANNOTATED BIBLIOGRAPHY

This appendix presents an annotated bibliography of the papers reviewed for this report; the papers were selected by the panel.

Bongaarts, J. (1978) A framework for analyzing the proximate determinants of fertility. Population and Development Review 4:105-132.

The stated goal of the paper is to present a model for analyzing the relationships between intermediate fertility variables and the level of fertility. A fertility level is parcelled out into proximate determining components via the relation $TFR = C_m C_c C_a C_i TF$, with the terms appearing estimated in ad hoc fashion. The relation is called a model, but it is not one in the usual statistical sense. Measures of uncertainty are not provided. A regression line is fit; however, the usual summary statistics are not provided, nor is there any indication that the assumptions of the regression were examined critically. The author states that "the model can be used in comparative fertility analysis to determine the intermediate fertility variables responsible for fertility differences among populations or among subgroups within a population." If all that is meant is that arithmetic may be carried out and some quantities computed that is one thing; however any inference that those quantities correspond to biologically relevant entities requires an assumption that many may not be prepared to make.

Heckman, J., and R. J. Willis (1976) Estimation of a stochastic model of reproduction: An econometric approach. In N. E. Terleckyj, ed., Household Production and Consumption Studies in Income and

Wealth, Vol. 40. New York: National Bureau of Economic Research.

The goal of this work is the development of an integrated theoretical and econometric model of fertility behavior within a sequential stochastic framework. The paper includes some economic theory, a dynamic stochastic model, and some empirical analysis (including significance testing and maximum likelihood estimation). There is a fair amount of discussion of the assumptions employed. One can certainly quibble about many things in the paper, but it does seem a good one.

Hermalin, A. I. (1978) Spatial Analysis of Family Planning Program Effects in Taiwan, 1966-72. Paper No. 48. Honolulu: East-West Population Institute.

"This paper uses regression analysis of areal data in Taiwan to examine the effects of the national family program on fertility . . . . Tests produce the general finding that the program did contribute to the decline in fertility . . . ." A clear assumption of this author, then, is that the units (after some correction for covariates) differ only in having the program or not. The traditional assumptions of regression analysis are apparently not examined critically for the data sets employed.

Hobcraft, J., and G. Rodriguez (1980) Methodological Issues in Life Table Analysis of Birth Histories. Paper presented at the Seminar on the Analysis of Maternity Histories, London.

This paper discusses methodological aspects of life tables for birth histories based on data from retrospective maternity histories from surveys. Some exploratory data analysis (EDA) is employed. Selection effects and censoring are discussed critically and the conclusions offered are mainly descriptive. Both the approach and conclusions seem quite reasonable. One interesting statement is the following: "A disadvantage of a model-based approach is that it raises doubts in the minds of many people about the validity of applicability of the model."

Lee, R. D. (1980) Aiming at a moving target: Period fertility and changing reproductive goals. Population Studies 34(2):205-226.

The author sets up a theory relating fertility targets and period fertility rates. A key assumption is made: Fertility of a cohort is closely related to the gap between their desired family size and the number of births to date. Implications of the assumption are derived and held up against some data. (The data studies referred to are not criticized although the analysis is a secondary one.) The contributions are mainly theoretical.

Mauldin, W. P., and B. Berelson (1978) Conditions of fertility decline in developing countries, 1965-1975. Studies in Family Planning 9(5):87-147.

"This paper is a macro-analysis of the correlates of fertility decline . . . . The analysis focuses on how much of the fertility decline is associated with . . . ." The analysis is based on secondary data. Regression is employed, apparently without checking assumptions. Individual coefficients are recorded and discussed extensively, without strong warnings about their dependence on the variates employed in the regression. (The value of a regression coefficient depends critically on which variates are included.) Some EDA is presented. The conclusions are typically carefully qualified: ". . . we have pushed them to their limits, well beyond the scientifically provable . . . ."

Mode, C. J. (1975) Perspectives in stochastic models of human reproduction: A review and analysis. Theoretical Population Biology 8:247-291.

This is a substantial and careful attempt to model fertility by means of age dependent branching processes. The model is simulated on a computer, but is not fit (in entirety) to any data sets. Statistical independence is crucial in the development of results for branching processes, and covariates are not easily included in the model. The general approach seems well worth pursuing. However, whether it is an effective one remains to be seen.

Potter, R. G., and J. Phillips (1980) Fitting and Extrapolating Contraceptive Continuation Curves by Logit Regression. Paper prepared for the IUSSP Seminar on the Use of Surveys for the Analysis of Family Planning Programs, Bogota, Colombia.

This paper considers the specific problem of modeling the probability of a contraceptor terminating use as a function of duration of usage. The model is logit and piecewise linear in duration; fitting is by maximum likelihood. Hierarchical hypotheses are examined, and a number of contraceptives are compared. The data are observational. The logit assumption is not examined in any detail, and there is apparently no examination of the extent to which groups using different contraceptives differed with respect to other variates; all contraceptors in a group are assumed to have the same probabilities.

Pullum, T. W. (1978) Standardization. World Fertility Survey Technical Bulletin No. 3/Tech. 597. The Hague: International Statistical Institute.

This paper provides some critical discussion of standardization, proxies, and path analysis in studying the impact of education on cumulative fertility in Malaysia. The data are observational. Comparisons are made by standardizing with respect to marital duration and ethnicity, for example, and interactions are examined. Although there is discussion of causation versus association, one reads: "For Chinese and Indians, an increase in education produces a decrease in fertility (holding duration constant)." It is unclear that standardization is a suitably efficient tcol for addressing the problem considered. It is recommended for use if resources are limited; this may be an improvement on doing nothing, but leads to scientific difficulties.

Retherford, R. D., and N. Ogawa (1978) Decomposition of the change in the total fertility rate in the Republic of Korea, 1966-70. Social Biology 25(2):115-127.

This is an attempt to understand a sharp decline in Korean fertility during the 1960s. The data are retrospective. The decline is partitioned out into several factors in a linear fashion, with no clear criteria of selection presented for those factors. The partitioning technique is also subject to a variety of criticisms: What are the measures of uncertainty? What about negative components? Moreover, unjustified statements are made: "Changes in education composition in turn largely explain the contribution of changes in residence composition to the decline in the TFR."

Rodriguez, G., and J. Cleland (1980) Socioeconomic Determinants of Marital Fertility in Twenty Countries: A Multivariate Analysis. Paper prepared for the World Fertility Survey Conference, London.

World Fertility Survey data are used to examine the relationship of level of childbearing to various socio-economic characteristics of couples. This work uses the technique of multiple regression extensively. There is some critical discussion of regression as a technique, yet the assumptions are not much explored. Several of the conclusions seem unjustified: ". . . the statistical analysis has shown that the expected rural-urban differences in fertility is universal" and ". . . we have shown the existence of a substantial effect of female labour force participation on marital fertility." Here, however, the authors add that "this effect remains after adjusting for all other variables in the model and therefore cannot be attributed to other socioeconomic factors."

Rosenzweig, M. R., and K. W. Wolpin (1980) Testing the quantity-quality fertility model: The use of twins as a natural experiment. Econometrica 48(1):227-240.

The proposition that the quantity and quality of children interact is analyzed. The paper begins with an extensive theoretical discussion. The theory is tested on twin data--a clever idea--by regression analysis. However, the conclusion is stated too strongly: "The results obtained are thus the first to confirm the hypothesis that exogenous increases in fertility decrease child quality . . . ."

Schultz, T. P. (1974) Birth rate changes over space and time: A study of Taiwan. Pp. 225-291 in T. W. Schultz, ed., Economics of the Family. Chicago: University of Chicago Press.

Six years of aggregate economic and demographic data for some 361 administrative regions of Taiwan are analyzed. Some economic theory is presented for a dynamic study of birth rates. A linear model with lags is set up, but its underlying assumptions are apparently not checked. Least squares is the estimation technique and there is discussion of the problems of aggregation.

Sheps, M. C., and J. A. Menken (1973) Mathematical Models of Conception and Birth. Chicago: University of Chicago Press.

The stated topic of this book is "the reaction of natality indices to variation in the physiological aspects of reproduction." It presents a number of elementary models for the birth process, but there is not much empirical work. In the time since the book was written, the statistical models described have been replaced by much broader ones, and in particular by models with covariates. If the authors' elementary models fit the data of interest, all is well, but it is not clear that this is the case.