

Integrating Theory and Research on the Institutional Determinants of Fertility

Herbert L. Smith

Population Studies Center, University of
Pennsylvania, 3718 Locust Walk, Philadelphia,
Pennsylvania 19104-6298

This article links recent conceptual theories regarding the determinants of fertility with research designs appropriate for testing those theories. The most important causal factors in these theories—typically social or cultural institutions, occasionally emergent properties of the collective behavior of individuals—are properly conceptualized at the macro level. Research designs must therefore feature variation at this level and are at a minimum *comparative*. Noncomparative micro-level research designs will be misspecified from the standpoint of theory. A case is also made for continuing to incorporate the observation of individual-level behavior into any comprehensive research design. A focus on macro determinants of fertility does not imply that fertility outcomes are determined at the institutional level. Instead, there remains some process whereby systemic properties are translated into individual behaviors. Recommendations for comparative community-level studies are discussed. The multilevel analysis framework is reviewed as a paradigm for the conceptual features of cross-contextual analysis.

It is widely agreed that we do not have an adequate *theory* of fertility, if by theory we mean a coherent body of analyses linking a characterization of society and economy, aggregate or local, to *individual* fertility decisions and outcomes, able to withstand scrutiny against the empirical record. (McNicoll, 1980:441; also cited in Entwisle and Mason, 1985)

Until recently, the dominant fertility research agenda stressed the socioeconomic perspective—the effects of socioeconomic structural factors on fertility and fertility decline. This perspective includes traditional demographic transition theory (Notestein, 1953; Thompson, 1929) and its elaborations (Coale, 1973; Teitelbaum, 1975). The socioeconomic perspective assumes that individual fertility behavior is affected by individual socioeconomic variables and that both fertility change and fertility differentials within populations are functions of shifts in the marginal distributions of the socioeconomic variables, induced by exogenous change in economic and social structures.

Recent years have not been kind to the socioeconomic perspective. The surfeit of data accumulated through the World Fertility Survey (WFS) and other surveys has failed to herald the triumph of the socioeconomic perspective and, instead, has led to calls for the consideration of other factors as determinants of fertility (e.g., Cleland and Hobcraft, 1985). “The fact that parental education and cultural factors, denoted by language, ethnicity, or region, emerge as major independent determinants of the onset of decline is more consistent with idealism than structural theories” (Cleland, 1985:247). Similarly, Gille’s (1985) summary of the policy implications of WFS research minces no words: “WFS has failed to identify any single aspect of development that is the crucial lever of demographic change. . . . Its

onset appears to be determined more by ill-understood cultural factors than by any objectively ascertainable development indicators" (p. 279).

Renewed emphasis in fertility theory on "cultural factors" and the cultural context of high fertility has not as yet had a great impact on fertility research. Nearly a decade later, McNicoll's (1980:441) characterization of fertility theory remains a *propos*, and for the same reason: the general failure to link the institutional determinants of fertility—which inhere at some aggregated level—with individual-level fertility behavior. That such a situation obtains in spite of recent developments in the statistical analysis of multilevel data (Mason, Wong, and Entwisle, 1983; Wong and Mason, 1985) suggests that the problem may lie at the interface of theory and research, that is, *research design*.

This article establishes linkages between recent conceptual theory regarding the determinants of fertility and research designs for testing those theories. The first section reviews several prominent theoretical statements pertaining to the institutional determinants of fertility. In the second section, a case is made for continuing to incorporate the observation of individual-level behavior into any comprehensive research design. The third section considers research designs that feature the macro variation deemed essential in the first section and the micro variation advocated in the second section.

Institutional Aspects of Fertility Theories

This section gives three examples of theories that relate aspects of the social and community structure to fertility. In each case, primary conceptualization is at the systemic, institutional, or macro level. Research designs lacking in macro variation are unable to shed any light on the validity of these theories and are unlikely to feature a meaningful translation of theory into empirical research.

1. Caldwell's (1982) theory of fertility and fertility decline brings into focus the institutional, structural character of the factors both supporting high fertility and, eventually, precipitating its decline. Patriarchy is a social institution subsuming large numbers of women within families, families within kin groups, and kin groups within communes or villages. The subordination of youth to their elders and of men to women is not a feature of particular households, families, or kin groups, but of the larger social structure. Individual variation (deviance) is of little account when arrayed against the larger forces militating for conformity to essential behaviors, including fertility (Caldwell, 1982:172).

When change comes, it comes not through the collective exercise of individual choice, but through the collapse of a larger system that had heretofore constrained all choices of behavior open to individuals. Theories of modernization that concentrate on personality change are criticized, due to their implication that "individuals could always have lived different ways of life by opting to do so, whether or not the needed economic and social institutions for the new way of life had yet come into existence" (Caldwell, 1982:280).

A primary source of fertility change is mass education. Mass education creates both educated (expensive, ungrateful, questioning) children and educated wives. Educated wives reduce the net wealth flow from wife to husband (and mother-in-law and father-in-law), strengthen the bonds between husband and wife (undermining the traditional family structure and its morality), and seek to avoid repeated pregnancies and periods with infants.

Education is easily measured at the individual level, and its incorporation into micro models of fertility and fertility-related behaviors is on occasion justified with reference to Caldwell's (1982) emphasis on education as a source of fertility decline. But Caldwell unambiguously points to the macro properties of education: "[T]he education of only half the community does not have the same effect on that half of the population, nor half the effect on the whole population" (1982:329). When there remain many in a community who have not attended school, strong forces maintaining the traditional family morality still

abound. “[T]he evidence suggests that the most potent force for change is the breadth of education (the proportion of the community receiving some schooling) rather than the depth (the average duration of schooling among those who have attended school).” Not only is the distinction manifest between mass schooling as a macro property and individual education as a micro property, but so too is the theoretical hegemony of the former.

2. Cain, Khanam, and Nahar (1979; see also Cain, 1980, 1981) considered the motivation for high fertility in rural South Asia. According to their analysis, there are two major systems of stratification in Bangladesh. One is grounded in class relations—that is, ownership of land and other resources. The other is a system of gender stratification, as maintained by the institution of patriarchy. Patriarchy is sustained via traditional kinship patterns (arranged marriages, postmarriage patrilocal residence), politics (exclusive male domination of formal political and quasi-political structures), and religion (laws of inheritance, *purdah*). The system of gender stratification is in many respects independent of the system of class stratification. The sexual division of labor is rigid across class lines. The variation in patterns of female labor that exists occurs by age/family role (daughter-in-law vs. mother vs. mother-in-law); and the age hierarchy of female labor is also relatively invariant across economic class.

Women’s dependence on men with respect to authority and material resources has as partial recompense the reciprocal obligations of men to women—to provide food, shelter, and clothing—plus the psychic, emotional, and prestige benefits conferred on women through the maintenance of *purdah*. However, whereas the subordination of women to men has a material basis, the responsibilities of men toward women are supported only by a normative order. Under conditions of rising poverty and landlessness, this normative order has eroded, although the material dominance of men has not.

Cain, Khanam, and Nahar (1979), like Caldwell (1982), discussed properties of social and economic systems and institutions. “The risk and insecurity that patriarchy imposes on women represent a powerful systemic incentive for high fertility. . . . [T]he institution of *purdah* confers social status upon women, while at the same time serving as an instrument of repression” (p. 432). The explicitly macro focus is also evident in the discussion of the role of labor markets in maintaining gender inequality: “The underemployment of these women and of poor women in general relative to men is primarily a consequence of highly segregated markets for male and female labor. . . . [M]arket segregation is both a consequence of and a means for perpetuating the system of patriarchy” (p. 428).

3. Mason (1984, 1986) described multiple paths by which various aspects of female status may impinge on both fertility and infant and child mortality. Crucial in delimiting the scope of the review is a definition of women’s status that focuses on gender inequality: The status of women obtains in comparison with that of men. Dimensions of gender stratification include control over material and social resources, prestige, and dependence/control/autonomy. Control over resources is antecedent to the other two dimensions, and only the dependence/control/autonomy complex directly affects fertility and mortality processes. All three dimensions of gender inequality derive from economic and kinship institutions at the community or societal level.

As in Cain, Khanam, and Nahar’s (1979) analysis, the conceptual distinction between institutions of gender stratification and those of class and socioeconomic status coincides with a disjuncture in levels of analysis. Class and socioeconomic status will vary among individuals in all settings, but gender stratification is an aspect of social settings that is constant across individuals, regardless of class or socioeconomic status (Mason, 1984:20, 77).

Although theory relating institutional factors to fertility is far from seamless, there is substantial support for the proposition that the key elements impinging on fertility inhere

at the systemic, institutional, aggregate level. Given the emphasis on macro properties of institutions (e.g., women's legal status, mass education, systems of gender stratification), Mason (1984:77) recommends that researchers concentrate more on aggregate-level analyses of the status of women and fertility. Dyson and Moore's (1983) study of the relationship between female autonomy and demographic regimes is an example of a descriptive analysis based on modal kinship patterns of Indian states, with a theoretical analysis extended to the regional and national level.

Whether purely national-level analyses are appropriate is debatable. Cross-national regression analyses of fertility decline have been successful in explaining variation in vital rates [e.g. Cutright, 1983; Mauldin and Berelson, 1978 (and references therein)]. This is "in spite of the lack of complete information on the motivational linkage between the indicator[s] and fertility" (Cutright, 1983:107); if policymakers know that "percent literate" is negatively associated with the crude birth rate, then the whys and wherefores of this association are presumably of little import. Without gainsaying the value of such studies, the time boundedness of the samples and the problems of operationalization of concepts and measurement of variables raise questions about whether the regression equations characteristic of such studies are structural in any real sense of the term. Nation-states are for the most part insufficiently homogeneous with respect to systems of culture and stratification (Tilly, 1984). Aggregate-level statistics may be insufficient proxies for integral properties of social institutions (e.g., Mauldin and Berelson, 1978:139). Moreover, most of the theory regarding the relationship among variables measured at the aggregate level pertains less to the emergent properties of these variables and more to the aggregated demographic response to change in individual circumstance (e.g., Davis, 1963). In the following section, I argue the case in favor of continuing to examine fertility at the micro level and of integrating macro explanatory variables at the micro level—neither of which is possible through purely aggregate-level research designs.

Keeping Individuals in the Picture

Locus of Causes and Locus of Effects

As argued in the preceding section, several important theoretical discussions of fertility have as primary causes variables that are explicitly conceptualized at the aggregate (systemic, society, community, institutional) level. This does not mean, however, that the endogenous variables—fertility and its proximate determinants—should be conceptualized and operationalized at the aggregate level. Instead, variation in fertility is typically observed at the individual level, as a function of both individual physiological differences and, more important, the attitudes and behaviors of women and their spouses.

Some arguments stress population and fertility goals from the standpoint of the society or community—demographic rationality at the aggregate, rather than individual, level (e.g., Wrigley, 1978). Consider Ryder's (1983) account of the macro-analytic model of the demographic behavior of a traditional society:

In the macro-analytical model . . . the group is the focus, and the requirements for survival of the group are paramount. The individual is viewed less as a decision-maker than as a servant of the society, charged with responsibilities to the group, such as the responsibility for producing replacements . . . for a population structure continually depleted by mortality. The individual is programmed by group processes to fill a role in the societal blueprint. (pp. 19–20)

The problem with reifying populations as actors with demographic interests was spelled out by Lesthaeghe (1980):

The problem with [this] proposition is that it refers to population growth regulation *sui generis* as a target, without exploring the extent to which demographic homeostasis could be the result of manifest short-run goal-setting (conscious rationality) by individuals with respect to *other facets* of social life. (p. 350)

Instead, demographic homeostasis is embedded in the larger exigencies of social organization:

In traditional societies the regulation of the right to reproduce constitutes an appropriation of female labor forces and of sexual gratification by those who, by virtue of age and descent, form the ruling group in an asymmetric authority structure and exchange system. (p. 350)

Various forms of gender inequality and female subordination are only aspects “of an overall pattern of autocratic control over people in lineages and crops on lineage land.” Thus a focus on social structure as the locus of gender stratification does not necessarily presuppose that populations or communities are the loci for fertility control.

Roles for the Individual in the Institutional Perspective

Although the institutional perspective on fertility has an explicitly macro-analytic orientation, there are elements of this perspective that suggest the need for micro (individual)-level data as well. Without ascribing to the authors cited in this section an explicit recommendation for a micro-oriented research agenda, the case can be made that there are at least implicit calls for consideration of the role of individual heterogeneity in the larger theoretical context.

According to Caldwell (1982), there are only two types of fertility regimes—those in which there is no economic rationality to limiting fertility and those in which economic rationality dictates complete fertility control. The only exception occurs in transitional societies, which presumably covers all cases of intermediate-level fertility. Intermediate levels of fertility among transitional societies are products not of the temporary economic rationality of intermediate numbers of children, but rather of the simultaneous presence of classes of families living under the old and new fertility regimes. This schema has at least two implications. The first is that any given population’s fertility behavior, when properly measured and analyzed, would give rise to two distinct, mutually exclusive, exhaustive classes of families. The second implication, given the validity of the first, is that aggregate-level analyses of fertility rates (regardless of whether they are based on intertemporal or cross-sectional data) are likely to lead to spurious reifications of gradations in fertility decline, which should more properly be ascribed to the (two-class) heterogeneity of the population. These are strong empirical propositions, but they appear to follow from the theoretical formulation. Insofar as they might be shown to be true (or proved false), they demand a hard look at the individual-level characteristics of the study population. Such a study might seek, for example, to identify the extent of emotional or economic nucleation within families/households/kin groups—a precursor of the shift from a pretransitional to a posttransitional way of life.

There is a second element of the institutional perspective that appears to involve a role for individuals and individual-level variation. It identifies the shift from traditional to transitional society with the elevated importance of the interests of the individual relative to those of the group. Thus Ryder (1983) equated modernization with “the growing tendency to regard the interests of the individual as a more significant test of the quality of a social system than any collective interest” (p. 24); and McNicoll (1978; cited in Ryder, 1983:24) described how “the structure of society that formerly held people in suspension receded, leaving individuals to struggle in a harsh neoclassical world.” Lesthaeghe (1980) touched on a similar theme in his discussion of “the interplay between individual calculation and institutional control,” noting that “problems of modernization deal in one way or another with the juxtaposition of individuals seeking more room for individual choice in response to new conditions, and the inheritance of the past operating on the level of the moral code and traditional forms of control” (pp. 534–535).

Other Considerations

Even if the institutional-determinants literature gave no recognition to the role of individual variability, there would still be several reasons for researchers to concern themselves with individual-level data collection and analysis and the link between macro variation in social institutions and micro variation in fertility behavior.

First, the dominant contemporary research paradigm is micro-analytic if not explicitly micro-economic. This is Easterlin's paradigm (1978; Easterlin and Crimmins, 1985). Its institutionalization is manifest in the National Research Council's framework for the study of fertility determinants (Bulatao and Lee, 1983), which avers: "The framework is developed at the level of the couple or household since it is assumed that they are the ultimate decision makers about fertility. If the community or society desire higher or lower fertility, they must somehow influence couples or households" (p. 2). Current interest in the institutional determinants of fertility is in large part a reflection of the perceived limitations of conventional theories and research strategies. It is therefore incumbent on this new school of research to reconcile findings with existing knowledge—much of which derives from the individualistic socioeconomic or micro-economic tradition. Only through the observation of individual-level data can it be determined whether existing fertility differentials reflect differences in desired family size (demand for children) as well as variation in the proximate determinants of fertility (supply of children).

Second, there is the issue of competing hypotheses. The selective review of the literature has been agnostic with respect to the validity and utility of various theories and empirical propositions and has focused instead only on the proper level of their operationalization. This has partially obscured the considerable debate surrounding some of these theories, including the presence of what are functionally alternative hypotheses. A good example is the role of education in fertility decline. The earlier discussion of Caldwell's (1982) theory of mass education focused on its irreducible, systemic properties, including an absence of concern for the relative duration of education across individuals. This contravenes an extensive and varied research tradition: Cochrane's (1979, 1983) reviews of the effects of education on fertility stressed the nonlinear association between the two variables at the micro level, with increases in education at low levels leading to increases in fertility; and the effect of education on the individual is prominent in discussions of fertility from a variety of perspectives, including economics (Willis, 1973), anthropology (LeVine, 1980), and public policy (Gille, 1985). Thus any test of Caldwell's (1982) mass education hypothesis at the community level should simultaneously test micro-level alternatives (Singh and Casterline, 1985:220).

Third, extensive observations on individuals are a prerequisite for the computation of so-called "X-bar" variables—areal or contextual averages of a given characteristic, designed to measure the emergent properties of aggregated individual-level characteristics (Blalock, 1984). For example, Casterline (1981) calculated community-specific levels of contraceptive use and used them in a series of individual-level analyses as proxies for community norms regarding contraceptive use (cf. Goldberg, 1976). The extent to which these methods can be applied to research into the institutional determinants of fertility hinges on further specification of the nature of hypothesized macro effects, but insofar as emergent properties are specified by at least some theories—for example, the mass education hypothesis—attention to micro data is again warranted.

Research Design and Strategies for Analysis

If investigators accept the proposition that fundamental properties of key theories are located conceptually at the macro level but agree that it is desirable to link these properties

to fertility behavior at the micro level, then some integrative, multilevel research design is clearly called for.

Multilevel analysis has increased in visibility in recent years, owing in particular to its direct application to the comparative analysis of WFS data. Hermalin and Mason (1980) provided a detailed introduction to the multilevel model and used it to investigate variation across 15 populations in the individual-level relationship between education and fertility. The forthcoming Demographic and Health Survey and the exhumation of various KAP (knowledge, attitudes, and practice) surveys (Hermalin, Entwisle, and Myers, 1985) give some promise that future multilevel analyses may contain an intertemporal comparative reference. This speaks to one criticism of the use of these large-scale surveys for testing theories of fertility—the absence of data on fertility *change* (e.g., Caldwell, 1982, 1985).

Nonetheless, comparative analysis of these national-level data sources is not the direction in which researchers into the institutional determinants of fertility should head—at least not yet. The multilevel analytic framework is very useful, but the national-level data sets to which it has been applied do not constitute optimal or even appropriate research designs. The units of analysis at the macro level are too large to sustain the assumption of homogeneity with respect to the institutional features deemed crucial in imparting variance in fertility and fertility-related behaviors. Instead, future research into the institutional determinants of fertility should hew toward the comparative analysis of selected communities.

The Comparative Analysis of Communities

What are the broad outlines of research designs for the comparative analysis of communities? Potter (1983) located his recommendation within the tradition of nationally representative data collection via sample surveys:

The use of heavily clustered samples in future surveys would permit a two-pronged investigative effort, collecting detailed community-level data on social and economic organization on the one hand, and individual and family-level data on fertility, economic activity, and the like, on the other. (p. 657)

Mason (1984) suggested “the intensive ‘controlled comparison’ study . . . a study that focuses on a relatively few communities or aggregates that share a number of factors in common but vary in key respects with regard to gender inequality” (pp. 77–78). Caldwell (1985) described a strategy for the intensive investigation of several community settings (anywhere from five to a dozen), including both rural villages and urban neighborhoods. Research begins with a census/survey, continues with extensive quasi-anthropological participant observation, and concludes with a second survey.

Justification for the comparative community approach begins with the weaknesses of alternative research designs. First, intensive studies of *single* villages yield both good empirical data and new working hypotheses, but they are difficult to generalize (Kish, 1987; Potter, 1983). It is difficult to gauge the context dependence of both data and theories. Caldwell (1985) acknowledged the limitations of single-village studies, noting that in India his “early work began by assuming a too purely anthropological approach, with increasing concentration on a single village and with nearly all effort expended on gaining knowledge of individual families to the point of saturation” (p. 52).

Second, the traditional approach to the study of institutional determinants of fertility—qualitative, discursive theoretical arguments, illustrated with selective, eclectic data drawn from around the globe—tells little about what is happening in particular countries. This approach also features theories that are essentially not falsifiable.

Third, purely individual-level surveys (e.g., those of the WFS) reveal little or nothing about the institutional environment surrounding the household. In contrast, “research concentrating on local institutional settings would supply much of the ‘mediation’ needed

between global and individual levels of analysis, whatever the theoretical framework adopted by the analyst" (Potter, 1983:659).

Fourth, the small-scale nature of a community study makes "the measurement of complex and elusive phenomena . . . more feasible than in a large-scale study" (Mason, 1984:78), because the structural features of culture, economy, and society are neither well apprehended nor well operationalized at the level of the nation-state. Better measurement of macro characteristics is possible at the community level. Comparative community studies also permit the measurement of variables such as levels of nutrition and contraceptive prevalence—micro-level variables that are often not available (or if available, less meaningful) for larger aggregates.

How should communities be selected for a comparative study?

With regard to sample design, there [are] important questions about the number, size, and character of the "communities" to be included . . . there [is] even . . . the important question of whether the primary sampling units should be defined geographically or according to some other criteria. (Potter, 1983:658)

Politically and/or geographically delimited jurisdictions may or may not represent communities in the sense that is crucial to the theories reviewed above—social systems sharing integral forms of structure, culture, and norms.

There is a tension between the need for homogeneity within communities and the empirical issue of whether institutions are all-encompassing with respect to geographic entities such as villages (McNicoll, 1975). A presupposition of the community-level design strategy is that institutional aspects of social stratification be coincident with geographic segregation. It would be disconcerting to discover that a selected "community" had no unitary marriage customs, rules for inheritance, and forms of kinship and residence; it would also be an important datum. The extent of heterogeneity *within* communities can be a source of variation *between* communities, as, for example, in the distinction between rural and urban opportunities and ways of life. Although researchers should seek communities that represent systemic and institutional features deemed important in relation to fertility, they must also be aware of the extent to which actual communities depart from ideal types. It is possible that in some circumstances geographically defined communities will provide less adequate frames for sampling social institutions than in others.

Assuming that some *modus operandi* can be reached by which communities are defined, the focus shifts to the issue of which communities are to be studied. Both Caldwell (1985) and Mason (1984) advocate purposive selection procedures, but toward different ends. "[T]he populations chosen for the micro studies [must] vary in type and represent as far as possible significantly different ways of life" (Caldwell, 1985:60)—a strategy consonant with the dictum that primary sampling units be as heterogeneous as possible. Mason's (1984) recommendation of a "controlled comparison" mirrors the basic principle of quasi-experimental design—that objects under study be as alike as possible, differing only with respect to a single "test" variable:

[C]ontrolled comparisons offer the advantage of permitting the researcher to examine the impact of gender inequality on fertility or mortality without have [sic] to deal with extraneous "noise" created by covariates. Because the units involved are chosen deliberately so as to minimize other sources of variation in fertility and mortality, attention can be focused on the source of variation of interest, namely, aspects of female status. (p. 70)

The distinction between these formulations has a basis in two desiderata of research design. The desirability of "wide spreads across important disturbing variables" allows for the "wider confirmation" of results (Kish, 1987:52); and in this sense results can be viewed as more representative, even if sites (communities) are not selected by a probability method. In contrast, the controlled comparison approach reflects the desirability of strong controls on known disturbing variables.

If communities “vary in type . . . as far as possible,” but the number of communities under study is not large, then it may be difficult to ascribe differences in fertility or its proximate determinants to, say, aspects of gender inequality, as opposed to ethnicity, religion, organization of agriculture, or the presence of health care clinics. On the other hand, strict adherence to the controlled comparison approach means that if important interactions exist among institutional characteristics with respect to their effects on fertility, they may not be observable within the framework of the design. As a result, site selection of communities with respect to key institutional features should adhere to basic principles in the design of experiments. If two institutional features are of primary importance—say, Muslim (A) versus non-Muslim (a) and urban (B) versus rural (b)—then a fully crossed four-site design may be appropriate (AB–aB–Ab–ab). If a third institutional feature is also of key interest, but expansion of the number of sites is impracticable, then the choice of sites can be balanced as well with respect to this third variable via the appropriate Latin square design (ABC–aBc–Abc–abC; Kish, 1987:52). In all instances variation should be maximized with respect to the exogenous institutional variables; choosing sites on the basis of fertility differentials (sampling on the dependent variable) is not recommended.

How are comparative community studies to be executed? What methods are to be employed? Mason (1984:78) suggests studying each community “intensively using a variety of data collection techniques, including surveys and participant observation.” Potter’s (1983) discussion of methods leans toward reliance on surveys, except for the measurement of institutional characteristics of communities, which would be accomplished via a combination of direct (external) observation, interviews with authorities, and relevant forms of content analysis (e.g., examination of educational curricula).

Caldwell (1985) stresses the quasi-anthropological approach and discourses at length on the limitations of surveys in particular and quantitative methods more generally (Caldwell, 1982:285–286). He does, however, recommend the use of surveys for generating baseline data (p. 52), and his inventory of essential survey items would be a good place to begin any future discussions of minimal data requirements in this area of research. Although all questions pertain to couples and/or households, the sampling design underlying such a survey must be attuned to families/kin groups as additional micro units of analysis.

Researchers involved in comparative community-level studies of the institutional determinants of fertility must also come to terms with the increased criticism of survey research in particular and quantitative methods in general. This critique is more often spoken than written, but it finds sufficient voice in Caldwell’s (1982, 1985) work for its particulars to be noted. Some of these criticisms—for example, that many surveys, by dint of their large sample sizes, yield “significant” results that lack substantive import (Caldwell, 1982:286)—are complaints against sloppy research practices or the paucity of the investigator’s imagination. They are not fundamental critiques of either surveys or quantitative methods per se. Other criticisms—for example, “surveys do not generate their own questions” (Caldwell, 1985:46)—can be alternatively read as saying that surveys are poor instruments for exploratory analysis and the formulation of hypotheses or as a general attack on positivism.

The truly troubling criticisms, however, pertain to the disjuncture between what survey questions are supposed to be measuring and what (if anything) they are actually picking up. Complex issues are tapped by single questions. Forced answers to questions do not permit reasoned responses. Questions have different meanings in different contexts. Answers to even the most basic questions on “objective” socioeconomic and demographic characteristics are passed through dense interpretive filters.

There is, however, reason to be sanguine that survey methods can be fruitfully applied to research designs such as the comparison of communities. It is worth recalling that one of the crucial advantages of the survey research methodology, especially as applied to comparative research, is *standardization*. In a well-executed survey, subjects receive the same

measurement stimulus both within and across contexts (communities). If there is reason to believe that the "meaning" of an item varies across contexts, then this should be (and can be) demonstrated empirically. This will take more time, effort, and attention to detail than has hitherto been characteristic of the formulation of survey research instruments in demography as well as other fields. It is, however, worth doing. In the absence of some general rules regarding what makes a (subjective) survey item relevant or meaningful, the adjudication of the adequacy of items becomes a function of the relative cleverness and moral authority of the various researchers involved.

The issue of the validity of measurement instruments in the study of fertility—an old one (e.g., Cicourel, 1974)—takes on increased importance in the context of comparative research. At the macro level, there is the question of how best to capture or represent systemic properties purported to be invariant within community settings. At the micro level, issues include how to make questions mean to the respondent what they mean to the investigator and how to ensure that questions have the same meaning across contexts. Shadowing the growth of interest in the institutional determinants of fertility is the growing dissatisfaction with the measurement limitations of traditional (survey-based) practices of data collection and analysis. For the gap between macro and micro effects to be closed, more attention must be given to the fundamentals of measurement and less to blanket condemnations of particular methodologies.

There is one limitation of the community comparisons approach that obtains regardless of methodology—the absence of a temporal perspective. Although the availability of several contexts may permit the construction of a synthetic dynamic perspective, problems remain. If a given community manifests substantial fertility variation according to some individual-level covariate—for example, education—it may represent the powerful properties of this variable in inducing heterogeneity in a small population. Alternatively, a longer perspective might reveal this association to be spurious—a transitory state of affairs glimpsed in the course of a much larger process of institutional and demographic change. Although it is wishful thinking at this stage to propose that comparative community studies incorporate a longitudinal design, it is reasonable to suggest that research protocols be made sufficiently explicit to permit genuine replication studies at some point in the future.

Strategies for Analysis

Until protocols for a comparative community study are established, it is of little value to engage in detailed prescriptions for data analysis. In fact, it is likely that analyses will take on many forms, depending on both level of theory and quality of data. Cain's (1981) study of risk environments in four villages (three Indian and one Bangladeshi) is exemplary with respect to the use of both quantitative and descriptive observational data in comparative perspective. It is a model for similar analyses, although other research questions will require other exemplars.

This is especially true in the realm of fertility and its proximate determinants: The following discussion is a brief introduction to the logic of the analysis of micro data that might be obtained from a well-designed comparative community study (cf. Hermalin and Mason, 1980:92–99). Because the focus is on the logic of analysis, the presentation is fairly abstract; basic identities are depicted via linear relationships, but no particular stance is adopted with respect to measurement of the dependent variable, functional form, or means of estimation (cf. Entwisle and Mason, 1985; Mason, Wong, and Entwisle, 1983; United Nations, 1985; Wong and Mason, 1985). Although the presentation makes use of a single dependent quantity (Y) and single individual-level regressors (I) and community variables (C), the criterion variable Y can vary across analyses, and both I and C can be thought of as representing sets of, respectively, micro- and macro-level variables.

The level of the response variable (Y) for the i th individual in the k th community can be expressed as

$$Y_{ik} = a_0 + b_k + g_k I_{ik} + e_{ik}, \quad (1)$$

where a_0 is a constant, for example, some minimal level of fertility, or the average fertility level across all communities; b_k is the “effect” of the k th context and is constant across individuals (or families) within that community; g_k is the coefficient, also constant for all individuals within the k th community, measuring the effect of the individual-specific variable(s) I_{ik} on Y ; and e_{ik} is an error term, representing individual variation in Y not captured elsewhere.

Expression (1) is potentially interactive, since if for some community contexts [$k = 1, \dots, K$ (the total number of communities)] $g_k \neq g_{k'}$, then the effect of the individual regressors(s) I on Y can be said to vary by context. If such an interaction is present, then the primary research focus pertains to the specification of $g_k = f(C_k)$; that is, what macro aspects of the community condition the relationship between I and Y ? *The problem of understanding cross-contextual variation in relationships measured at the individual level has so far been paramount in multilevel fertility research* (e.g., Hermalin and Mason, 1980:111–141).

If, however, $g_k = g_{k'}$, for all k , then equation (1) can be re-expressed as

$$Y_{ik} = a_0 + b_k + g I_{ik} + e_{ik}, \quad (2)$$

which is identical to expression (1) save for the deletion of the subscript on the g coefficient, thereby indicating that individual-level effects *do not* vary by community. This means that the effects of micro and macro regressors are *additive* rather than interactive.

Analysis of equation (2) is abetted by the introduction of

$$Y_{ik} = a_0^* + b_k^* + e_{ik}^*, \quad (3)$$

in which individual-level effects have been omitted. Variation in Y is thus a function of community-level variation alone (the b_k^* s); and equation (3) is the familiar expression for the analysis of variance. If $b_k^* \neq b_{k'}^*$, for some k and k' , then there is community-level variation in Y , and attention should focus on the b_k of equation (2). If they are all equivalent ($b_k = b_{k'}$, for all k), then there is no net effect of context on Y . Instead, variation in Y is a function of variation of the average of I , the individual-level regressor (or set of regressors). This does not mean that community differences in macro phenomena are unimportant; rather, it means that the entire gross effect of community (b_k^*) is interpreted through the effects of community (C) on the distribution of individual-level characteristics (I). The analyst's focus might thus shift to the specification of $I_k = f(C_k)$, the effect of community differences on the distribution of individual attributes across communities. *This situation—complete explanation of community-level phenomena via their effects on the conditional distributions of individual-level characteristics associated with fertility—is consistent with the socioeconomic perspective regarding the determinants of fertility and fertility decline.*

If for all k , $b_k \neq b_{k'}$, then there are direct community effects on Y , net of the inclusion of the individual-level regressors (I). Variation in community effects may be studied via the specification of $b_k = f(C_k)$, that is, community-level variables that account for the observed pattern of context effects (the estimated b_k). If in comparisons across expressions (2) and (3), $b_k = b_k^*$, then net and gross community effects are equivalent, and community effects are functionally independent of simultaneously included individual-level regressors. *This is the implicit prediction of those segments of the institutional determinants of fertility literature that ignore or downplay the role of individual-level factors in the determination of fertility and fertility-related behaviors.*

Thus there are three avenues through which community-level variables can manifest their effects on the fertility behavior of individuals. The three are not mutually exclusive, and it may be that community-level variation in individual-level fertility can be attributed to all three sources. Still, the logical exercise of locating hypothetical community effects within this framework is of considerable value. Theories that inhere at the community (macro) level, but cannot be linked to individual fertility behavior via this or any comparable framework, are likely to be of little value as guides to further empirical research.

Summary

To date, the literature on the institutional determinants of fertility has featured theories and research hypotheses residing primarily at the aggregate level. Research based on the operationalization of these hypotheses has been tentative and scant, in part owing to the newness of this line of inquiry, but also to the misfit between theoretical and operational levels of analysis. This article has attempted to (a) draw out the *macro* nature of salient conceptual discussions of the institutional determinants of fertility, (b) argue the need to judge theories of fertility with respect to their ability to explain fertility variation as observed at the *individual* level, and (c) outline a general approach by which relevant macro causes and micro outcomes might be integrated and investigated empirically. The research design behind this approach is a comparative community study; the framework for the analysis of data is the multilevel model. The foregoing is less an outline for further work in this area than it is a checklist—a reminder of the conceptual elements that must be considered in the design of empirical research in this area.

Acknowledgments

This article is a revised and abridged version of the paper "Theory and Research on the Status of Women and Fertility: Some Necessary Linkages," presented at the Rockefeller Foundation Workshop on Women's Status and Fertility, Mt. Kisco, New York, July 1986. Funding for the preparation of that paper was provided by the Population Sciences Division of the Rockefeller Foundation. From that paper to this, I have benefitted from the comments and encouragement of Tom Gierny, Susan Watkins, Pat Caldwell, Mary Kritz, Parker Mauldin, and Phil Morgan plus the editors of this journal, although none bear responsibility for the views expressed herein.

References

- Blalock, H. M. 1984. Contextual-effects models: Theoretical and methodological issues. *Annual Review of Sociology* 10:353–372.
- Bulatao, R. A., and R. D. Lee (eds.). 1983. *Determinants of Fertility in Developing Countries* (2 vols.). New York: Academic Press.
- Cain, M. 1980. Risk, fertility, and family planning in a Bangladesh village. *Studies in Family Planning* 11:219–223.
- . 1981. Risk and insurance: Perspectives on fertility and agrarian change in India and Bangladesh. *Population and Development Review* 7:435–474.
- Cain, M., S. R. Khanam, and S. Nahar. 1979. Class, patriarchy, and women's work in Bangladesh. *Population and Development Review* 5:405–438.
- Caldwell, J. C. 1982. *Theory of Fertility Decline*. New York: Academic Press.
- . 1985. Strengths and limitations of the survey approach for measuring and understanding fertility change. Pp. 45–63 in J. Cleland and J. Hobcraft (eds.), *Reproductive Change in Developing Countries*. New York: Oxford University Press.
- Casterline, J. B. 1981. *Community Effects on Fertility Intentions: The Effect of Aggregate Levels of Contraceptive Use on Individual Intentions to Use*, WFS Technical Paper No. 1703. London: World Fertility Survey.
- Cicourel, A. V. 1974. *Theory and Method in a Study of Argentine Fertility*. New York: Wiley.
- Cleland, J. 1985. Marital fertility decline in developing countries: Theories and the evidence. Pp. 223–252 in J. Cleland and J. Hobcraft (eds.), *Reproductive Change in Developing Countries: Insights From the World Fertility Survey*. New York: Oxford University Press.

- Cleland, J., and J. Hobcraft (eds.). 1985. *Reproductive Change in Developing Countries: Insights From the World Fertility Survey*. New York: Oxford University Press.
- Coale, A. J. 1973. The demographic transition. Pp. 53–72 in *International Population Conference, Liege, 1973*. Liege, Belgium: International Union for the Scientific Study of Population.
- Cochrane, S. H. 1979. *Fertility and Education: What Do We Really Know?* Baltimore, Md.: Johns Hopkins University Press.
- . 1983. Effects of education and urbanization on fertility. Pp. 587–626 in R. A. Bulatao and R. D. Lee (eds.), *Determinants of Fertility in Developing Countries* (Vol. 2). New York: Academic Press.
- Cutright, P. 1983. The ingredients of recent fertility declines in developing countries. *International Family Planning Perspectives* 9:101–109.
- Davis, K. 1963. The theory of change and response in modern demographic history. *Population Index* 29:345–366.
- Dyson, T., and M. Moore. 1983. On kinship structure, female autonomy, and demographic behavior in India. *Population and Development Review* 9:35–60.
- Easterlin, R. A. 1978. The economics and sociology of fertility: A synthesis. Pp. 57–113 in C. Tilly (ed.), *Historical Studies of Changing Fertility*. Princeton, N.J.: Princeton University Press.
- Easterlin, R. A., and E. M. Crimmins. 1985. *The Fertility Revolution: A Supply–Demand Analysis*. Chicago: University of Chicago Press.
- Entwisle, B., and W. M. Mason. 1985. Multilevel effects of socioeconomic development and family planning program on children ever born. *American Journal of Sociology* 91:616–649.
- Gille, H. 1985. Policy implications. Pp. 273–295 in J. Cleland and J. Hobcraft (eds.), *Reproductive Change in Developing Countries: Insights From the World Fertility Survey*. New York: Oxford University Press.
- Goldberg, D. 1976. Residential location and fertility. Pp. 387–428 in R. G. Ridker (ed.), *Population and Development: The Search for Selective Interventions*. Baltimore, Md.: Johns Hopkins University Press.
- Hermalin, A. I., B. Entwisle, and L. G. Myers. 1985. Some lessons from the attempt to retrieve early KAP and fertility surveys. *Population Index* 51:194–208.
- Hermalin, A. I., and W. M. Mason. 1980. A strategy for the comparative analysis of WFS data, with illustrative examples. Pp. 90–168 in *The United Nations Programme for Comparative Analysis of World Fertility Survey Data*. New York: United Nations Fund for Population Activities.
- Kish, L. 1987. *Statistical Design for Research*. New York: Wiley.
- Lesthaeghe, R. 1980. On the social control of human reproduction. *Population and Development Review* 6:527–548.
- LeVine, R. A. 1980. Influence of women's schooling on maternal behavior in the Third World. *Comparative Education Review* 24:S78–S104.
- Mason, K. O. 1984. *The Status of Women: A Review of Its Interrelationships to Fertility and Mortality*. New York: Rockefeller Foundation.
- . 1986. The status of women: Conceptual and methodological issues in demographic studies. *Sociological Forum* 1:284–300.
- Mason, W. M., G. Y. Wong, and B. Entwisle. 1983. Contextual analysis through the multilevel linear model. Pp. 72–103 in S. Leinhardt (ed.), *Sociological Methodology* 1983–84. San Francisco: Jossey-Bass.
- Mauldin, W. P., and B. Berelson. 1978. Conditions of fertility decline in developing countries, 1965–75. *Studies in Family Planning* 9:90–147.
- McNicoll, G. 1975. Community-level population policy: An exploration. *Population and Development Review* 1:1–21.
- . 1978. Population and development: Outlines for a structuralist approach. Pp. 79–99 in G. Hawthorn (ed.), *Population and Development*. London: Frank Cass.
- . 1980. Institutional determinants of fertility change. *Population and Development Review* 6:441–462.
- Notestein, F. W. 1953. Economic problems of population change. Pp. 13–31 in *Proceedings of the Eighth International Conference of Agricultural Economists*. New York: Oxford University Press.
- Potter, J. E. 1983. Effects of societal and community institutions on fertility. Pp. 627–665 in R. A. Bulatao and R. D. Lee (eds.), *Determinants of Fertility in Developing Countries* (Vol. 2). New York: Academic Press.
- Ryder, N. B. 1983. Fertility and family structure. Pp. 15–34 in *Population Bulletin of the United Nations*. New York: United Nations Dept. of International Economic and Social Affairs.
- Singh, S., and J. Casterline. 1985. The socio-economic determinants of fertility. Pp. 199–222 in J. Cleland and J. Hobcraft (eds.), *Reproductive Change in Developing Countries: Insights From the World Fertility Survey*. New York: Oxford University Press.
- Teitelbaum, M. S. 1975. Relevance of demographic transition theory for developing countries. *Science* 180:420–425.
- Thompson, W. 1929. Population. *American Journal of Sociology* 34:959–975.
- Tilly, C. 1984. *Big Structures, Large Processes, Huge Comparisons*. New York: Russell Sage Foundation.
- United Nations. 1985. *Women's Employment and Fertility: A Comparative Analysis of World Fertility Survey*

- Results From 38 Developing Countries*, Population Studies No. 96. New York: United Nations Dept. of International Economic and Social Affairs.
- Willis, R. 1973. A new approach to the economic theory of fertility behavior. *Journal of Political Economy* 81:S14–S64.
- Wong, G. Y., and W. M. Mason. 1985. The hierarchical logistic regression model for multilevel analysis. *Journal of the American Statistical Association* 80:513–524.
- Wrigley, E. A. 1978. Fertility strategy for the individual and the group. Pp. 135–154 in C. Tilly (ed.), *Historical Studies of Changing Fertility*. Princeton, N.J.: Princeton University Press.