Chapter III

THE MULTI-LEVEL APPROACH: THEORY AND CONCEPTS

*Albert I. Hermalin**

A. Overview

1. *Definition and background*

Multi-level analysis is a strategy for combining information from more than one level of observation in studying the determinants of various forms of behaviour. A basic tenet of social science is that people's behaviour not only is influenced by their individual goals and characteristics but is shaped by their social and economic environment. Multi-level analysis, by combining elements from both levels of social reality, permits greater concordance between the theoretical views and the models employed for studying behaviour.

Multi-level analysis, also called "contextual analysis", has a long tradition within certain subfields of social science, such as voting behaviour, school performance and social deviance.[1]

In recent years, there has been an upsurge of interest in contextual or multi-level analysis in population studies. This may be seen as reconciling major themes within the field and building upon valued traditions within social research. One important theme in population studies, the demographic transition, views changes in fertility and related behaviour as emanating from large-scale societal transformations, such as urbanization, rising literacy levels and movement from an agricultural to an industrial economic structure. In addition, in many countries there have been structured social interventions to influence various aspects of demographic behaviour, as through health and sanitation programmes, family planning programmes and land development schemes. Another theme may be traced to the growth in the power and popularity of survey research, which has permitted the collection of detailed information about individuals and families and which has prompted a number of influential models about the manner in which individual characteristics influence demographic behaviour.

Attention to contextual analysis recognizes that both types of factors play a role and attempts to reconcile the disjunction between over-reliance on the individual or micro approach, or on a macro approach that focuses unduly on the environmental factors. As such, multi-level analysis has found application in studies of mortality (Rosenzweig and Schultz, 1982), migration (Findley, 1982; Hugo, 1985) and fertility (Entwisle, Hermalin and Mason, 1982). It is particularly suitable where explicit programmes have been instituted to influence behaviour. In these cases, it is clear that response to the programme can depend upon specific features of the programme, characteristics of the targeted population and other aspects of the social environment. Multi-level analysis permits one to assess simultaneously the importance of various programme features and other factors influencing the behaviour in question. In subsequent sections, attention is confined to the use of multi-level analysis in studying aspects of fertility and evaluating the role of family planning programmes, although much of the discussion is applicable to other types of behaviour and programmes.

2. *Advantages of multi-level analysis*

There are several advantages to pursuing programme evaluation from a multi-level framework. They are briefly enumerated here and taken up in more detail in appropriate sections below.

The major advantage has been noted above: by combining programme elements with characteristics of the targeted population as well as other aspects of their social environment, multi-level analysis is more likely to identify the full range of factors impinging on fertility than analysis that is strictly micro or strictly macro. By the same token, the question of family planning evaluation is properly placed in the broader context of the determinants of fertility.

As a technique for programme evaluation, multi-level analysis does not impose a standard set of measures to be applied uniformly to each setting, as is the case with several other approaches. In designing a multi-level analysis for a particular country, it is possible to achieve closer collaboration among policy-makers, administrators and researchers so that programme and community factors unique to each country will be incorporated into the analysis, and results of direct policy relevance are more likely to emerge.

From the standpoint of programme administrators, the multi-level strategy can provide guidance about specific features of their programme. It allows for analysis of key programme factors while controlling for the individual and community characteristics also known to affect fertility-related behaviour. As a result, it can provide insights similar to those derived from ex-

* Director, Population Studies Center, University of Michigan, Ann Arbor. Many of the ideas presented were developed in collaboration with William Mason, who provided very helpful comments on an earlier draft, and Barbara Entwisle, as part of the Michigan Comparative Fertility Project. Their contribution is substantial but they are not responsible for the formulation presented here. The author also wishes to express appreciation to Lora Myers and Mary Scott in preparing the manuscript.

perimental design, by combining aspects of operations research with careful statistical modelling.

The structure of multi-level analysis makes it highly cost-effective. Given the existence of a survey, the macro data of interest can usually be collected at small cost. In some cases, the necessary macro data may already exist in administrative records and censuses. Even where new data are required, they need to be gathered only for those localities which served as sampling points in the micro-level survey.

3. *Basic representation of multi-level analysis*

It is useful to conceptualize multi-level analysis as a two-stage process, although, as will be clear, actual estimation does not necessarily proceed in this manner.

At the first stage, one has comparable data on individuals in a variety of social settings, which may be countries or communities within a country. At this stage, interest centres on representing the appropriate dependent variable (e.g., use or non-use of contraception) as a function of individual characteristics (i.e., the independent variables). Assume that a regression equation to estimate the effect of each independent variable has been carried out in each setting. This will yield a set of parameter estimates (intercept plus regression coefficients) for each setting.

These estimates will differ across settings. If the between-setting variability is slight, then multi-level analysis is unnecessary. Usually, however, the parameter estimates will vary considerably across settings, in which case interest centres on understanding the variability. Is it merely random fluctuations around a single set of coefficient values that hold for all contexts? Or is there also a systematic component to this variability? Answers to these questions can be obtained with a second level of analysis, in which the regression parameters (not their estimates) of each context become the dependent variables of interest and the independent variables are the relevant social-setting characteristics. Finding differences in the regression parameters across settings is equivalent to saying that the individual independent variables interact with the social-setting characteristics, since their effect varies with the level of the macro variables. And, depending upon the exact specification, it may also be saying more, such as that the micro parameters depend upon other setting characteristics.

The foregoing process can be represented by a series of equations which further illustrate the process and introduce several additional considerations. For simplicity, it is assumed that the dependent variable is affected by only one individual characteristic and that the regression parameters are affected by only one social-setting characteristic. Extension to additional micro and macro variables is straightforward:

$$Y_{ik} = b_{0k} + b_{1k}I_{ik} + e_{ik}; \quad (1)$$

$$b_{0k} = c_0 + d_0C_k + a_{0k}; \quad (2)$$

$$b_{1k} = c_1 + d_1C_k + a_{1k}; \quad (3)$$

$$Y_{ik} = c_0 + d_0C_k + (c_1 + d_1C_k)I_{ik} + (e_{ik} + a_{0k} + a_{1k}I_{ik}); \quad (4)$$

$$Y_{ik} = c_0 + d_0C_k + c_1I_{ik} + d_1C_kI_{ik} + [f_{ik}]. \quad (5)$$

Equation (1) represents the individual-level equation in which the subscript i denotes individuals, and the subscript k denotes communities or social settings; I is the individual explanatory variable, b_0 and b_1 are the regression coefficients (intercept and slope) and e represents the error term. In the two-stage representation, it is assumed that this equation is estimated in each of the K social settings. This results in K regression coefficients, which become the dependent variables in equations (2) and (3), with C_k the explanatory variable—a characteristic of the community or social setting. The regression coefficients at this stage are c and d, and the a symbols represent the error terms, since it is not likely that any set of community characteristics will fully explain variation in the individual regression coefficients, b.

If one substitutes for b_{0k} and b_{1k} in equation (1) their equivalent expressions from equations (2) and (3), one obtains equation (4), which expresses the individual-level dependent variable as a function of the individual and community explanatory variables and a complex error term, represented by the term in brackets. Several features of equation (4) are worth noting. The effect of I_{ik} represented by $(c_1 + d_1C_k)$ will vary from community to community, depending upon the value of C_k. This arises from the fact that b_{ik} in equation (3) depends upon C_k and demonstrates the nature of the interaction between the individual and community variables. The ability to capture this interaction is one of the strengths of multi-level analysis, because it is often hypothesized that reactions to a programme will differ among individuals with different characteristics. For example, where a programme is stronger, differentials in contraceptive use among educational strata may be less, on the assumption that the programme will facilitate use among the less educated to a greater extent than among the more educated, who may already have other sources for their contraceptive needs. This pattern of diminished differentials in contraceptive use by education has been observed over time in a number of countries.

Also deserving mention is the structure of the error term in equation (4), particularly the fact that the magnitude of the error is dependent in part upon the value of the individual variable. As a result, the error term is not likely to be independent of the value of the individual characteristic, I_{ik}, violating one of the assumptions of ordinary least-squares estimation.

Equation (5) rewrites equation (4) in an alternative form which shows that the two-stage process can be presented as a single equation containing individual and community characteristics, and interactions between them, as well as an error term, represented by the bracketed item. The nature of the error term, of course, though shown by a single term here, remains the same as in equation (4). Thus, although it is possible to represent the multi-level approach as a single equation, the underlying dynamics of the process must be kept in mind to provide adequate specification and estimation. Methods for estimating structures characterized by equations (1)-(4) have been developed by Mason, Wong and Entwisle (1983) and by Wong and Mason (1985) for the case of Gaussian (normal) error in equation (1) and

the case of a dichotomous dependent variable in equation (1).

4. *Basic steps of multi-level analysis*

The foregoing representation implicitly spells out the basic steps in carrying out a multi-level analysis, and this section defines them more explicitly. Additional detail for many of these steps is then reviewed below.

A multi-level analysis begins with the identification of the levels of analysis and development of the micro and macro models. Comparable data must be available on individuals as well as the macro units selected. For a cross-national study, the data for individuals will generally come from surveys; and the potential for comparative analysis has been greatly enhanced by two international survey programmes—the World Fertility Survey (WFS) and the Contraceptive Prevalence Survey (CPS)—which provide reasonably comparable data for a large set of countries. The national data on the social, economic and demographic characteristics of the country generally will be taken from its statistical system as reported in official publications or various international compendiums. Special studies that provide national data of particular relevance, such as family planning programme effort scores, may also be employed (Mauldin and Berelson, 1978; Lapham and Mauldin, 1984). On occasion, it may also be desirable to aggregate the individual-level data to develop macro-level measures appropriate to the model.

The multi-level strategy may also be applied within a single country, with communities serving as the macro level. In this case, a single survey of sufficient size with a probability sampling plan will generally provide comparable data on the individuals located in the communities sampled. The macro data concerning the communities may come from a variety of sources: official published statistics; administrative records of the family planning programme and other relevant activities; community data collected in the course of the survey, such as through the WFS community-level modules; and special data collected from the communities as part of the multi-level analysis.

Although theoretically one may carry out a multi-level analysis at three or more levels of aggregation (e.g., individual, community and country), the problems in so doing are formidable and analysis currently tends to be restricted to two levels.

After selecting the individual-level behaviour to be explained (i.e., the dependent variable), the key step is specification of a model of the underlying dynamics. In formulating these models, it is useful to proceed in the two-step fashion outlined above, first developing the micro-level explanatory variables and their hypothesized effects; and then selecting the macro characteristics hypothesized to affect the regression coefficients (intercept and slopes) of the micro equation. An adequate micro model is crucial because omission of a strategic individual-level variable may lead to omitted variable bias.

Attention should also be given to the causal ordering of the micro and macro variables, although at present it is difficult to incorporate macro indirect effects on the micro variables as well as the interaction effects already inherent in the multi-level approach (shown in equations (4) and (5)).

After development of the appropriate models and operationalization of the specified variables, the next step is estimation. The preferred approach is represented by equation (4), which recognizes the complex error term and makes use of an appropriate procedure, such as the restricted maximum-likelihood/Bayes procedure. If necessary programs for a full multi-level analysis are not available, a second-best alternative would be to estimate equation (4) or (5) with ordinary least squares or a logistic regression, depending upon the nature of the dependent variable. This alternative, which implicitly assumes that the regression coefficients given in equations (2) and (3) are determined by the macro variables without error, is often described as a "fixed-effects model".

Although the two-stage procedure can be valuable for exploratory purposes, rarely will one use it for estimation purposes, first carrying out the micro estimation indicated by equation (1) and then estimating the macro equations (2) and (3). This situation arises from both pragmatic and theoretical reasons. Where the multi-level analysis is applied within a single country, with communities serving as the macro stage, there will rarely be more than from 30 to 50 interviews per community. As a result, the estimates of the regression coefficients within community, equation (1), will be subject to large sampling errors; in addition, there will be little inherent interest in these coefficients rather than those indicated by equation (5). From a theoretical standpoint, since the regression coefficients are stochastic variables, their estimation in equations (2) and (3) will also require special techniques.

B. Development models

1. *Example of models employed*

To make more concrete the various issues involved in model construction, this section begins by sketching the essential elements of a multi-level analysis carried out on data from Thailand for the purpose of ascertaining the role of accessibility to family planning programme service points in contraceptive use (Hermalin and Chayovan, 1984). The major elements of the model may be set forth as follows:

(a) *Individual level*:
  Dependent variable: current use of contraception;
  Independent variables: desire for more children; wife's education;

(b) *Community level*:
  Independent variables:
  (i) *Programme characteristics*: accessibility (time and distance measures to various types of outlets);
  (ii) *Other characteristics*: region; distance to district centre; distance to secondary school; presence of electricity;
  Interactions: region × accessibility; accessibility × education; accessibility × desire for more children.

The analysis was carried out for the women interviewed in the 51 rural villages of the 1979 National Survey of Fertility, Mortality, and Family Planning in Thailand. (The paper cited also reports on a similar model employed with the rural villages covered in a 1972 survey.) This survey provided the data for the dependent variable and the micro-level independent variables. The community-level data were obtained from a special investigation designed to collect detailed and comprehensive information on accessibility to various sources of contraceptive services as well as on socio-economic and cultural conditions of the sample villages.[2]

The model displayed was estimated separately for two age groups of women, those under 30 and those 30 and older. (Additional age categories might be desirable when sample size permits.) For each age group, an equation similar to equation (5) was estimated using ordinary least squares. This method of estimation might be questioned on two grounds: it does not take into account the complex error term associated with this specification; and it ignores the dichotomous nature of the dependent variable. At the time of this study, the program for the special estimation technique called for in multi-level analysis was not yet available for within-country analysis, and subsequent testing revealed that the special nature of the sampling plan (with many villages having relatively few interviews) made application of the restricted maximum-likehood approach problematical. Since the dependent variable is binary, a logit regression is more appropriate than ordinary least squares. In this case, however, where contraceptive use is around 50 per cent, little difference in the magnitude of the regression coefficients is expected between these two modes of estimation, although the standard errors of the coefficients may well differ. (A comparison of ordinary least squares and logit regression for the 1972 survey, where contraceptive use was only 19 per cent, produced practically identical estimates of the coefficients.) A related multi-level analysis of Thai CPS data, which employs a logit regression, shows similar findings, given the differences in data and specifications (Entwisle and others, 1984).

The emphasis here is on the general structure of a multi-level analysis and not on estimation *per se*. The annex to this chapter presents the results of another analysis of contraceptive use. In that annex, two types of multi-level modelling are carried out—one with, and the other without, errors in the macro equations. In both cases, a logistic response formulation is used. In general, the appropriate form of estimation will be a function of the nature of the dependent variable and the structure of the posited equations. If the dependent variable were normally distributed, one would still have to decide whether to estimate a fixed-effects model or a stochastic-parameter model. The factors underlying this choice are discussed further in the illustrative example presented in the annex.

Several features of the analysis are worth noting. The models employed only two individual-level independent variables although many characteristics were collected in the surveys. The theoretical justification for the variables selected is discussed below in the section on the choice of micro models. Likewise, there are often many community characteristics that can be incorporated; strategies for selecting among them are taken up in the subsequent discussion of the macro model. A distinguishing feature of the analysis was the use of multiple indicators of accessibility. Most studies develop one or two simple measures based on the availability of facilities within a village, or the distance (or travel time) to a specific source, or the nearest distance to any outlet (see review by Tsui, 1985). In this study, 11 different measures were constructed, ranging from time or distance to the nearest health station to weighted indices of times and/or distances to different types of outlets (*tamban* health stations, district health centres, district and provincial hospitals).[3]

In the multi-level analyses, six different measures of accessibility were employed at some points to observe the extent to which different operational measures of accessibility affected the conclusions about the role of the programme. Many other features of a family planning programme can also affect contraceptive use and can be tested in a multi-level analysis, and a number of these features are described below.

As indicated earlier, a particularly valuable feature of multi-level analysis is its focus on the interactions between the macro and micro variables. The model under discussion hypothesized two interactions between community characteristics and individual variables: one between the level of accessibility and desire for more children; and the other between accessibility and education. In the first case, one assumption is that differentials in contraceptive use between women who want more children and those who wish to stop childbearing will differ according to the level of accessibility. If higher accessibility leads women who want no more children to adopt contraception more quickly than it encourages women who want more to space their births, the differentials between these two categories should increase with accessibility. On the other hand, it can be argued that those who want no more children are highly motivated and likely to be little influenced in their contraceptive behaviour by differing degrees of accessibility, while those who still want more children are more likely to adopt contraception for spacing purposes in an environment of high accessibility. This would result in diminishing differentials between potential spacers and limiters as accessibility increases.

The other interaction effect, between accessibility and education, assumes that in villages with higher accessibility there will be smaller differences in contraceptive use among educational categories, because the less educated will gain in relation to the more educated in terms of knowledge of methods and sources and the ability to obtain supplies and services; that is, the less educated will see their costs of fertility regulation reduced faster than the more educated. Changes in differentials of contraceptive use by education have occurred in Thailand (Knodel and others, 1982), although this change over time is no assurance of a similar contraction in the cross-section.

It is possible that other community characteristics also interact with the individual variables in the same manner as accessibility. For example, educational differentials in contraceptive use may differ according to the level of ur-

banization or modernization, but these possibilities were not tested in this analysis.

Interactions among the macro characteristics must also be considered. The effect of accessibility, for example, may differ in areas that differ according to ethnicity or language, or by degree of modernization. In this analysis, this possibility was explored by testing for interaction between region of the country, which captures many of the cultural and social differences across areas, and accessibility. It should be noted that in terms of equations (1)-(4), this is equivalent to allowing an interaction between macro variables to affect the intercept of the micro equation (represented by equation (2)) but not including this interaction term in equation (3) as an effect on the slope of the micro equation. If interaction between macro variables is introduced into equation (3), this would result in a three-way interaction term in equation (4), incorporating the two macro variables and the relevant individual variable. Underlying theory may point to a model of this type.[4]

The results of the analysis were somewhat different for the two age groups. For those aged from 15 to 30, the regression coefficients indicated no significant differential in contraceptive use by educational category and no interaction between accessibility and education. There were sharp differentials in use by desire for more children and, for many of the measures of accessibility tested, a significant interaction between accessibility and this motivation variable. The nature of the interaction is to increase the differential between those who want to stop childbearing and those who want additional children in areas of high accessibility. The main effect of accessibility (i.e., its coefficient alone) was not significant, indicating that higher accessibility mainly promoted higher contraceptive use among those who wanted no more children but did little to promote spacing. Of the other community-level variables, there were significant differences in contraceptive use by region and by the level of electrification; there were no significant differences among villages according to distance to district centre or distance to a secondary school. Tests of the interaction between accessibility and region revealed that this was a significant factor only in the southern part of the country.

Among the older women, those aged from 30 to 45, the major difference in pattern from the younger was the absence of any significant interaction between accessibility and desire for more children. When a model without this interaction was estimated, it revealed significant main effects of accessibility on contraceptive use for many, but not all, of the accessibility measures developed, indicating that greater accessibility promoted contraceptive use. This effect, however, was not different among those who did and those who did not want more children. Stated otherwise, the differential in use among those who desired more children and those who did not was not significantly altered in areas with high accessibility compared with areas of lower accessibility.

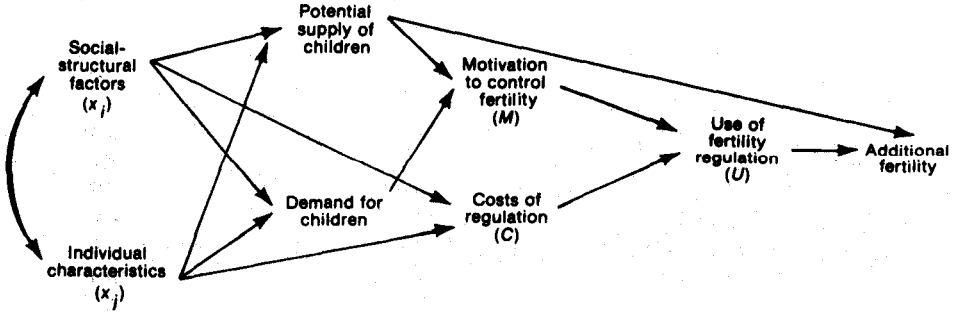
These results illustrate how multi-level analysis elucidates the factors affecting contraceptive use. A strictly micro analysis would not reveal the effect of the programme and of the other community characteristics; a strictly macro model would not capture the important interaction between accessibility and the desire for more children (among younger women). Knowledge of the magnitude of the programme effect and of its interactions with other variables in the model provides the policy-maker and analyst with important information for assessing programme effectiveness and may provide clues for strengthening programme operations.

2. *Development of the micro model*

As stated above, it is important to have a reasonably complete model of the individual factors to avoid omitted variable bias. The view taken here is that existing micro theory concerning contraceptive behaviour is more fully developed than macro theory. The selection of the individual variables in the example just presented was guided by the formal model developed by the National Academy of Sciences Panel on the Determinants of Fertility in Developing Countries (Bulatao and Lee, 1983), as elaborated for the study of fertility regulation (Hermalin, 1983). This model is sketched in figure IV.

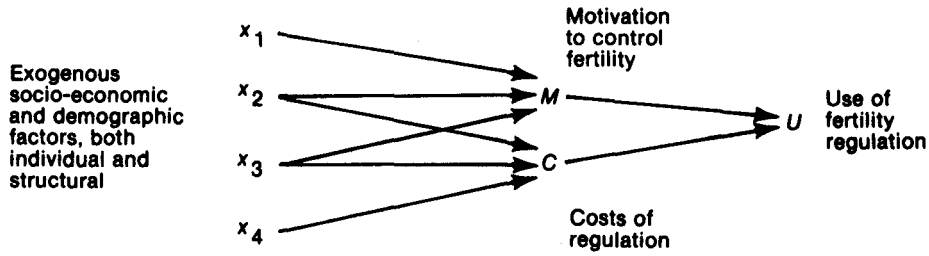
The model assumes that each couple has a demand for some number of surviving children that serves as a basis for assessing the sufficiency of their supply (or likely number of surviving children) at any given point. Once supply reaches or exceeds the desired number, the couple is motivated to some degree to control their fertility. Whether they do so is affected by the means of fertility regulation available and their costs. Although it is possi-

Figure IV. Model of factors determining additional fertility



Source: National Academy of Sciences model taken from Hermalin (1983), p. 3.

Figure V. Reduced model of factors affecting fertility regulation



Source: Hermalin (1983), p. 5.

ble in principle to estimate the entire structural model, it is common to focus on selected parts. The "semi-reduced" form used in the Thai analysis is illustrated in figure V. From this standpoint, the likelihood of using contraception (or another means of fertility regulation) is determined by motivation and costs. Motivation to limit childbearing is a function of the demand for and supply of children (see figure IV) and can be measured by a woman's expressed desire to have more children or by comparing the total number of children reported as ideal or desired with the actual number living.

The concept of costs is multidimensional—including economic factors (money and time to obtain knowledge and services), social attitudes (the possibility of violating current norms and facing sanctions) and health and psychic elements (the fear of trying something new which may be to some degree risky or unpleasant). These categories make clear that costs are a function of both individual and social structural characteristics. They will be determined by a couple's economic position, social standing and other traits; at the same time, a community may raise or lower costs by its prevailing social attitudes, which encourage or discourage fertility, and through the structure of the family planning programme, which determines the prices, accessibility and quality of the services provided. In the Thai example, the wife's level of education was used to capture many of the individual dimensions associated with costs—degree of knowledge, willingness to innovate, ability to meet direct and indirect monetary expenses etc. Other characteristics of the couple available in most surveys, such as husband's education and occupation, wife's labour force status and number of living children, were not included in the micro model either because they overlap considerably with wife's education or because the direction of causality is ambiguous (e.g., between number of children and contraceptive use; see Hermalin, 1983). Although the Thai example employs only two individual variables, the theoretical framework does not point to any omitted characteristics that are likely to bias either the micro or macro effects.

It should be noted that since the model was derived to explain the adoption of fertility regulation to limit births, its application to the use of contraception for spacing is more problematical. This is one reason that the analysis was carried out separately for two age groups, because contraceptive users in the older age group are more likely to be using for limiting than those in the younger group. Stratifying by age also helps control for life-cycle factors (duration of marriage, numbers of children) which affect the likelihood of contracepting.

The theoretical approach followed in the Thai example is, of course, not the only one that can be adopted. The main purpose of presenting this example was to demonstrate the importance of developing a conceptual framework to guide the selection of a parsimonious number of individual-level variables and to avoid circularities in causation that render the results meaningless. Even within the National Academy of Sciences framework shown in figure IV, models other than that presented can be developed. For example, it may be desirable to use as independent variables the exogenous individual characteristics that determine motivation (the X_i variables shown in figure V). The framework can prove helpful here as well since figure IV indicates that any exogenous variable selected must be justified as a determinant either of demand or of supply of children, or as a determinant of cost.

3. *Development of the macro model and choice of interactions*

This section takes up the problems associated with selecting the community and programme characteristics to include in a multi-level analysis and the interactions to be incorporated. The social and economic characteristics are taken up first, followed by the choice of programme characteristics.

If the reduced model presented in figure V is followed, one is interested in aspects of the community that may have a bearing on the costs of adopting contraception, where costs represent the complex of dimensions described above. Current theory provides less guidance for the choice of macro characteristics than for the micro variables. Many aspects of community structure could have an effect on costs and it is difficult to specify a parsimonious list. One broad category concerns the extent to which the community has contacts with the "outside world" through the mass media. This would include not only the prevalence of movies, television, newspapers and magazines but the content of the messages received. Programmes and materials that foster aspirations for new life-styles should not only increase the motivation to adopt fertility regulation but reduce the subjective costs of experimenting with con-

traception (Freedman, 1979). The various measures of modernization commonly obtained, such as educational and health facilities and the availability of electricity and modern goods, operate in a similar fashion.

Other measures of relevance would attempt to discern structural factors that affect the likelihood of a critical mass of innovators and the speed of diffusion. Depending upon the society, it is suggested (Hermalin, 1985) that possible indicators of this process might include:

(*a*) Size of the middle class and measures of income distribution or measures of the concentration and distribution of land holdings;

(*b*) Distribution of the labour force, with particular attention in some settings to the degree to which agricultural families engage in off-farm employment;

(*c*) Ethnic distribution of the community to identify the possible existence of enclaves that hinder community-wide diffusion of ideas;

(*d*) Existence and utilization of farmer associations and other organizations that might serve to readily diffuse new information;

(*e*) Type of political structure and the nature and popularity of the leadership;

(*f*) Ecology of the community, including its distance from urban centres which might serve as a source of diffusion of new ideas.

Thus, a key issue in multi-level analysis is data reduction among the potentially large number of macro variables. There is no simple solution to this problem. Strategies sometimes employed, such as arbitrary index construction, stepwise regression or factor analysis, are atheoretical and do little to illuminate the basic processes. A preferred approach would be first to inform the choice of macro characteristics through study of a country's culture and socio-economic structure, and discussions with knowledgeable observers about local and regional variations. A second step might include a separate analysis of all the macro data to identify the degree of correlation across items and to detect distinct dimensions. This might be followed by the utilization of confirmatory factor analysis (Long, 1983) in which one tests for the coherence of indicators around underlying constructs, specified in advance.

In using multi-level analysis for evaluation of the impact of a family planning programme, particular attention should be paid to those aspects of the programme which may contribute to variations in contraceptive use. The Thai example described above permitted assessment of indices of travel time and distance to various types of outlets. But many other aspects of a programme may influence use. One suggestion is that the concept of accessibility be conceptualized as the objective supply environment whose components include not only distance and travel time but the ease, convenience and cost of travel, as well as the range, cost and quality of services offered (Hermalin and Entwisle, 1985). Each of these characteristics may directly influence the cost of adopting contraception. In addition, other aspects of a programme may affect the demand for contraceptive services, particularly those associated with the informational and educational efforts of the programme.

Which of these programme features should be incorporated into the analysis depends upon the specific country situation. Some of these dimensions may not vary across communities and thus are not relevant to the analysis. For example, the price of contraceptive services or the hours of clinic operation may be uniform across communities. The goal is to include salient programme features that vary from community to community. Discussions with programme administrators and knowledgeable observers can assist in identifying the appropriate programme variables.

Attention must also be paid to availability of contraception from non-programme sources because this factor can influence both overall contraceptive use and the level of use from the programme. In the former case, variation across communities in availability of services from non-programme outlets can be an important effect on the likelihood of contraceptive use. But even if the dependent variable is contraceptive use from programme sources, the level of non-programme availability is salient, since non-programme outlets may reinforce and enhance the utilization of programme outlets or may compete with them. One possible way to capture this process is to include an interaction term between programme and non-programme accessibility, as well as each level of accessibility separately.

Other possible interactions among the macro variables were noted above in the Thai example. In particular, it is important to be alert to aspects of the community that might interact with programme characteristics. These aspects may be cultural (ethnicity, language), ecological (terrain, transport facilities, marketing patterns) or related to levels of development.

Examples of interactions between macro and micro variables also were given above. Identification of plausible cross-level interactions is aided by considering the two-stage process represented by equations (1)-(3). Inclusion of a macro-micro interaction term is equivalent to asserting that the macro variable has an effect on the micro regression coefficient, so underlying theory should provide a rationale for changing differentials in the micro variable according to levels of the macro variable. In setting forth the interactions, it is desirable to follow a hierarchical design. This means that macro variables included as interaction terms (with micro variables or other macros) also appear as main effects in equation (5). For cross-level interactions, this is equivalent to saying that any macro variable hypothesized to affect a micro regression slope in equation (3) should also be included as an effect on the micro equation intercept in equation (2). It is reasonable to assume that any variable that alters the way another variable influences the dependent variable also has a direct effect on the dependent variable.

In developing the macro model, it is also important to take into account the dynamic aspects of the process under study. For example, prior levels of accessibility have implications for current patterns of use. Two villages currently equal in their travel time to outlets may well have displayed different levels of accessibility in the past, with consequences for contraceptive use that continue through time. A similar problem arises in modelling the effect of elements of modernization on fertility

behaviour. In so far as the theory points to lags in the response to broad social structural changes, this should be reflected in data collection and modelling. A related need for more effective modelling is to collect information on the processes that lead to differentiation among communities on key characteristics. To illustrate from the Thai example, one may ask to what extent family planning facilities are located in accord with other services and amenities in the area, or the perceived receptivity of the community; and what factors lead some communities to have a higher percentage of homes electrified while others have no electricity. Attention to questions of this type will help ensure a properly specified causal model and point to the correct dating of the key macro variables.

Up to this point there has been an implicit assumption that the dependent variable of interest is current use of contraception or current use of contraception from programme sources. In evaluating a programme, however, the focus may be on the use of specific methods, because methods vary in terms of the facilities and personnel needed to distribute them. This shift in the dependent variable complicates matters because the availability of competing methods, as well as of the method in question, must be taken into account. In addition, the dependent variable in such models must accommodate multiple options (e.g., use of a particular method, use of some other method, no use). One author (Jones, 1984) uses multinomial logistic models for these purposes.

C. Sources of data

Brief reference has already been made to the sources of micro and macro data for multi-level analysis. This section takes up a few special issues that arise in carrying out multi-level analysis within a country for the purpose of programme evaluation.

1. *Individual-level data*

Surveys are likely to be the major source of individual data required for programme evaluation analyses, although public-use tapes of census data or service statistics might provide the basis for specialized investigations of such topics as factors affecting discontinuation of contraceptive use. Surveys already conducted, either as part of an international programme like WFS or CPS or from local auspices, may be used. They must, however, contain the requisite information, be based on a representative sampling plan and contain a sufficient number of interviews in a reasonable number of communities. The words "sufficient" and "reasonable" cannot be defined precisely but a range of 30-50 interviews in 50 or more communities serves roughly as a lower bound. Attention should also be given to the definition of the sampling cluster (Casterline, 1984). In some cases, these are actual political or social units, which also have administrative identity so that census and other data are likely to be available. In other cases, the units chosen as sampling clusters are not independent arenas of social and economic interaction and one may want to aggregate to a larger unit. Aggregation may also be necessary if the number of interviews per sampling cluster is too small, although this will reduce the number of macro units available for the multi-level analysis.

The foregoing considerations should be taken into account when designing a survey that will also serve as a basis for multi-level analyses. The number of interviews per community, the number of communities and the definition of the sampling cluster all require attention along the dimensions just addressed. For example, given the same average number of interviews per community, a sampling design that produces less variance in sample size per community is to be preferred to one that produces greater variance, since the latter will yield more communities with a small number of interviews.

The content of the questionnaire should produce the information needed for each individual woman or couple. In general, this will include items on current contraceptive use, the method used and the source of the method, particularly whether it is a programme or non-programme source; and information on status, attitudes and behaviour likely to serve as explanatory variables in the micro model. Additional data collected from each couple or household may prove helpful in structuring the macro model or in investigating the relation of the macro and the micro characteristics. This information would include data on the history of contraceptive use and the presence of various amenities and durable goods in the household. Items of this type might be incorporated directly into the micro model or might serve as a more accurate guide to community-level characteristics than similar data obtained from other sources. For example, the item on the history of contraceptive use would indicate the recency of contraceptive use and assist in identifying the appropriate lags for the macro programme variables; a household response on the presence of electricity is likely to provide a more accurate estimate of the proportion of homes electrified in a community than the report of a village headman. The same strategy can be employed to measure certain dimensions of programme accessibility. As an illustration, consider the question of programme quality, involving such aspects as waiting time at clinics, privacy, courtesy and perceived professionalism of the staff. These facets of programme operations are difficult to obtain from programme personnel and administrators because the responses may be self-serving, do not exist in programme records and are likely to be unreliable when based on the impressions of one or two informants or on brief observation. In situations where prevalence is reasonably high, responses from individual women on aspects of programme quality can be aggregated to form a community-level measure of programme quality. The guiding idea is that, although each respondent may be fallible to a degree, collectively their responses probably serve to rate communities accurately by the quality of services provided.

The converse situation should be noted. In CPS and many WFS efforts, it is common to ask respondents about travel time or distance to various sources of contraception; and these data are sometimes employed as individual characteristics in accounting for contraceptive use. This practice tends to produce perceptions based on sources known, preferred and utilized which are partially a function of a couple's pattern of con-

traceptive use (Hermalin and Entwisle, 1985). For this reason, it is circular to use these data as determinants of use. The authors recommend that data on travel time and distance as well as several other facets of programme operation be collected at the macro level.

2. *Macro-level data*

With respect to possible sources of macro data, it is useful to distinguish between situations in which the analyst must use existing data and cases in which new data can be collected. In the former situation, there still may be a wide array of data sources: census and administrative information; data from the family planning programme; and special investigations that may have been conducted as part of a survey, such as the community-level module employed by WFS on a number of occasions. The analyst must, of course, evaluate the suitability of the data for the proposed analyses as well as their reliability. In some cases, the desired macro data may not be available or be too inaccurate. Special attention should be paid to the methods of data collection.

As an example, most observers agree that the community data obtained through the WFS modules were not collected with the same care and thoroughness which characterized the individual interviews, and the field procedures used are not well documented. The general practice was to collect the community data while the survey team was interviewing in the community, usually by having the team supervisor interview the village headman at the end of the team's stay in the village. This practice limited the supervisor's ability to be knowledgeable about the community or to become acquainted with the village leaders. The few tests of the reliability and validity of village data carried out suggest potentially serious problems of reliability on key variables (Casterline, 1984). In addition, Chayovan and Knodel (1985), in pre-testing the feasibility of group interviews, note sizeable variation in the responses from different informants on the existence of or proximity to various services. In some countries, on the other hand, community data were collected by special investigators after the completion of a CPS or WFS, and this added attention probably improved reliability.

Many of the distinctive strengths of multi-level analysis come to the fore when the analyst can influence the content of the macro data and the modes of data collection. These advantages include:

(*a*) The ability to refine hypotheses about salient macro characteristics by interviewing knowledgeable observers and programme administrators, conducting infomal individual and group interviews etc. The goal here is to assemble a succinct list of the community and programme characteristics that are thought to influence the behaviour in question (e.g., contraceptive use);

(*b*) The ability to develop appropriate strategies for gaining accurate measures of the hypothesized macro characteristics. A wide battery of data collection techniques may be employed including: actual measurement in the field of distances from community or neighbourhood centres to specified types of outlets; interviews with groups of knowledgeable informants (Chayovan and Knodel, 1985) or with particularly influential leaders; utilization of programme records on staffing and other features of outlets; and aggregation of individual survey responses, as described above.

Most of these advantages, it should be noted, hold regardless of whether the individual-level survey has already been undertaken. In cases where the micro survey has been completed, however, there will be no opportunity to influence the content of the questionnaire so that some items of relevance will be lost either for the micro model or, through aggregation, for the macro model. A good example of the collection of community-level data well after the micro surveys is illustrated by two researchers (Chayovan and Knodel, 1985), who obtained data for the years 1969, 1972 and 1979 in an investigation carried out in 1983.

These approaches to macro-data collection also illustrate the inherent cost-effectiveness of multi-level analysis. Given the existence or plans for a survey, the macro data need only be collected for the communities covered by the individual survey, so that even extensive data-collection activities can be carried out at fairly modest costs. It is desirable to ascertain that the macro units selected in the sampling plan are representatively chosen.

D. Additional perspectives on multi-level analysis

This paper has reviewed the potential and limitations of multi-level analysis with special reference to the study of contraceptive behaviour. Conceptually, this strategy avoids the disjunction between strictly micro and strictly macro analyses, permitting the use of both types of variables in a manner compatible with theoretical frameworks. In addition, it reduces the over-reliance upon surveys for data on group characteristics that are more easily measured at the macro level, such as travel time or distance from a village to a specified outlet.

Several other benefits are associated with using the multi-level strategy for the study of contraceptive behaviour. It encourages close co-operation between researchers and policy-makers to develop appropriate models and to identify the salient programme factors which need to be taken into account. This should also ensure greater utilization of research results. Contraceptive behaviour is investigated in the broader context of the determinants of fertility by combining community-level programme data with social structural factors. At the same time, the multi-level strategy can produce direct guidance to administrators and policy-makers about specific features of their programmes that are more or less successful.

Despite the advantages, several limitations must also be mentioned. Multi-level analysis is not a strategy that can be applied mechanically. There are no set formulae and no pre-specified measures. The development of appropriate micro and macro models and careful attention to their interrelationships are necessary first steps. Results are likely to vary with the models employed. Problems of estimation and data reduction must also be recognized. The present data-collection abilities often produce a wide array of community-level data, but the

social structure theory used here is often too general to provide guidance for the selection of specific variables. Improved theory will also be needed to capture the increasing use of contraception for spacing, because existing theory focuses on its use for the purpose of limiting childbearing.

On balance, multi-level analysis is a useful tool in the evaluation of programmes because it can measure programme effects while controlling for many other factors, and it can assist in identifying specific programme elements that contribute to variation in contraceptive prevalence.

NOTES

[1] Eeden and Hüttner (1982) provide a general treatment and an extensive literature review, while Boyd and Iversen (1979) treat a number of the methodological issues.

[2] This investigation is described in Chayovan and Knodel (1985).

[3] The construction of these measures is described in some detail in Chayovan, Hermalin and Knodel, 1984.

[4] An example from a cross-country multi-level analysis is given in Entwisle, Mason and Hermalin (1984).

REFERENCES

Boyd, Lawrence H. and Gudmund R. Iversen (1979). *Contextual Analysis: Concepts and Statistical Techniques*. Belmont, California: Wadsworth Publishing Company.

Bulatao, Rodolfo A. and Ronald D. Lee, eds. (1983). With Paula E. Hollerbach and John Bongaarts. *Determinants of Fertility in Developing Countries*. 2 v. Report of the National Research Council Committee on Population and Demography, Panel on Fertility Determinants. New York: Academic Press.

Casterline, John (1984). The collection and analysis of community data: the WFS experience. Paper submitted to the World Fertility Survey Symposium, London, 1984.

Chayovan, Napaporn, Albert I. Hermalin and John Knodel (1984). Measuring accessibility to family planning services in rural Thailand. *Studies in Family Planning* 15(5):201-211.

_____ and John Knodel (1985). Improving the collection of village-level data: an experience from Thailand. In John Casterline, ed., *The Collection and Analysis of Community Data*. Proceedings of the World Fertility Survey Seminar on the Collection and Analysis of Data on Community and Institutional Factors, London, 20-23 June 1983.

Eeden, P. and H. J. M. Hüttner (1982). Multi-level research. *Current Sociology* (London) 3(3).

Entwisle, Barbara, Albert I. Hermalin and William M. Mason (1982). *Socioeconomic Determinants of Fertility Behavior in Developing Nations: Theory and Initial Results*. National Research Council Committee on Population and Demography Report No. 17. Washington, D.C.: National Academy Press.

_____, William M. Mason and Albert I. Hermalin (1984). *The Multilevel Dependence of Contraceptive Use on Socioeconomic Development and Family Planning Program Strength*. Population Studies Center Research Report No. 84-60. Ann Arbor: University of Michigan.

_____ and others (1984). A multilevel model of family planning availability and contraceptive use in rural Thailand. *Demography* 2(4):559-574.

Findley, Sally E. (1982). Methods for linking community-level variables with migration survey data. In *National Migration Surveys; Survey Manual X: Guidelines for Analysis* (ST/ESCAP/203). Bangkok: Economic and Social Commission for Asia and the Pacific, 276-311.

Freedman, Ronald (1979). Theories of fertility decline: a reappraisal. *Social Forces* 58:1-17.

Hermalin, Albert I. (1983). Fertility regulation and its costs: a critical essay. In Rodolfo A. Bulatao and Ronald D. Lee, eds., *Determinants of Fertility in Developing Countries*; vol. 2, *Fertility Regulation and Institutional Influences*. New York: Academic Press, 1-53.

_____ (1985). Integrating individual and community data in the study of contraceptive behavior. In John Casterline, ed., *The Collection and Analysis of Community Data*. Proceedings of the World Fertility Survey Seminar on the Collection and Analysis of Data on Community and Institutional Factors, London, 20-23 June 1983.

_____ and Napaporn Chayovan (1984). *The Effects of Individual, Village and Program Characteristics on Contraceptive Use in Rural Thailand*. Population Studies Center Research Report No. 84-67. Ann Arbor: University of Michigan.

_____ and Barbara Entwisle (1985). Future directions in the analysis of contraceptive availability. In *Proceedings of the International Population Conference, Florence, 5-12 June 1985*. Liège: International Union for the Scientific Study of Population, vol. 3, 445-456.

Hugo, G. (1985). The nature of community effects in the study of migration and their empirical investigation. In John Casterline, ed., *The Collection and Analysis of Community Data*. Proceedings of the World Fertility Survey Seminar on the Collection and Analysis of Data on Community and Institutional Factors, London, 20-23 June 1983.

Jones, Elise F. (1984). *The Availability of Contraceptive Services*. World Fertility Survey Comparative Studies, No. 37. Voorburg, The Netherlands: International Statistical Institute.

Knodel, John and others (1982). *Fertility in Thailand: Trends, Differentials, and Proximate Determinants*. National Research Council Committee on Population and Demography Report No. 13. Washingon, D.C.: National Academy Press.

Lapham, Robert J. and W. Parker Mauldin (1984). Family planning program effort and birthrate decline in developing countries. *International Family Planning Perspectives* 10(4):109-118.

Long, S. (1983). *Confirmatory Factor Analysis*. Sage University Paper Series on Quantitative Applications in the Social Sciences. Beverly Hills, California: Sage Publications.

Mason, William M., George Y. Wong and Barbara Entwisle (1983). Contextual analysis through the multilevel linear model. In Samuel Leinhardt, ed., *Sociological Methodology 1983-1984*. San Francisco: Jossey-Bass, 72-103.

Mauldin, W. Parker and Bernard Berelson (1978). Conditions of fertility declines in developing countries, 1965-75. *Studies in Family Planning* 9(5):89-147.
Section by Zenas Sykes.

Rosenzweig, M. and T. P. Schultz (1982). Child mortality and fertility in Colombia: individual and community effects. *Health Policy and Education* 2:305-348.

Tsui, Amy Ong (1985). Community effects on contraceptive use. In John Casterline, ed., *The Collection and Analysis of Community Data*. Proceedings of the World Fertility Survey Seminar on the Collection and Analysis of Data on Community and Institutional Factors, London, 20-23 June 1983.

Wong, G. Y. and W. M. Mason (1985). The hierarchical logistic regression model for multilevel analysis. *Journal of the American Statistical Association* 80(391):513-524.

Annex

THE MULTI-LEVEL APPROACH: ILLUSTRATIVE EXAMPLE

*William M. Mason**

This annex illustrates multi-level analysis with fixed- and stochastic-parameter models. In it, the dependent variable, current use of an

---

* Professor, Population Studies Center, University of Michigan, Ann Arbor.

The author gratefully acknowledges the aid of A. I. Her-

efficient method of contraception, is measured dichotomously (using or not using); and logistic response formulations are used to estimate relationships based on actual data.

Although the differences between fixed-effect and stochastic-parameter multi-level models are discussed below in detail, an initial summary of the key difference may be helpful. Multi-level modelling assumes two levels of observation – micro and macro (contextual).[a] At the micro level, the same regression specification is applied within each context. At the macro or contextual level, the parameters of the micro specification are treated as functions of macro variables, one function for each type of micro parameter (e.g., intercept, coefficient of the first regressor). Each of these functions constitutes a macro equation. If the macro equations are conceptualized to allow for errors (disturbance terms), then the multi-level model is said to be a stochastic-parameter formulation. If the macro equations are conceptualized to have no disturbance terms, then the multi-level model is said to be a fixed-effect or fixed-parameter formulation.

Whether the analyst chooses to allow for macro errors has implications for subsequent uses of the information, as well as for estimation and computation. This illustration emphasizes the conceptual differences between fixed- and stochastic-parameter models for multi-level analyses, since these differences will eventually, and ultimately, decide estimation and computation decisions.

At present, the decision to use fixed- or stochastic-parameter multi-level modelling does not rest on conceptual issues alone; the relative costs of implementing the two perspectives also matter. It is possible and practical to estimate multi-level fixed-parameter models using standard estimation procedures, such as are available in widely distributed statistical computer packages (e.g., SAS, BMDP, GLIM and SPSS). It is less practical, although possible, to estimate multi-level stochastic-parameter models because the major statistical packages cannot be used easily (or at all, in most instances) for this purpose.[b] Further, the computations involved in fixed-parameter approaches are less time-consuming, and hence less costly, than those involved in stochastic-parameter approaches. The importance of these differences is likely to decline, however, as access to computers and software continues to diffuse.

A major goal of this annex is to illustrate the perspective from which multi-level stochastic-parameter formulations seem to arise. In so doing, the conceptual reasons suggesting the need for a way to estimate such models should become more clear. It is not essential to describe estimation procedures in order to achieve this expository goal. However, because the illustration requires the use of differing methods of estimation, it may be helpful to mention them at the outset. The method used here for the estimation of multi-level stochastic-parameter models with a dichotomous dependent variable is that of Wong and Mason (1985), which is based on a combination of maximum-likelihood and Bayes procedures. The method of estimation used for multi-level fixed-parameter (or fixed-effect) models is that of maximum likelihood for logistic regression. This is a standard treatment for regression models with dichotomous dependent variables, and it poses no special problems for multi-level analysis if the researcher is willing to make certain restrictive assumptions, which are discussed below.[c]

Because the empirical illustration used is based on a dichotomous response variable, only one type of stochastic-parameter model for multi-level modelling is discussed. Other models are appropriate under different conditions. In particular, if the response variable is normally distributed conditional on the micro regressors (i.e., the case of normal errors), it is possible to estimate a stochastic-parameter model for multi-level regression. Paralleling the distinction between fixed- and stochastic-parameter approaches for modelling a dichotomous dependent variable in multi-level analysis, separate treatments for the case of normally distributed errors are also available. Mason, Wong and Entwisle (1983) present a combined maximum-likelihood Bayes procedure for estimating multi-level stochastic-parameter regressions. The usual method of estimation of multi-level fixed-parameter models is that of maximum likelihood for the case of normally distributed errors. The standard computational procedure for this is ordinary least squares. Restrictive assumptions are again necessary for this type of multi-level fixed-parameter model.

malin, who made the data for this annex available; L. J. Neidert and M. Miele, who assisted in data preparation; A. F. Anderson, who programmed the multi-level computations described here; C. Crawford, who TEXed the document; and B. Entwisle and G. Wong, who provided helpful comments at various stages.

A. Data

The data for this problem concern country A. For this country, a large cross-sectional sample survey was conducted in 1972. The focus of the survey was on fertility and related behaviour. The respondents were married women in their reproductive years. Information was collected not only about the individual respondents but about the villages in which they lived. Thus, there are two levels of observation – individual and village. The analyses reported here are based on 4,327 women living in 56 villages.

The subsample of women used in these analyses were aged from 19 to 42 and not pregnant at the time of the survey, did not believe themselves infecund; had no missing information on their age, education and current contraceptive use; and, if sterilized, had undergone the procedure for contraceptive reasons.

The village-level data are the same as those used by Hermalin (1979), who presents and discusses the actual observations for 25 variables measured on the villages.

B. The problem

Numerous theoretical and empirical studies consider the determinants of contraceptive use in countries undergoing, or poised for, fertility transition (e.g., Entwisle, Hermalin and Mason, 1982 and 1984; Entwisle and others, 1984; and Entwisle, Mason and Hermalin, 1984). On the basis of this prior knowledge, contraceptive use should be more common among respondents with more education and among older respondents. More highly educated women are more likely to be innovators in a variety of areas, including fertility behaviour. They are more likely to gain from contraception, especially for fertility limitation as distinguished from spacing purposes, because they are more likely to occupy positions in the labour force, and in society more generally, in which limiting fertility is perceived to be economically beneficial for the family. Similarly, older women are more likely than younger women to want to terminate childbearing, having already borne children in acceptable numbers and, in relation to younger women, deriving somewhat less benefit from continuing to reproduce.

The possibility of variability across social settings in the impact of micro characteristics on contraceptive use is addressed by a number of authors (Entwisle, Hermalin and Mason, 1982 and 1984; Entwisle and others, 1984; Entwisle, Mason and Hermalin, 1984; and Wong and Mason, 1985). Consonant with much other research on the determinants of contraception, these authors suggest two potentially important dimensions that could underlie setting variability, should it exist, in the micro parameters: socio-economic conditions and family planning programme inputs. In addition, these authors provide *a priori* sign hypotheses for the effects of these dimensions on contraceptive use and provide the rationale for the hypotheses. Although this annex cannot rehearse the theoretical development of the particular substantive problem, the task of developing reasoned hypotheses is, if anything, more important in multi-level analysis than in single-level analysis, because so many of the potential parameters in the model describe (cross-level) interactions as distinguished from more readily interpreted additive effects.

C. Micro specification

In multi-level analysis as presented here and as conceived by Mason, Wong and Entwisle (1983) and by Wong and Mason (1985), it is essential that the micro specification be identical in each context. It is possible to relax this constraint to allow for contextually unique characteristics (Wong and Mason, 1985), but such complexity is unnecessary for present purposes, which pertain to within-country analyses. The need to allow for contextually unique characteristics is more likely to occur in comparative multi-level analysis.

How is the micro specification to be arrived at? For the problem at hand, analysis began with inspection, based on the entire sample of in-

dividuals, of the relationship between the logit of currently contracepting and educational attainment, with education scaled by years of schooling completed. Data exploration suggested that this relationship could be treated as linear. Inspection of the relationship between the logit of contraceptive use and age suggested a monotonic but non-linear relationship, with the rate of increase in contraceptive use decreasing markedly after age 30. Empirical consideration of several formulations for the age effect led to the following transformation of age: the natural logarithm of (age − 18). The principal advantages of this transformation, as opposed to a squared term in age or splines in age, are that collinearity among the regressors is minimized, the number of parameters to be modelled at the macro level is kept to a minimum and it is unnecessary to partition the data by age group. These are significant advantages, because part of the analysis consists of estimating the micro model within contexts. For the small context sample sizes typically available for individual-village analyses, it is helpful to reduce collinearity and to keep the context sample size as large as possible, in order to be able to estimate the within-context parameters.

D. Analysis

Analysis begins with data exploration based on the entire micro data set. As noted above, a logistic regression formulation, appropriate for a dichotomous response variable, is used with previously selected regressors (education and age). At issue is the exact functional form of the logistic regression. Balancing simplicity and parsimony against the slightly better fit that more complex specifications provide, the micro specification employs education scaled in years of school completed and a monotonic but non-linear transformation of age. The estimated form used is of a fixed-effect logistic regression[d] of current use of an efficient method of contraception on education and transformed age. The results are given below:

	Estimated logit (EFF_{ij}) =
Intercept	−2.719
	(0.169)
Education measured as years of schooling completed, *ED*	0.0698
	(0.000907)
Log_e (age − 18), *TAGE*	1.136
	(0.0633)
Goodness of fit	377
Degrees of freedom	2

Goodness of fit is the difference between −2 ln (likelihood) computed for the model estimated here and the null model. The null model contains an intercept but no regressors. Degrees of freedom is the difference between the null model and the model estimated here in the number of regressors. *EFF* refers to efficient contraceptive methods. The subscripts *i* and *j* refer to individual and village characteristics, respectively.

This logistic regression shows that, for these data, current contraceptive use increases with education and age. To clarify the exact nature of these effects, table 14 presents the predicted logits and probabilities of contracepting, conditional on education and age, evaluated at low, middle and high values.[e] The table shows that the education effect is modest compared with the age effect, which increases markedly from ages 20 to 30 and moderately from ages 30 to 40.

Having established the exact form of the micro model, it is now possible to consider whether there is variability, and systematic variability, in the parameters of the micro specification. Estimating the micro model separately for each village is a way of initiating this assessment. Table 15 presents descriptive information on the variability of the estimated micro coefficients across villages. Only the upper panel of the table is of concern at this point in the discussion. The upper panel of table 15 shows that over villages, all three coefficients range from positive to negative, with averages similar to the corresponding coefficients in the pooled logistic regression presented above.

Despite the relatively large village sample sizes, summarized by the stem-and-leaf diagram (Tukey, 1977) presented below, many of the within-village coefficients, especially the education coefficients, are not statistically significant by usual criteria. This is not necessarily a critical problem for multi-level analysis, nor can any conclusions about the outcome of a multi-level analysis be drawn at this point.

In the stem-and-leaf diagram, which is a listing of all village sample sizes and describes the distribution of the village sample sizes, there are, for example, two villages with 43 persons and two villages with 129 persons. The left-hand side of the stem (vertical line) represents 10s; the right-hand side represents units (leaves). The first line displays one 4 (10s) and two 3s (units) and is interpreted as signifying two times 43. Likewise, one 12 (10s) and two 9s (units) reads two times 129 (Tukey, 1977, chapter 1):

```
 4 | 334
 5 | 014555556678
 6 | 2233444789
 7 | 0012245566
 8 | 112235
 9 | 168
10 | 3335
11 | 2556
12 | 7799
```

Is there systematic between-village variability in the parameter estimates of the micro model? This possibility can be checked by plotting the coefficient estimates against potentially relevant village-level variables. Indeed, to assess the potential effects of macro variables, the within-village intercepts, education coefficients and age coefficients

Table 14. Estimated logits and probabilities of current use of an efficient contraceptive method, conditional on education and age[a]

Education	*Age* 20	30	40
Secondary-school graduate (12 years)	−1.09389	0.941008	1.62939
	[0.25]	[0.71]	[0.83]
Primary-school graduate (6 years)	−1.51269	0.522208	1.210596
	[0.18]	[0.63]	[0.77]
No schooling (0 years)	−1.93149	0.103408	0.791796
	[0.13]	[0.52]	[0.69]

Note: Numbers in brackets are estimated probabilities; other numbers are logits.

[a] Using the logistic specification for which the estimated equation is presented in the text.

Table 15. Characterization of within-context coefficients

Coefficient	*Minimum*	*Maximum*	*Mean*	*Standard deviation*
A. *Classical maximum likelihood within-context estimator*				
Intercept	−7.33	1.92	−2.82	1.79
Education effect	−0.17	0.30	0.056	0.092
Age effect	−0.48	2.85	1.20	0.64
B. *Posterior within-context estimator*				
Intercept	−4.83	−0.80	−2.66	0.94
Education effect	0.033	0.074	0.056	0.0088
Age effect	0.50	1.91	1.13	0.33

Note: Computations for the means and standard deviations use the usual sample definitions. These computations are unweighted and based on 56 villages. The label "Classical maximum likelihood within-context estimator" refers to the standard fixed-effect logistic regression model applied separately to each context. The label "Posterior within-context estimator" refers to the method of Wong and Mason (1985) for computing strengthened within-context coefficients based on a multi-level specification. See text of this annex for further discussion of strengthened coefficients.

were plotted against 25 village-level characteristics, for a total of 75 plots (not shown here). That is, all of the intercepts were plotted against each village-level variable, as were all of the education and age coefficients. Inspection of these plots showed that one dimension, the educational level of the village, was clearly an important determinant of two of the types of estimated micro coefficients, namely, the intercepts and age coefficients. So important is this variable that several different operationalizations of it yield equivalent results. The operationalization selected here is a measure of the percentage of females aged 15 and over, not in school, who were primary-school graduates or higher in 1970 (*PSG*). Of the family planning programme variables, the most important is this input measure: cumulative number of woman-months of programme field-work per 10,000 married women aged 20-44 (*FPW*). Whereas a village-level education variable would have been selected from a purely atheoretical search, none of the family planning programme variables would have been, because their associations with the micro coefficients are weak. The *FPW* measure is carried along in this illustration, however, because it is important to assess its impact in a multi-level specification. That is, it is insufficient to rest the case of a null effect on a scatter plot alone, especially when the micro coefficients have been estimated using the standard logistic regression model.

Table 16 consists of a village-level correlation matrix, the upper part of which summarizes the associations of the education and family planning programme variables with each other and with the micro coefficients. Only the first four rows of table 16 are of concern at this point in the discussion. From rows 1-4 of table 16 it can be seen that education composition, *PSG*, is more highly correlated than programme input, *FPW*, with the intercepts and age effects, and that *PSG* and *FPW* are more highly correlated with each other than *FPW* is with either the intercept or age coefficient. This suggests that in a multi-level logistic regression, however estimated, *FPW* will have no effect. The low correlations between the education coefficient and both *PSG* and *FPW* suggest that the education coefficient varies randomly across villages.

TABLE 16. SELECTED CORRELATIONS INVOLVING WITHIN-CONTEXT COEFFICIENTS ESTIMATED TWO WAYS AND TWO MACRO VARIABLES

No.		$\tilde{\beta}_0$	$\tilde{\beta}_1$	$\tilde{\beta}_2$	*PSG*	*FPW*	$\hat{\beta}_0$	$\hat{\beta}_1$
1.....	$\tilde{\beta}_1$	−0.28	-	-	-	-	-	-
2.....	$\tilde{\beta}_2$	−0.96	0.13	-	-	-	-	-
3.....	*PSG*	0.54	0.08	−0.51	-	-	-	-
4.....	*FPW*	−0.30	0.19	0.22	−0.43	-	-	-
5.....	$\hat{\beta}_0$	0.65	-	-	0.98	−0.43	-	-
6.....	$\hat{\beta}_1$	-	0.67	-	0.13	0.04	−0.02	-
7.....	$\hat{\beta}_2$	-	-	0.66	−0.97	0.42	−0.99	0.07

NOTE: $\tilde{\beta}$ denotes use of the maximum-likelihood estimator for the classical fixed-effect logistic regression model and $\hat{\beta}$ denotes use of the posterior within-context estimator defined by Wong and Mason (1985) (see text for further discussion of the posterior, or strengthened, within-context estimator). The subscripts 0, 1 and 2 refer, respectively, to the intercept, the education effect and the age effect. *PSG* = primary-school graduates; *FPW* = programme field-work.

By this point in the analysis, the difficult decisions have been made; those which remain include the choice of a procedure with which to estimate the multi-level model and final specification of the variables to be included in the complete multi-level model. To take stock of the progress, it helps to refer to several equations. The micro model is given by

$$\text{Logit}(EFF_{ij}) = \beta_{0j} + \beta_{1j}ED_{ij} + \beta_{2j}TAGE_{ij} \tag{1}$$

where $i = 1, \ldots, \eta_j$ refers to individual observations within contexts and $j = 1, \ldots, J$ refers to contexts (in this analysis, $J = 56$); $EFF = 1$ if the respondent is currently using an efficient method of contraception and $= 0$ otherwise; *ED* denotes education; and *TAGE* denotes transformed age.

The parameters of the micro model are conceptualized to vary as a function of macro characteristics. Original substantive concerns suggested that both *PSG* and *FPW* should affect the micro coefficients. Under that hypothesis, the macro model is given by

$$\beta_{0j} = \eta_{00} + \eta_{01}PSG_j + \eta_{02}FPW_j + \alpha_{0j}; \tag{2.1}$$
$$\beta_{1j} = \eta_{10} + \eta_{11}PSG_j + \eta_{12}FPW_j + \alpha_{1j}; \tag{2.2}$$
$$\beta_{2j} = \eta_{20} + \eta_{21}PSG_j + \eta_{22}FPW_j + \alpha_{2j}. \tag{2.3}$$

In these equations, the α_{kj}, $k = 0, 1, 2$, are macro errors assumed independent over j for fixed k, but allowed to co-vary across macro equations, just as is usually the case for a block of related equations. For this reason, the full multi-level model is also known as a "covariance components model".

The full multi-level model is given by equation (1) together with equations (2.1)-(2.3). Equivalently, substitution of (2.1)-(2.3) into (1) yields

$$\begin{aligned}\text{Logit}(EFF_{ij}) = {} & \eta_{00} + \eta_{01}PSG_j + \eta_{02}FPW_j + \eta_{10}ED_{ij} + \eta_{20}TAGE_{ij} \\ & + (\eta_{11}ED_{ij} + \eta_{21}TAGE_{ij})PSG_j \\ & + (\eta_{12}ED_{ij} + \eta_{22}TAGE_{ij})FPW_j \\ & + (\alpha_{0j} + \alpha_{1j}ED_{ij} + \alpha_{2j}TAGE_{ij}).\end{aligned} \tag{3}$$

Comparison across the two forms of the full set of equations, ((1) and (2.1)-(2.3), as against (3)) shows that the coefficients of the macro equations sustain these interpretations: η_{00} is the overall intercept of the multi-level model; η_{01} is the main effect of *PSG*; η_{02} is the main effect of *FPW*; η_{10} is the main effect of education; η_{11} is the micro-macro education interactive effect; η_{12} is the micro-macro education-programme interaction effect; η_{20} is the main age effect; η_{21} is the micro-macro age-educational composition interactive effect; and η_{22} is the micro-macro age-programme interactive effect. Since all of the ηs except η_{00} are cofficients describing either main or interactive effects in the logit metric, it follows that one can develop hypotheses about them and interpret the ηs just as would be done in standard logistic regression. Furthermore, one can also plot estimates of the β_{kj} against macro variables in order to gain impressions of likely direction and strength of particular ηs.

The initial examination of zero-order macro plots suggests that estimation of the full multi-level specification will result in: $\hat{\eta}_{01} > 0$, $\hat{\eta}_{02} \approx 0$, $\hat{\eta}_{11} \approx 0$, $\hat{\eta}_{12} \approx 0$, $\hat{\eta}_{21} < 0$, and $\hat{\eta}_{22} \approx 0$. Also, the positive mean education effect reported in table 15 suggests that $\hat{\eta}_{10}$ should be positive. In addition, *a priori* substantive considerations, the results of the pooled logistic regression, the clear positive mean age effect over villages (table 14) and the negative interaction effect suggested by the correlation between the education coefficient and *PSG*, as well as the scatter plot between these variables, all suggest that $\hat{\eta}_{20}$ should be positive. Lastly, on *a priori* grounds and the evidence thus far, $\hat{\eta}_{00}$ should be negative (corresponding to a probability less than 0.5): contracepting illiterate young women living in environments in which most people are illiterate and in which no family planning programme effort is being made should be a minority.

E. CHOICE OF STATISTICAL MODEL

Testing the tentative conclusions reached thus far requires a statistical model. The two alternatives available (apart from the non-fundamental choice between, for example, probit and logit models) are the standard fixed-effect logistic response model and the stochastic-parameter logistic response model.

In a fixed-effect model, the specific contexts are assumed to be a focus of interest, just as are the groups or categories in a classical fixed-effect analysis of variance. From this perspective, the use of macro variables instead of indicator variables for contexts is simply a way to structure parsimoniously the between-context contrasts. A fixed-effect model containing micro and macro variables, as well as micro-macro interactions, can be arrived at without a multi-level conceptualization and without postulating the dependence of micro parameters upon macro characteristics.

A fixed-effect specification can also be arrived at by assuming that the macro equations do not contain error terms, that is, that the α_{kj} are uniformly zero. This is evident from the equivalence of equation (3) with the combination of (1) and (2.1)-(2.3). Does this mean that specifying no macro error terms is equivalent to assuming that the contexts are fixed and that interest is confined solely to the observed contexts in a particular analysis problem? Not necessarily. However, specification of no macro error terms does require the analyst to postulate not only that any new contexts added to the analysis, or collected in a replication, would have the same values on the macro predetermined variables (e.g., *PSG* and *FPW*), but that these new contexts would have the same micro

coefficients. That is, the specification of no macro errors requires the analyst to assume that any additional contexts observed would merely provide replicates of, for instance, the combinations (β_{kj}, PSG_j, FPW_j), for $j = 1, \ldots, J$; $k = 0, 1, 2$. This assumption is unrealistic on two grounds: First, it does not allow for the possibility that the micro coefficients for a new set of contexts with the same macro variables could ever be different. Secondly, it assumes that the analyst has specified perfectly the determinants of the micro coefficients. Thus, suppression of the macro error terms exacts a toll in the realism of the underlying assumption of the model, given that the analyst wishes to treat micro coefficients as endogenous with respect to macro characteristics, or given that the analyst's ultimate focus remains with the particular set of contexts studied in a particular analysis.

The stochastic-parameter approach can be arrived at in three ways. Equation (3) is an instance of a mixed model, in which there are fixed effects (the ηs) and random effects (the αs). Mixed models arise from situations in which groups or contexts are sampled, and there is interest in generalizing from the sample to a larger population. From the standpoint of the macro equations (e.g., equations (2.1)-(2.3)), allowing for the αs is equivalent to assuming that the β_{kj} can differ even when the macro characteristics remain the same in, for example, a replication. In addition, the αs allow the analyst to distinguish between systematic and random variability in the β_{kj}. Lastly, from a Bayes perspective, the stochastic-parameter model assumes that the errors are exchangeable across contexts, even if contexts are not sampled.[f] Thus, even if the analyst's interest is ultimately focused on just the set of contexts available in a particular set of data, it is still possible to treat micro parameters as endogenous with respect to macro characteristics and to allow for error in the macro specification.

This sketch of the fixed-effect and stochastic-parameter perspectives indicates that, all other things being equal, the choice of approach should be guided by the analyst's goals. If the analyst does not wish to treat the micro parameters as endogenous with respect to other macro characteristics and if inference is limited to the set of contexts currently available for analysis, then the fixed-effect approach seems appropriate. In that case, the analyst can make use of widely available computer programmes for maximum-likelihood estimation of fixed-effect logistic response models, in order to estimate equation (3) with all terms in a excluded. If the analyst wishes to treat the micro parameters as endogenous with respect to other macro characteristics, or the contexts are sampled, then the stochastic-parameter approach is appropriate.

The two approaches can lead to coefficient estimates that are quite similar, as is shown below. Since the underlying assumptions of the approaches differ fundamentally, however, interpretations based on the different assumptions must reflect these differences if the analyst is to be logically consistent.

F. Results: fixed-effect logistic regression

Table 17 presents two estimated fixed-effect logistic regressions. Regression 1 corresponds to equation (3), except that no macro error terms are conceptualized or estimated. For this regression, all coefficients involving the family planning programme input variable are small in relation to their standard errors. Regression 2 omits all terms involving *FPW* as well as the micro education-educational composition interaction. Comparison of the goodness of fit statistics for the two regressions shows no appreciable deterioration in fit (the difference 431 − 426 = 5 is chi-square distributed with four degrees of freedom under the null hypothesis that all coefficients involving *FPW* are zero). All terms included in regression 2 are significant at conventional levels.

Regression 2 in table 17 indicates that education is positively related to current contraceptive use, controlling age and educational composition in the local environment. The education effect controlling for programme input and the age-education composition interaction is only slightly lower than the education effect observed in the results of the pooled logistic regression given in section D. The age and education composition effects are inseparable, because of their interaction. For these villages, the age effect is 2.519 − 0.027 *PSG* and the education composition effect is 0.078 − 0.027 *TAGE*. Evaluation of these partial derivatives for a possible extreme value within the ranges of the variables shows, first, that the age effect is always positive but decreasing as the educational composition of the local environment increases—this decrease is in addition to the decrease captured by the functional form for age itself (log (age − 18)). Secondly, the effect of educational composition in the local environment is positive and appears to peak at about age 37, with a slight decline thereafter. By virtue of the fixed-effect assumption one does not, for example, treat the rate of change in the age effect as a consequence of changes in educational composition. Moreover, statistical inference about the age, education and educational composition effects is conditional on the sampled villages.

Table 17. Fixed-effect logistic regressions of current use of efficient contraception, based on data pooled across contexts and incuding contextual characteristics as regressors

Regressor	Coefficient symbol	Logistic regression 1	Logistic regression 2
Intercept	η_{00}	−5.66	−6.7129
		(1.2335)	(0.7384)
PSG	η_{01}	0.064235	0.07779
		(0.01676)	(0.01345)
FPW	η_{02}	−0.001850	-
		(0.002837)	
ED	η_{10}	−0.08920	0.05623
		(0.07043)	(0.009619)
ED·PSG	η_{11}	0.001766	-
		(0.0009149)	
ED·FPW	η_{12}	0.0002879)	-
		(0.0001715)	
TAGE	η_{20}	2.3778	2.5186
		(0.4607)	(0.3000)
TAGE·PSG	η_{21}	−0.02442	−0.02663
		(0.006309)	(0.005486)
TAGE·FPW	η_{22}	0.0001056	-
		(0.001046)	
Goodness of fit		431	426
Degrees of freedom		8	4

Note: Figures in parentheses are estimated standard errors. *PSG* = the compositional measure of education measured at the contextual level; *FPW* = the family planning programme input variable measured at the contextual level; *ED* = education of respondent; *TAGE* = transformed age of respondent. Goodness of fit is the difference between −2 ln (likelihood) computed for the model estimated in the text and the null model. The null model contains an intercept but no regressors. Degrees of freedom is the difference between the null model and the model estimated here in the number of regressors.

G. Results: stochastic-parameter multi-level logistic regression

Table 18 presents three stochastic-parameter regressions. These regressions are computed using the iterative procedure proposed by Wong and Mason (1985). The beginning values for the iterations are the classical maximum-likelihood within-context coefficient estimates. That is, standard logistic regressions are computed separately within each context. The resulting coefficients are then used to initiate the iterative computations described by Wong and Mason.

The first regression given in table 18 is an estimated random coefficient logistic regression model. It is the estimated form of (3) with all coefficients involving *PSG* and *FPW* constrained to be zero. Equivalently, it is the estimated form of equations (2.1)-(2.3) allowing for αs and with only the η_{k0} ($k = 0, 1, 2$) allowed to be non-zero. This is the baseline equation, and its coefficients are quite similar to those of the fixed-effect equation presented in the pooled logistic regression.

Regression 2 in table 18 is an estimate of the full model given by (3). For this regression, the coefficients involving *FPW* are small in relation to their standard errors, suggesting that all terms in *FPW* can be omitted. Regression 3 omits the terms in *FPW* and allows the education effect to be random. The coefficients in regression 3 are all large in relation to their standard errors. Moreover, the outcome of a Wald test of the null hypothesis that $\eta_{02} = \eta_{11} = \eta_{12} = \eta_{22} = 0$ sustains the simultaneous omission of the four terms.

TABLE 18. STOCHASTIC-PARAMETER MULTI-LEVEL LOGISTIC REGRESSIONS OF CURRENT USE OF EFFICIENT CONTRACEPTION

Regressor	Coefficient symbol	Logistic regression 1	2	3
Intercept	η_{00}	−2.6366	−5.6575	−6.6692
		(0.2030)	(1.3179)	(0.7766)
PSG	η_{01}	-	0.06427	0.07741
			(0.01793)	(0.01416)
FPW	η_{02}	-	−0.001884	-
			(0.003044)	
ED	η_{10}	0.06023	−0.08830	0.05505
		(0.01023)	(0.07586)	(0.01034)
ED·PSG	η_{11}	-	0.001702	-
			(0.0009953)	
ED·FPW	η_{12}	-	0.0002966	-
			(0.0001831)	
TAGE	η_{20}	1.1180	2.3940	2.5210
		(0.07510)	(0.4985)	(0.3191)
TAGE·PSG ..	η_{21}	-	−0.02449	−0.02663
			(0.006821)	(0.005850)
TAGE·FPW ..	η_{22}	-	0.00006197	-
			−(0.001141)	

Source: Calculated using the method of Wong and Mason (1985).

NOTES: Figures in parentheses are standard errors. *PSG* = the contextual measure of education; *FPW* = the contextual measure of family planning programme input; *ED* = education; *TAGE* = transformed age of respondent.

For description of regressions 1, 2 and 3, see text.

The coefficients of regression 3 in table 18 show that the effect of education varies randomly over villages and has a positive mean of 0.055. The age effect also varies, with both a systematic and a random component. The partial derivative for the effect of age is 2.521 − 0.0266 *PSG*, which indicates that although the effect of transformed age is positive, the rate of increase in contraceptive use decreases as a function of educational composition. That is, the multi-level conceptualization is consistent with the view that a change in the level of educational composition leads to a change in the age effect. In addition, since *TAGE* is the natural logarithm of translated age (shifted by 18), there is a decline in the rate of increase of the age effect that exists apart from the decline captured by *PSG*. Whether the aspect of the decline implicit in *TAGE* is due to other macro phenomena, and if so, what they might be, is unclear without further research. For example, it may be that the decrease in the age effect implicit in the use of *TAGE* is due to cohort differences in willingness to use contraception.

The effect of educational composition, as judged from regression 3 in table 18, is also assayed by a partial derivative, which is 0.077 − 0.0266 *TAGE*. Evaluation of the partial derivative shows that there is a point of inflection at 36 years of age. The multi-level statement of this result is that educational composition has its greatest positive impact on contraceptive use at the youngest ages of married women and that this impact actually becomes negative for persons approaching 40. This negative effect is quite small. For example, at age 40, the *PSG* partial derivative is − 0.005. It is unclear how to interpret this negative portion of the *PSG* effect. Perhaps the most important aspect of the *PSG* effect is that, over most of the years relevant to childbearing, people appear to experience the impact of the educational composition of the local environment positively, but this aspect of local environment appears to become less relevant at the upper end of the age range for childbearing. In sum: the younger the person, the stronger the composition effect on contraceptive use.

The estimated macro-error variances and covariances provide additional information about the nature of the fit of regression 3. These are presented in table 19, which shows that the maximum-likelihood estimate of the between-village variance in the education effect is virtually zero (0.000588), that the cross-equation error correlation between the education-effect and age-effect equations is about 0.5 and that the errors of prediction in the intercept macro equation are inversely correlated with the errors of prediction for the remaining macro equations. That the error correlation for the intercept and age-effect equations is so large is a reflection of the strength of the age effect in the within-context logistic regressions. The surprisingly strong correlation estimated between α_1 and α_2 suggests that there remains between-village variability in the effects of education and age yet to be accounted for. This variability need not be systematic. Indeed, it is clear that the variables studied in the exploratory phase of this analysis (referred to, but not displayed here) cannot account for this between-village variability. On the strength of the evidence thus far, it is a reasonable working hypothesis that the extent to which the errors of the macro equations are correlated is nothing more than a reflection of the characteristics that make each village unique.

TABLE 19. ESTIMATED ERROR VARIANCES AND STANDARDIZED MACRO ERROR COVARIANCES OF REGRESSION 3 IN TABLE 18

	a_0	a_1	a_2
α_0	0.213	-	-
α_1	−0.630	0.000588	-
α_2	−0.945	0.522	0.0362

NOTE: The off-diagonal terms are the estimated macro error covariances divided by the positive square roots of the products of the respective estimated macro error variances, which are presented on the main diagonal.

The stochastic-parameter approach also allows for posterior estimates of the within-context regressions (Wong and Mason, 1985). These estimates are strengthened in relation to classical maximum-likelihood fixed-effect logistic regressions, because they take into account information from all other villages. The classical within-context logistic regressions simply apply the standard logistic regression model separately to each context. Strengthening is justified to the extent that the assumption of random as is valid, or to the extent that the assumption of exchangeability of the as is valid. In the present instance, the assumption of exchangeability seems to be reasonable—on the basis of prior knowledge of the contexts. To some degree, the validity of this assumption can be checked empirically, for example, by plotting the macro residuals against other macro variables.

The posterior estimates of the micro regressions are summarized in the lower panel of table 15. From table 15 it can be seen that, for the posterior estimates, the intercept is always negative, and the education and age effects are always positive. This is in marked contrast to the classical estimates (upper panel of table 15). Moreover, as the shrunken ranges of the posterior estimates suggest, the dispersion of the posterior estimates as measured by the classical sample estimator of the standard deviation should be smaller than those for the classical maximum-likelihood within-context coefficients. This is in fact the case.

The posterior estimates differ from the classical maximum-likelihood estimates by more than just shrinkage. This is made clear in figures VI and VII, which plot the posterior intercepts and age effects, respectively, against *PSG*. Inspection of these figures, as well as the correlations involving *PSG* in table 16, shows that the classical within-context coefficient estimates are not nearly so strongly related as the posterior estimates are to *PSG*. If additional macro data with which to model coefficient variability were to become available, the posterior estimates of the within-context coefficients could be used for additional exploratory purposes. Further, the differences between posterior and classical estimates can also be used for additional study of the data. For example, if differences between the two forms of estimates for a given village are large, this may indicate an unusual village. Consideration of this village could lead to a re-evaluation of the validity of the assumption of randomness (or exchangeability), or it could lead to new sustantive insight and the alteration of the terms included in the macro or micro equations. If within-context micro equations were needed for prediction purposes, as they sometimes are (e.g., when attempting to predict student success in school admissions processes), then the posterior estimates could be used in preference to the classical estimates.

Figure VI. Plot of posterior micro intercepts against contextual education composition, *PSG*, based on regression 3 in table 18

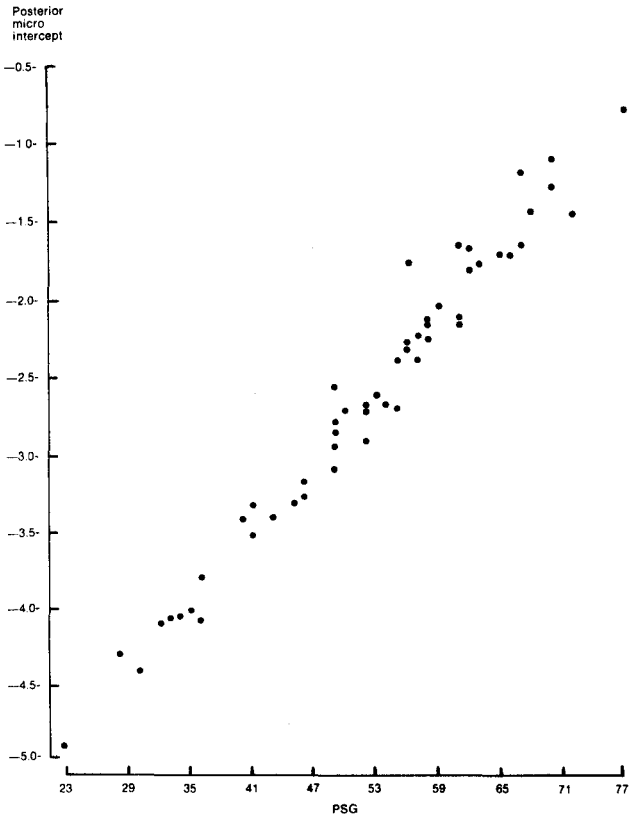
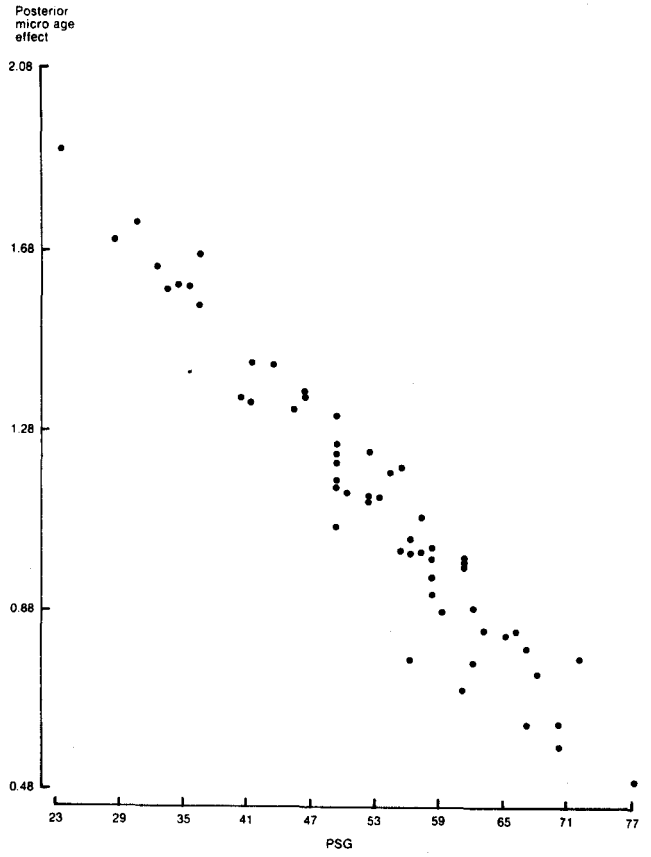


Figure VII. Plot of posterior micro age effects against contextual education composition, *PSG*, based on regression 3 in table 18



Summary and conclusions

The steps in multi-level modelling consist of the following: (*a*) selection of a micro model on the basis of *a priori* substantive considerations; (*b*) theoretical development of micro-macro hypotheses concerning the main effects of macro variables and the effects of cross-level interactions; (*c*) data exploration—at both the micro and macro levels. At the micro level, data exploration is used to refine the initial specification. At the macro level, data exploration consists of graphical examination of scatter plots of micro coefficients against macro variables, as well as application of simple methods of multivariate analysis, such as multiple regression using ordinary least squares. Lastly, true multi-level statistical estimation is carried out.

The choice between use of a fixed-effect approach and a stochastic-parameter approach is fundamentally conceptual, although until appropriate stochastic-parameter software is routinely available and computer costs are reduced from current levels, the lesser expense of fixed-effect modelling will remain appealing. The conceptual choice can be clarified with a schematic table that summarizes the earlier comments about the differences between the fixed-effect and stochastic-parameter approaches:

	Interest is in a specific set of contexts	*Interest is in a universe of contexts, which has been sampled*
The micro coefficients are conceptualized to depend upon contextual variables	(1)	(2)
The micro coefficients are not conceptualized to depend on contextual variables	(3)	(4)

Cases 1, 2 and 4 pertain to situations in which stochastic-parameter models are appropriate, case 3 to a situation in which a fixed-effects approach is appropriate.

In the empirical example presented here, the estimated h-coefficients of the fixed-effect and stochastic-parameter approaches are virtually identical. This cannot be expected to occur in all instances. Moreover, even when the coefficient estimates are similar, the standard errors estimated by the two approaches can be quite different. Relevant factors appearing to contribute to the convergence of results include the relatively large number of contexts, large sample size per context, and the degree of cross-context variability in the coefficients. The relative weights of these factors in determining the convergence of the two approaches have yet to be studied systematically.

Notes

[a] More levels of observation are possible, but appropriate data are scarce and statistical methods for dealing with them are not currently available in a practical form.

[b] FORTRAN programs for stochastic-parameter multi-level modelling are available from the author.

[c] When the dependent variable is dichotomous, the regression formulation assuming normally distributed errors is, in general, incorrect. The logistic response model is a preferred alternative in this case, regardless of whether the empirical problem is multi-level or single level. Logistic response models are exposited in textbooks, such as those by Hanushek and Jackson (1977), Fox (1984), Cox (1970) and Maddala (1983). In this annex, the standard logistic response model, or logistic regression model, is also referred to as the classical fixed-effect logistic regression model. This additional terminology is used to help distinguish the conventional model from the stochastic-parameter model for multi-level logistic regression.

[d] Widely distributed statistical computer packages, such as SAS, BMDP, SPSS and GLIM, include the capability of estimating standard fixed-effect logistic regressions using the method of maximum

likelihood. The SAS program also includes the minimum logit chi-square method, which is an alternative to maximum likelihood for cases in which there are replicates for each combination of values of the predictor variables. This alternative method was not used here, as it is less generally applicable than maximum likelihood.

[e] The logit is defined as log (p/q), where $p = Pr(Y=1)$ and $q = 1 - Pr(Y=1)$, and log denotes natural logarithm. For expository discussion of logits, logistic regression and the connection between estimated logits and estimated probabilities, the presentations of Hanushek and Jackson (1977) and Fox (1984) are quite useful.

[f] If the exchangeability assumption is satisfied, one would be indifferent to any permutation over j of $(\alpha_{0j}, \alpha_{1j}, \alpha_{2j})$. That is, one would be indifferent to arbitrary mixing-up of the macro errors—it would make no difference if the errors for contexts j and j' were switched with each other.

References

Cox, D. R. (1970). *The Analysis of Binary Data*. London: Methuen.

Entwisle, Barbara, Albert I. Hermalin and William M. Mason (1982). *Socioeconomic Determinants of Fertility Behavior in Developing Nations: Theory and Initial Results*. National Research Council Committee on Population and Demography Report No. 17. Washington, D.C.: National Academy Press.

_____ (1984). A model for the comparative analysis of WFS contraceptive use data. In John A. Ross and Regina McNamara, eds., *Survey Analysis for the Guidance of Family Planning Programs*. Proceedings of a seminar organized by the International Union for the Scientific Study of Population. Liège: Ordina Editions, 127-146.

_____, William M. Mason and Albert I. Hermalin (1984). *The Multilevel Dependence of Contraceptive Use on Socioeconomic Development and Family Planning Program Strength*. Population Studies Center Research Report 84-60. Ann Arbor: University of Michigan.

_____ and others (1984). A multilevel model of family planning availability and contraceptive use in rural Thailand. *Demography* 21(4):559-574.

Fox, J. (1984). *Linear Statistical Models and Related Methods*. New York: John Wiley.

Hanushek, Eric and John E. Jackson (1977). *Statistical Methods for Social Scientists*. New York: Academic Press.

Hermalin, Albert I. (1979). Multivariate areal analysis. In *Manual IX. The Methodology of Measuring the Impact of Family Planning Programmes on Fertility*. New York: United Nations, 97-111. Sales No. E.78.XIII.8.

Maddala, G. S. (1983). *Limited-Dependent and Qualitative Variables in Econometrics*. Econometric Society Monographs in Quantitative Economics. Cambridge: Cambridge University Press.

Mason, William M., George Y. Wong and Barbara Entwisle (1983). Contextual analysis through the multilevel linear model. In Samuel Leinhardt, ed., *Sociological Methodology 1983-1984*. San Francisco: Jossey-Bass, 72-103.

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

Wong, G. Y. and W. M. Mason (1985). The hierarchical logistic regression model for multilevel analysis. *Journal of the American Statistical Association* 80(391):513-524.