

ECONOMIC GROWTH CENTER

YALE UNIVERSITY

P.O. Box 1987, Yale Station
27 Hillhouse Avenue
New Haven, Connecticut 06520

CENTER DISCUSSION PAPER NO. 460

STUDYING THE IMPACT OF HOUSEHOLD ECONOMIC AND COMMUNITY VARIABLES
ON CHILD MORTALITY

T. Paul Schultz

July, 1984

Notes: Prepared for a Workshop on Strategies for Research on Child
October 11-15, 1983, Rockefeller Foundation Studies Center,
Bellagio, Italy. This research was supported in part by grants
from NIH, Center for Population Research HD-12172 and from the
General Service Foundation. I have benefited in revising this
paper from the comments of the conference participants, David
Feeny and Ethel Churchill.

Center Discussion Papers are preliminary materials circulated
to stimulate discussion and critical comment. References in
publications to Discussion Papers should be cleared with the
author to protect the tentative character of these papers.

Abstract

Empirical study of the biological and behavioral determinants of child survival may proceed in two directions: either to estimate the reduced form relationships between child survival and variables that are assumed to be exogenous to the households' many choices; or to estimate both the behavioral demand equations for health inputs and the biological production function linking health inputs to child survival. The reduced-form equation for child survival may be estimated without imposing a great deal of structure on the problem, whereas the biological production function for child survival can only be estimated after more structure is imposed, because in the presence of health heterogeneity, self-selected health inputs will be correlated with the child's unobserved biological health endowment. To eliminate the bias due to self-selected inputs and heterogeneity, a two-stage estimation procedure is outlined that estimates first the health input demand equations, and then estimates consistently the health production function by simultaneous equations methods. In the estimation of either the reduced-form equation or the health input and production relationships, it is valuable to have data collected both for a sample of individuals of households and on prices, programs, and characteristics of the communities in which these households are located. Community aggregate variables on prices, wages, programs, and environmental conditions are essential for the analysis, because they are outside the control of the household, and hence, they are legitimately treated as exogenous instruments.

Empirical studies of child survival make use of many types of information and are guided by many research paradigms. Statistical analyses of the determinants of child survival may deal with observations on regional conditions and population aggregates, characteristics of systematically selected populations such as those encountered in a clinical practice, or data from representative household surveys on children and families.

Epidemiological research often measures the direct association between inputs to health, such as nutrition and medical care, and health outcomes, such as child morbidity, disability, and mortality. Economic research, on the other hand, generally evaluates how the constraints on peoples' opportunities are associated with child health, presumably by affecting behaviorally some of the direct inputs. The purpose of this paper is to show why the former direct association of health inputs and an individual's health is a statistically biased and potentially misleading indicator of the corresponding causal effect. This bias arises because people are essentially different in their health endowments, and their use of health inputs tends to be affected by what they know about their endowment or heterogeneity. A stochastic framework is proposed for studying jointly the biological determinants of health in the presence of such heterogeneity, and the economically constrained selection of health inputs. Within this integrated framework the requirements for merging several types of data and for maintaining various "structural" assumptions will become clear. The goal is to estimate without bias both the underlying biological and behavioral relationships that will allow one to assess the benefits of "technology based health interventions" as well as the consequences of "socioeconomic change" that affects the prospects for child survival by influencing peoples' opportunities and thereby their use of health related

inputs.

Information used in the study of child health is classified in Figure 1 according to whether the variable is viewed here as independent/exogenous or dependent/endogenous, whether a variable is unobserved by the researcher, and the appropriate level of aggregation, namely the household or community. There are two basic sources of data: (1) information from surveys or census records on a representative sample of households and individuals and (2) characteristics of regions where these individuals live, of markets in which they obtain and exchange goods and services, and of the physical environment that constrains their health and productive opportunities. To implement a research design based on these data, it is necessary first to lay out the conceptual framework guiding the research. It is important then to specify the structure and source of statistical disturbances and unobserved factors that enter the central relationships, for these features of the problem suggest how to estimate without bias the parameters in some of these relationships.

General Framework for Estimating the Determinants of Child Health

The relationship between child mortality or morbidity for the i th mother, Y_i , is assumed to be approximately a linear function of a vector of proximate biological inputs to child health, I_i , and a vector of persistent biological endowments of the child, B_i , and a random disturbance, e_{1i} , and is hereafter called a health "production function" (see Figure 2):

$$Y_i = c_0 + c_1 I_i + c_2 B_i + e_{1i}, \quad (1)$$

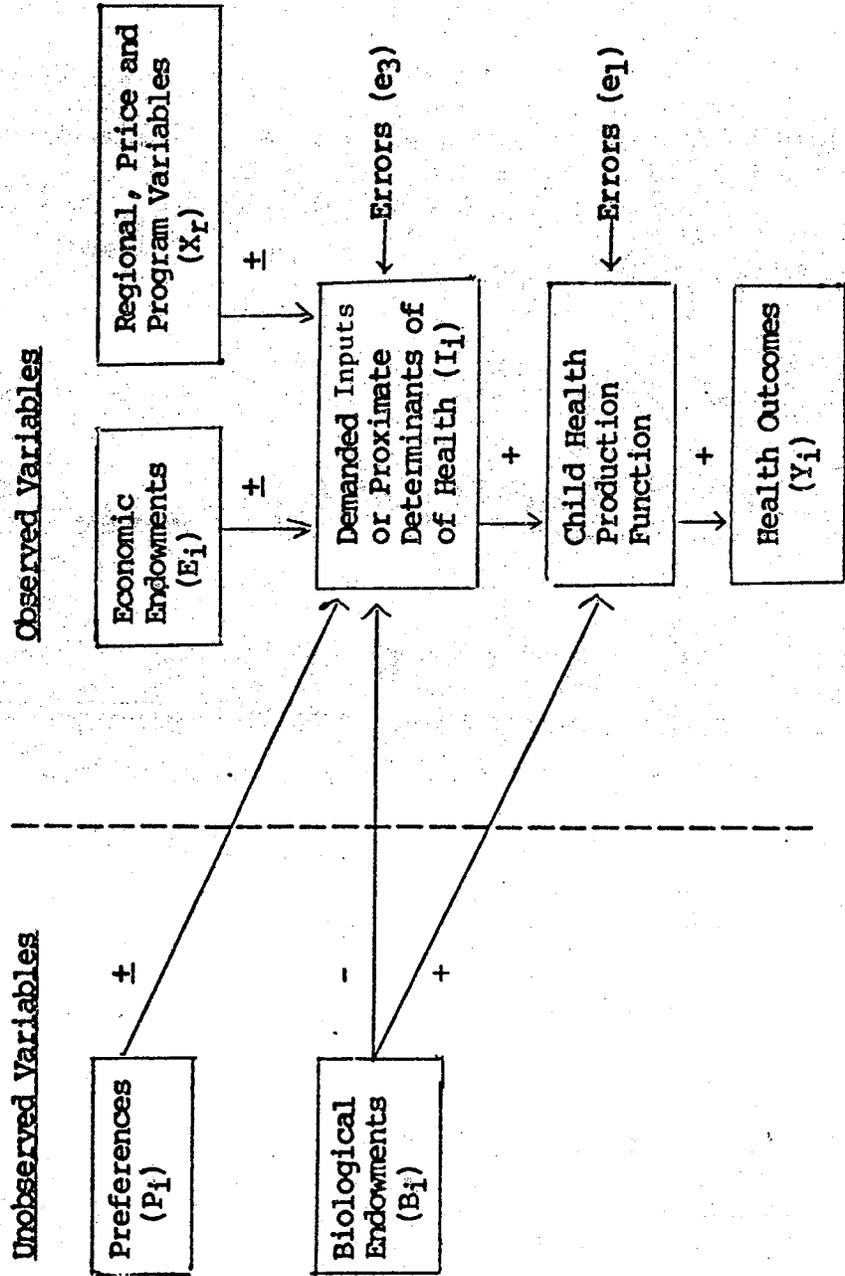
where the c 's are the parameters of this linearized biological/technical

Figure 1

Parallel Classifications of Variables
Used in the Study of Child Mortality Determinants

- I. Independent or Exogenous Variables
 - a. Preferences or Individual/Family Goals (unobserved by researcher) (Pi)
 - b. Individual Endowment Variables (Xi)
 - Economic (Ei)
 - Education-Labor
 - Nonhuman Assets
 - Adult Lifetime Wealth
 - Biological (unobserved by researcher and not controlled by individual) (Bi)
 - Healthiness or Frailty
 - Fecundity
 - c. Community or Regional Variables (Xr)
 - Market exchange prices and wage rates
 - Public programs—access, money and time costs to use
 - Availability of information
 - Infrastructure for production, e.g. roads
 - Climate and exposure to disease
- II. Dependent or Endogenous Variables
 - a. Proximate or Intermediate Input Variables (Ii) related to the technical determination of child health, e.g.
 - Nutrition, breastfeeding, and consumption patterns
 - Hygiene and child care
 - Water and sanitation environment
 - Preventive and curative medical care
 - Fertility and child spacing
 - b. Child Health Outcomes or Output (Yi)
 - Child survival
 - Child acute and chronic morbidity: incidence, severity, and duration of episode

Figure 2
Flow Diagram of Factors Determining Child Health



relationship.

The I_i are chosen by the woman and her family to minimize Y_i , and to achieve other goals with her limited resources; B_i is again the child's health endowment, that component of child health due to either genetic or environmental conditions which cannot be influenced by the family's behavior but which is partially known to it, called health heterogeneity; the economic endowments of the individual including human and nonhuman capital, E_i ; the regional prices, programs and environmental constraints, X_r ; preferences, P_i , of the woman and her family; and e_{2i} , a stochastic disturbance that is independent of the exogenous variables in Figure 1:

$$I_i = a_0 + a_1E_i + a_2B_i + a_3X_r + a_4P_i + e_{2i}. \quad (2)$$

This relationship between the use of health inputs and all relevant exogenous variables is referred to as a reduced-form input demand function. The parameters, a 's, reflect behavioral responsiveness to socioeconomic and biological constraints, as well as the parents' understanding of the health production function.

Finally, there implicitly exists reduced-form equations for the health outcomes that have the same arguments as equation (2):

$$Y_i = b_0 + b_1E_i + b_2B_i + b_3X_r + b_4P_i + e_{3i}, \quad (3)$$

where the parameters, b 's, represent the combined effects of technological constraints and demand behavior, for example in this simple linear model, $b_1 = a_1c_1$, and $b_3 = a_3c_1$.

The parameters of these reduced-form equations (2) and (3) can be estimated

consistently by standard statistical methods, such as ordinary least squares, if the dependent variable is continuous and normally distributed, because the inability of the researcher to observe preference (P_i) or biological endowments (B_i) need not bias the remaining estimates. This follows from the conventional assumption that tastes and biological endowments are distributed independently of economic endowments, prices, programs, etc., namely, P_i and B_i are uncorrelated with E_i and X_r .

The estimation problem is to obtain consistent estimates of the health production function (1). Because the unobserved biological variability in a child's health endowment, say its frailty, may both increase the likelihood of morbidity ($c_2 > 0$) and increase the use of beneficial health inputs ($a_2c_1 < 0$), one expects compensating responses in observed health inputs for low (unobserved) health endowments. Direct estimation of the health production function by standard epidemiological methods is unsuitable in this case because input use, I_i , and health heterogeneity, B_i are likely to be related. The estimation problem posed by heterogeneity in health endowments across a population is likely to be more serious in the evaluation of the effect of curative health inputs for acute conditions and less serious in evaluating the benefits of long term preventive health measures. As in the conventional simultaneous equations model, some instrumental variable(s) is required that is independent of the biological endowment (B_i), but which will account for some of the variation in health input demands. The exclusion restriction that is needed to identify statistically the production function parameters might be based on the economic expectation that regional prices and health programs influence health input demands, but do not enter directly as arguments in the production function.^{2/} In this paper, the identification restriction proposed is to

assume that the input selection process (input demand equations) are conditional on regional prices, etc. (X_r), but that regional prices are not in the health production function. The mechanics of this two-stage instrumental variable estimation of the health production function are illustrated in the lower panel of Figure 3. There are other methods for estimating the "structural" production function, such as by using panel data to estimate directly the role of the individually persistent biological endowments, i.e., fixed or serially correlated individual effects. It is preferable to use qualitative structural model estimates if the dependent variable is discrete, such as mortality, or if the health outcome is disproportionately concentrated at particular points, such as zero episodes of illness in the last month (Manski and McFadden, 1981; Olsen and Wolpin, 1983).

It may always be argued that important health inputs to the health production function are omitted from any empirical study and they might be correlated with the included inputs. This situation also leads to biased estimates of the effects of the subset of proximate health determinants that can be observed and jointly analyzed (Griliches, 1957). This omitted input problem may be serious in the case of child survival because the inputs that might reasonably exert an effect on child survival are numerous and difficult to quantify. The most one can do is try to measure and include all of the major determinants, particularly if those determinants covary with the policy interventions being evaluated. Only true randomized experimental design can free one of this complication.

With this introduction to the general framework, the next section outlines how the microeconomic approach to household behavior helps one specify the regional constraints and individual economic endowments. The following section shows how interactions between the mother's education and regional program activity may be used to evaluate distributional implications of policy.

Figure 3

Alternative Strategies for Statistical Study of Merged Household and Regional Variables Defined in Figure 1

Type of Relationship Represented	Linear Approximation of Relationship	Suitable Statistical Technique for Estimating Relationship
----------------------------------	--------------------------------------	--

1. Unconditional or Reduced-Form Demand Functions determining:

- a. Proximate determinant or Input to Child Health (I)

$$I_1 = a_0 + a_1 E_1 + a_2 B_1 + a_3 X_r + a_4 P_1 + a_2 I_1$$
- b. Health Outcomes (Y) with production and demand parameters not separately identified

$$Y_1 = b_0 + b_1 E_1 + b_2 B_1 + b_3 X_r + b_4 P_1 + e_3 I_1$$

Single equation methods, such as ordinary least squares regression or for discrete dependent variables logit or probit functions might be estimated

2. Production Functions determining health outcomes (Y), or a demand function conditional on the input choices (I) that are not systematically correlated with biological endowments (B) or preferences (P)

$$Y_1 = c_0 + c_1 \hat{I}_1 + e_{11}$$

where

$$\hat{I}_1 = \hat{a}_0 + \hat{a}_1 E_1 + \hat{a}_3 X_r$$

and the a's are estimated by standard methods as in la above.

Structural equation methods, such as instrumental variable simultaneous equation estimators

Subsequent sections consider data requirements and report the empirical application of the methodology to the study of infant health determinants in the United States. The final section recapitulates the main points of the paper.

The Microeconomic Model of the Family

Microeconomic analysis of family economic and demographic behavior rests on the hypothesis that people allocate their time and other economic resources in response to the value of the time of each family member, the amount of the family's nonhuman capital endowments, and the relative prices of the family's market inputs and outputs. The marketplace sets wage rates for various types of labor, and thereby determines the value of the time of those persons working in the market or hiring additional labor to supplement their own efforts within the family's productive activities. Despite the diversity in family and market structures and the different degrees of coordination in production, consumption, and reproduction within the family/household in different societies, this greatly simplified economic approach to the interrelated labor market, demographic, and demand behavior of family members has provided a useful framework for much empirical research in high- and low-income countries.

The economic framework suggests that families are likely to respond differently to variation in male and female wages. The wage rate paid to women relative to men is emphasized in family time allocation studies as a relative price for household goods and services that are primarily produced by women. The care of children in many cultures consumes much of the time of their mothers, and, therefore, an increase in the opportunity cost of women's time relative to men's is likely to increase the opportunity cost of children and child care and thus depress fertility, even though it provides the family with increased income opportunities. Consequently, the simplified model of household

economic and demographic demands provides some suggestive predictions for how the labor market is likely to affect the long run evolution of fertility and possibly child mortality.

While good data are not always available on market wage opportunities of men and women or the shadow value of time of household workers, women's education—a close correlate of women's wages—is associated with decreases in fertility and increases in child survival in both urban and rural areas in countries at virtually all levels of development. Numerous studies also confirm the importance of a mother's education for understanding the survival rates of her children even after the effect of husband's income or education is held constant (Caldwell, 1979; Schultz, 1980; Cochrane et al., 1981; Farah and Preston, 1982).^{3/} Moreover, education is a good predictor of the probability of interregional migration for women, as it is for men. Because most migration in low income countries is from rural to urban areas, and rural conditions appear to favor higher child mortality, higher fertility, lower labor force participation by women, and lower school enrollment rates, the increased probability that an educated woman will migrate from a rural to an urban region reinforces the tendencies for her to give her fewer children better health and greater educational opportunities. All of these changes in behavior associated with increasing women's education lead to reinforcing improvements in the "quality" of the future population or the human capital embodied in the next generation (Schultz, 1983). Understanding how a mother's education affects child mortality is a priority issue for research.

Mother's Education

Economics suggests at least five distinct ways that a mother's education might affect child health.^{4/} (1) Education may affect the productivity or

effectiveness of the health inputs used in the production of child health, as hypothesized by Grossman (1972). Thus, more educated mothers may obtain more benefits from a given use of health services; higher "productivity" would also increase their demand for health services, because the mother's resources are thereby effectively increased for allocation to any or all activities. (2) Education may affect perceptions about the best allocation of the health inputs. In this case, more educated mothers would have healthier children because they have better information on the optimal allocation of health resources, and thus are able to produce health at lower cost. Moreover, such mothers would be particularly advantaged where information on "best" input allocations is scarce or costly to acquire, as with new technologies. (3) Education may increase total family resources. Even where women tend not to work in the market, more educated women may have greater market resources because, due to assortative mating, they marry wealthier men. Such women might be expected to invest more resources in the production of child health and thus would have healthier children. (4) More educated women may assign a higher value to their own time, particularly, but not only, if they work in the market and receive a higher wage rate. If mother's time is an essential "input" in the production of child health, education could then be negatively related to the health of children. (5) Education may residually affect preferences for child health and family size, given total resources, prices, and technology.

To ascertain precisely how education affects child health, estimates are needed of the actual technology associated with the production of health. Such estimates require, as noted above, that all behavioral inputs be observed for each child—breastfeeding, nutrient intakes, use of medical services, and so on. In the presence of heterogeneity in health endowments, information would

also be required on prices of all inputs and on the household environment. In the absence of such reasonably complete knowledge of the biological and behavioral determinants of child survival, it is still possible to learn about the effects of policy interventions and to explore the distributional implications of policies by interacting regional policy variables with educational variables in generalized reduced-form health equations (Figure 3, row 1b).

Interactions Between Public Program Interventions and Education

In the jargon of economics, demand equations for health inputs are derived from the maximization of household utility subject to the production functions and the budget constraint. The demands for health inputs are explained by the conventional types of exogenous variables—prices, income, education, age, the local environment, and the unobserved endowments of health. Public programs may play three distinct roles in affecting health, and related behavior, through the demand equations:

1. They may reduce the prices of the health inputs, directly through subsidization of the goods or services or indirectly by increasing access to them, thereby reducing the time or travel costs to evaluate and use the service.

2. They may provide information on how to produce health more efficiently. This might include information on new inputs or on efficient practices with traditional inputs—how long to breastfeed, how to sterilize baby formula, etc.—which yield greater survival rates for a given expenditure.

3. They may alter the health environment, without directly affecting other opportunities available to people. The control of malaria and eradication of smallpox are often cited as examples of exogenously introduced change in health conditions or technology that occurred for all persons living in an area,

regardless of their economic status and educational attainment.^{5/}

The behavioral framework suggests that programs which reduce the costs of health inputs will lead to greater investments in health and thus greater survival rates. An implication of the household framework is that any program which affects the cost of goods consumed, whether directly useful to the production of health or not, may influence the demand for health inputs and thus indirectly affect health or survival. Thus, prices of all major goods and certainly the prices of goods that are thought to be close complements or substitutes for child health belong in the demand equations for child health inputs.

The effects of program interventions may differ depending on the mother's education, according to the relative importance of the (five) roles of education in affecting health and on the predominance of (1) the user subsidy or (2) the productive information or (3) the disease eradication effects of the programs. Table 1 outlines the distributional consequences of these possibilities.^{6/}

Estimates of the effects of education and public health programs on child health are likely to be misspecified, if the interactions between education and the health "infrastructure" are ignored. Moreover, while estimates of such interactions, in the absence of information on the characteristics of the health production function, cannot conclusively pin down the most important roles either of education or of the health programs in augmenting health, they can eliminate some possibilities.^{7/}

Implementation of the General Framework

The general framework proposed here to structure study of child survival is primarily a way to classify relationships and variables, but this taxonomy has important implications for how variables should be measured. A distinction has been made between (1) technical or biological relationships that link inputs to

Table 1
Who Benefits More from Health Program Interventions,
According to the Roles of the Programs and of Mother's Education

Role of Mother's Education	Roles of Health Programs		
	A. Information Raises Techni- cal Efficiency	B. Modern In- put Subsidy Encourages Use	C. Improvement in Healthi- ness of Environ- ment
1. Increase Productivity of Health Inputs	Unknown	Unknown	Less Educated
2. Reduce Costs of Information ON TECHNOLOGY	Less Educated	More Educated	Less Educated
3. Increase Family Income	Unknown	More Educated	Less Educated
4. Increase Mother's Price of Time	Unknown	Less Educated	More Education
5. Affect Preferences for Health	Unknown	Unknown	Unknown

the production of health outcomes and (2) demand relationships that translate individual preferences into the demand for consumption goods and intermediate inputs, conditional on the technological or biological relationships, the prices of inputs and outputs, and the initial factor endowments, education and age of the individual. The production functions contain little information about economic or demographic behavior, per se; they represent physical-technological limitations of the human environment and knowledge of how to use this environment. The demand functions contain the economic core of the model.

Strong assumptions are usually made regarding the optimizing behavior of the individual or the family in choices of economic and demographic behavior, over a budget of given resources, whose exchange values (prices) are fixed in the marketplace, subject to technically fixed production possibilities. Household demand economics embeds the traditional consumer demands for market produced goods into a household firm that also produces the basic commodities that satisfy individual needs. In our case, the household produces the commodity of child survival. The consumer's demands are thus extended to encompass demographic events that are quite different from market goods; because there is no direct market for children or for child health, the consumer must be her or his own producer.

Few researchers have given empirical content to this distinction between production technology and consumer demand. If the assumptions required to disentangle household technology and consumer demands by two-stage estimation procedures prove to be too restrictive, then another more limited approach to studying child survival can be pursued. The socioeconomic and environmental determinants of child survival can be directly estimated unconditional on

production technology. These unconditional demand relationships are called reduced-form demand equations for health (Figure 3, row 1b), and they mix together the underlying structural parameters that characterize the production functions and the conditional input demand functions discussed earlier.

Estimation of these reduced-form demand equations loses track, however, of the interesting linkages across input allocations and the specific details of household production processes. For example, what determines when an individual adopts a child survival input, such as oral rehydration therapy, or how effective it will be when it is adopted? On the other hand, the reduced-form health equations do indicate what many policymakers want to know. For instance, if the local wage rate for women increases in the labor market (i.e., there is a change in this market price), what will be the consequences for child survival? If a hospital bed is added per thousand persons in a locality, (i.e., the price of some forms of medical care is reduced), what will be the consequences for child survival in that region, and also for which groups of parents in that locality will the increase in child survival be greatest? Not only do reduced-form demand equations answer such questions, they can also be estimated without bias by standard multiple regression or multiple logit/hazard techniques. Consistent estimates are, thus, readily obtained at relatively low cost from commonly available data, if one can forego the separation of technological from behavioral parameters.

The controversial aspect of this approach to studying the determinants of household economic and demographic behavior is in deciding precisely what types of behavior are appropriately viewed as demand determined (i.e., endogenous), and what variables are legitimately interpreted as predetermined, such as market prices and individual endowments (Figure 1).

In principle, all forms of behavior that are influenced or constrained by household production and consumption can be treated as partially determined by the equilibrium system of household demand behavior. It is largely an empirical question whether the degree of influence of the demand determinants is substantial or negligible.^{8/} Of course an included explanatory variable may be capturing the effect due to an omitted explanatory variable, or joint and simultaneous effects may be present, clouding any causal insight, or any of a number of the other specification errors could arise if the estimated model were wrong. But conditional on the specification of a model, empirical analysis may proceed to determine whether the factors emphasized by a theory are in fact helpful in explaining variation in behavioral outcomes.^{9/}

Specification and Measurement of Variables

One important insight of household economics is also a limitation to empirical analysis of variation in child survival. Many forms of household economic and demographic behavior are viewed as endogenous to the broader, longer term set of choices that include child health and survival. One endogenous choice cannot generally be used to "explain" another choice or related behavioral outcome.^{10/} First, a line must be drawn at some point in an individual's lifetime, before which events are treated as predetermined or exogenous for the purposes of looking at a specific set of subsequent outcomes. For understanding the determinants of a couple's choices with respect to labor supply, fertility, or child health, one might draw the line when the couple is married. But their timing of marriage could be closely related to the later events under study and probably reflects the heterogeneous unobserved biological characteristics and tastes of persons.

Alternatively, it may be more satisfactory to select one spouse, say the

woman, and treat her characteristics as predetermined when she has completed her childhood and schooling. The argument for this specification is that parents play a dominant role in inducing or preventing their children from attending school, and consequently, the schooling decision reflects heavily the preferences of the family rather than that of the child. It then follows that, given what we know about the individual's characteristics on leaving school, the endogenous choices and behavior that are to be explained start with age at marriage or age at first birth, as well as the characteristics of the husband, if she is married.

Prices

What are appropriate variables to use to explain household demand behavior and outcomes related to child health? The first major class of factors is prices. Prices are predetermined, if they are set by a market over which the individual has negligible control. The prices must not vary with the quantity demanded by the consumer, or they would thereby be contaminated by consumer choices and tastes.^{11/}

Market price data is unfortunately rarely collected by household demographic-economic surveys. The level of local wages may be available, and household demand theory emphasizes the relevance of wage rates for men, women and children to the choice of activities people undertake, particularly market labor supply behavior. Other market prices are proxied by the availability of private and public services, where access is often a major part of "price" variation. For market prices to differ across a population at one point in time, there must be something separating markets, perhaps the costs of transporting goods to consumers or the costs of migrating to consume immobile services, or both. This suggests the concept of market price implicitly relates

to an aggregation of households in a market area. Cross-sectional studies require a sufficient number of market aggregates to provide a basis for estimating the effect of "market prices" on household demands across these local market areas. The appropriate level of aggregation to measure market prices depends on the good or service and what delimits its market. Clearly it is not appropriate to measure the household's price by observing what the household pays, as this will embody the household's demand for "quality" in that good, which is unlikely to be independent of its demand for "quantity". For example, the family that spends a great deal per visit to a doctor for their children cannot be viewed as confronting a higher price for medical care than neighbors who spend much less per visit to a doctor for their children. Conversely, a household located in a town that spends more on public health programs than does the average town will actually confront a lower "market" price of medical services, presumably of quality adjusted health care, that does not necessarily reflect the individual household's preferences. Market prices must be measured, then, at a level of aggregation above the household to assure that the prices are exogenous to the household's demands and behavioral outcomes which axiomatically depend also on the preferences of the household.^{12/}

Residential community characteristics may be particularly useful in estimating input demand relationships because they generally can be assumed exogenous from the household's point of view. Merging data on small community markets—their wages, prices, and service programs—is a natural step toward empirical application of the household production-demand approach to child survival. Extraction of "price" type variables from regionally aggregated data, from a census or consumer expenditure survey, for example, to explain individual household behavior is increasingly common in economic demography.

Endowments

The second class of demand determinants is the household's ownership of factors of production. The prior discussion suggested why one might want to treat the individual's education as such a predetermined labor endowment. It may be useful to also treat vocational training as an exogenous determinant of demands. Inherited nonhuman wealth may, by the same logic, be viewed as given to the individual or at least largely outside of the individual's control. Although the ownership of land and reproducible capital is generally viewed as exogenous to short run household consumer behavior and to producer input allocation decisions, these stocks of nonhuman wealth reflect lifetime accumulation behavior and may not be predetermined from the point of view of long-term demographic behavior. Again, drawing the line between exogenous constraints and endogenous choices becomes a critical matter of model specification. In this case, the characterization of an individual's initial endowments might be inferred from family history questionnaires that seek to distinguish among sources of inherited wealth and subsequent accumulation over the individual's adult lifetime. Clearly, one should distinguish between inherited wealth brought to the household by husband and wife and that brought by other members.

Environment

The third and final determinant of household demands is the productive environment that is outside the behavioral control of the individual or household. A local climate conducive to malaria, for example, may increase child mortality. The area may benefit from a public irrigation project that alters cropping possibilities and raises the market value of land, but spreads schistosomiasis. Resulting changes in cropping patterns may then affect women's employment opportunities, through encouraging, for example, rice or cotton

cultivation that assigns greater value to the tasks that women appear to have a comparative advantage in performing. These technological characteristics of the productive environment are related to geographical and administrative aggregates that need to be specified for each productive process or outcome. There are likely to be many ramifications of a change in the local productive environment, obviously, and the net effect of these on child survival may not always be clarified by theoretical insights but may simply require extensive empirical study.

Health Inputs

Another tier of data required to estimate the structural model relates to child health inputs. The standard medical care inputs must be summarily measured, for the statistical inefficiency of the two-stage estimation approach argues against the evaluation of a very long list of inputs, particularly when the prices of the inputs and the income elasticities of their demand are similar. In addition to medical care inputs, birth order of the child, spacing, number of siblings, and age of mother at birth are relevant data. All of these aspects of the timing of family formation are thought to influence the parent time and market resources available for child care and other consumption needs. Duration and extent of breastfeeding and nutritional inputs should be summarized, perhaps by anthropometric measures, such as height, weight and skinfolds. Time inputs of mothers and other persons performing child care functions should be quantified, and the educational attainment of these individuals recorded. Household consumption patterns that may affect child survival include whether the family has a protected water supply or modern sanitation facilities, and various aspects of quality of housing. The productive activities engaged in by household members, including the mother and child, may

affect the provision of child care or exposure to particular diseases or occupational hazards.

All of these potential inputs to child survival are to some degree choice variables, even though these household choices may not be strongly affected by their consequences for child survival. Occupational choices, for example, may be determined primarily by other factors (e.g., wage rates), even though they often entail exposure to different health risks. The income effects associated with the occupational choices may neutralize these differences in risks to child survival, or they may not. Direct correlational evidence is not likely to clarify these issues, just as we have slowly learned that the simple correlation of fertility and child mortality is not necessarily a satisfactory estimate of a causal relationship.

The third type of health programs, those that alter the exogenous health environment, can be treated as an input at the community level eradicating or controlling a local disease and thereby reducing child mortality risks for all persons. As suggested in Table 1, however, one has reason to expect such programs to benefit differentially the population according to their preprogram capacity to protect themselves through other inputs, such as mother's education, or father's wages, or household nonhuman wealth. The production function should be specified to include interactions between inputs in these cases where it is hypothesized that inputs are substitutes or complements for one another. The earlier discussion of mother's education and program effects could be generalized to many pairs of inputs. Empirical study of such potential interactions is possible whenever more than one of the inputs is measured at the household level. However, where both inputs are measured at the community level, multicollinearity across regions in various inputs and input prices may

make it difficult to observe sufficient independent variation in the inputs so as to test empirically for cross-substitution effects in the production function. Experimental design may ultimately be needed in the implementation of public programs across regions to introduce independent variation in the mix of programs that would permit one to confidently measure synergistic effects among policy inputs of this form. Nothing in the approach precludes, however, estimating cross-substitution effects among program inputs in the reduced-form child survival equation.^{13/}

An Empirical Example: U.S. Infant Health Determinants

Experience in estimating "structural" health production functions is still very limited. Only one study of infant health determinants has dealt with the statistical problem of health heterogeneity (Rosenzweig and Schultz, 1982a). The timing and extent of prenatal medical care is not consistently associated with infant health outcomes in large random samples of U.S. births (Eisen, 1979). This finding could be interpreted to suggest that prenatal medical care does not systematically increase birth weight or improve prospects for child survival. However, heterogeneity in health endowments could lead women with difficult pregnancies to secure earlier prenatal medical care and still report below average child health outcomes. If this form of "adverse selection" of compensatory early prenatal medical care were widespread, health heterogeneity would bias downward direct estimates of the "benefits" due to prenatal care.

To test this reasoning and apply the estimation strategy outlined in this paper, Table 2 reports the simple linear determinants of birth weight, gestation, and infant survival for a representative sample of over 9000 legitimate births in the United States from 1967 to 1969. Columns (1), (3), and (5)

Table 2

Proximate Determinants of Infant Health Indicators: U.S. 1967-69*

Dependent Variable:	Birth Weight (grams)		Gestation (weeks)		Infant Mortality (1 = died)	
	OLS (1)	IV ^a / (2)	OLS (3)	IV ^a / (4)	OLS (5)	IV ^a / (6)
<u>Explanatory Variables:</u>						
Delay in Prenatal Medical Care (months) ^a / ₂	-1.96 (3.11)	-23.6 (4.16)	-.0091 (3.21)	-.0512 (2.18)	-.00015 (.81)	.0043 (2.49)
Smoking while Pregnant (cigarettes per day) ^a / ₂	-10.0 (15.3)	-20.0 (3.84)	-.0082 (2.83)	-.0159 (.74)	.0000 (.04)	.0043 (2.74)
Age at Birth (years) ^a / ₂	3.05 (2.41)	.486 (.12)	-.0085 (1.52)	.0080 (.48)	-.0015 (3.12)	.0017 (1.37)
Parity at Birth ^a / ₂	22.9 (3.85)	67.0 (4.15)	.0044 (.26)	.0734 (1.10)	.0159 (13.4)	.0051 (1.05)
Black	-250. (16.6)	-242. (10.1)	-.725 (10.9)	-.639 (6.46)	.0235 (5.06)	.0297 (4.12)
1967	16.8 (1.21)	13.7 (.86)	.540 (8.89)	.522 (8.00)	-.0084 (1.98)	-.0097 (2.04)
1968	13.2 (.95)	13.5 (.90)	.110 (1.81)	.119 (1.91)	-.0011 (.27)	-.0017 (.39)
Intercept	3245. (110.)	3323. (31.2)	39.2 (302.)	38.8 (88.1)	.0282 (3.12)	.0612 (1.91)
R ²	.0540	-	.0242	-	.0287	-
F	78.39	32.37	34.12	30.13	40.52	14.73
Sample size = 9484						
Dependent Variable Mean (standard deviation)	3288. (568.)		39.06 (2.47)		.0306 (.172)	

^a/Endogenous variables estimated by instrumental variable (IV) method. Instruments include those cited in Rosenzweig and Schultz (1982a) plus husband's education by five discrete levels, milk prices, share of labor force in manufacturing, government, agriculture, construction, unemployment rate, population per ob-gyn, sex of child, part time share of jobs, ease of divorce laws, AFDC payment average, Medicaid includes family planning services.

*The absolute value of t ratios is reported in parentheses beneath the regression coefficients. Note that in the case of columns 5 and 6 the t ratios are not unbiased, because the dependent variable is a binary variable (infant mortality).

report ordinary least squares (OLS) estimates of the child health production functions that would be biased by the hypothesized form of health heterogeneity and self-selection of health inputs; the inputs analyzed are the mother's delay in seeking prenatal care, smoking while pregnant, age, and parity. Columns (2), (4), and (6) report the instrumental variable estimates of the same child health production functions that are statistically consistent in the presence of health heterogeneity. The following variables are selected as instruments: education of mother and father, father's income, and merged characteristics of the residential area such as prices, health and family planning programs, and labor force characteristics.

Health heterogeneity appears to mask almost completely the significant beneficial impact of early prenatal medical care on these three infant health indicators. The child health effect of postponing prenatal care is increased twelve fold on birthweight, five fold on gestation, and the sign of the effect on infant mortality is reversed and becomes statistically significant when the appropriate IV estimates replace the standard OLS estimates. A mother who delays one month in obtaining prenatal care raises her expected infant mortality in this sample by 14 percent according to the instrumental variables estimates (Col. 6) that allow for the health heterogeneity.

The strong effects of smoking on birth weight that are noted in many direct epidemiological studies of birth outcomes are also evident in this sample (see Col. 1). But allowing for health heterogeneity doubles the estimated effects of smoking while pregnant on birth weight, while the effect on infant mortality is also increased substantially.

Although the age and parity of the mother exhibit a strong linear (and non-linear) association with birthweight and infant survival in this sample, the

instrumental variable estimates suggest there is relatively little linear biological effect of mother's age on birth weight, gestation, or infant mortality, though parity continues to contribute to greater birth weight. Interactions of age and smoking, however, are deleterious to the child's birthweight (Rosenzweig and Schultz, 1982b).

The empirical analysis of a probability sample of U.S. births matched to residential regional characteristics suggests that inferences drawn regarding the child health effects of health inputs can be sensitive to whether or not heterogeneity in health endowments is present and correlated with health input behavior. In this empirical example, child health effects of early prenatal care are obscured by standard epidemiological methods, as has been the case in past studies based in individual data. Allowing for the self-selection of health inputs and the interactive role of health heterogeneity, one obtains strong evidence that early prenatal medical care is beneficial in this U.S. population.

Summary and Conclusions

Empirical study of the biological and behavioral determinants of child survival may proceed in two directions: either to estimate the reduced form unconditional relationships between child survival and individual- and community-level variables that are assumed to be exogenous to the many demographic, consumer, and producer choices a household makes over time (Figure 3, row 1b); or to estimate both the demand equations for health inputs (Figure 3, row 1a) and the production function linking health inputs to child survival by simultaneous structural equation methods (Figure 3, row 2).

The reduced-form equation for child survival may be estimated without imposing a great deal of structure on the problem, though specifying which

variables are exogenous and endogenous to child survival may prove controversial. The approach proposed here is that all events after the woman reaches economic independence of her parents and completes her schooling are reasonably treated as endogenous, and in particular the timing of her first birth may be seen as the first life-cycle decision that may subsequently be an important endogenous input to the health of her children.

The biological production function for child survival can only be estimated after more structure is imposed on the problem, because in the presence of health heterogeneity self-selected health inputs will tend to be correlated with the child's unobserved biological health endowment. To eliminate the bias due to self-selected inputs and heterogeneity, a two-stage estimation procedure is outlined that estimates first the health input demand equations, and then estimates indirectly the health production function (Figure 3).

In the estimation of either the reduced-form equation or the health input and production relationships, it is valuable to have data collected both for individuals from a representative sample of households and on prices, programs, and characteristics of the communities in which these households are located. Since the health inputs are often household choice variables, they are not independent of the health endowments of the children, nor to the preferences of household members. Consequently, the community aggregate variables on prices, wages, programs, and environmental conditions are essential for the analysis, because they are outside the control of individuals in the household, and hence, they are legitimately treated as exogenous instruments. This merger of micro and macro variables facilitates both types of empirical research strategies and draws the line more distinctly between endogenous and exogenous variables for the purposes of studying the determinants of child health. The

joint analysis of micro and macro variables also permits the study of plausible interactions between the two levels, such as that illustrated above between the mother's education and the community level of program activity. Micro-level observations on child survival, observed for the individual child or averaged for each mother over all her children, also provide a firmer basis to explore nonlinear response patterns than is feasible when analysis relies on aggregate measures of child mortality across regions.^{14/} On the other hand, market prices and certain forms of variation in the productive environment are conceived in terms of aggregates over geographic and administrative regions. Other aggregates of individuals might also be considered, where it is hypothesized that group identification influences the productive possibilities available to individuals, such as their membership in a certain caste, class, racial, or religious group. The absence of individual data, however, might conceal the interactive structure and nonlinear form of relationships, and confound simultaneous relationships and feedbacks.

If the general hypothesis of health heterogeneity is accepted, as I think it must be, the widely cited evidence of direct associations between proximate health inputs and child health outcomes cannot be accepted at face value as evidence of causal relationships. The bias embodied in these associations may be minimal, but I suspect it is frequently substantial. The problem of evaluating health inputs may be particularly serious when it comes to policy interventions that subsidize the self-selected curative medical services for acute illness.

More generally, if social and biological scientists are to proceed along convergent research strategies in their analyses of the determinants of child survival, they must move toward a common framework. This framework needs to

draw clearly the stochastic distinction between independent or exogenous variables and dependent or endogenous variables. Child survival should be viewed as one among many endogenous household variables that are modified by economic and demographic choices. Child survival may be usefully interpreted as directly conditioned on exogenous constraints and environmental conditions, such as regional public programs, prices, wages, and individual biological and economic endowments. To estimate the structure that links these exogenous variables to child survival will require added information about the problem. This paper has outlined one possible estimation strategy for clarifying this structure.

Notes

1/ Examples abound of this form of bias misleading the researcher who tries to characterize health technology from nonexperimental evidence. Efforts to estimate the effect of vaccinations on child survival, for example, are bedeviled by the tendency for children who "naturally" get vaccinated to be otherwise different from nonvaccinated children, namely they tend to receive other unmeasured beneficial health inputs.

2/ It is possible to identify the selection correction solely on the basis of functional form assumptions (Heckman, 1979), but this seems to place excessive reliance on nonlinear specification choices, such as implied by the probit or logit formulation.

3/ Mark Rosenzweig and I have examined the possibility that education of the mother may achieve an increase in birth weight by three means: altering the measurable mix of inputs to child health, enhancing the productivity of specific inputs differentially, and exerting a neutral productivity enhancing effect on all inputs. Thus far we have found most of the effect of mother's education on birth weight in the United States can be attributed to education's effect on the mix of inputs used by mothers, though the exception may be in the timing of the first visit to the doctor for prenatal care, which appears to be of enhanced value for more educated mothers.

4/ This section draws on the joint research the author has been engaged in

with Mark Rosenzweig over several years that is summarized in Rosenzweig and Schultz, 1982b.

5/ For further discussion with respect to mortality, see Schultz, 1980.

6/ Column A, second row (i.e., A.2) indicates that if the informational roles of both public health and family planning programs and maternal education are predominant, such programs are likely to have a greater effect on child health in families with less educated women, compared to families in which the mother has a higher level of schooling. This difference reflects the fact that the availability of information is only of value to households which have not already acquired the information, namely, to those who face higher costs of information acquisition. If programs mainly lower the user cost of modern health inputs (column B), however, higher educated families, who are already more aware of the "benefits" of such inputs, will be the ones benefiting most, given the informational role of education (B.2). With respect to the four other roles of education, it cannot be established a priori who will benefit more from the public provision of technological information (A) relating to the production of health, as the effects will depend on the type of information generated and on unknown differential income effects by mother's education level.

Programs that cause the remission or eradication of disease from the community environment, without otherwise affecting the regimen of people, (column C) have the potential for narrowing mortality differentials by educational class, assuming that before the program disease is more common among the least educated. Broadly based eradication campaigns are, therefore, likely to benefit most the least educated, since it is implicitly assumed that education's

effect on the pre-program incidence of disease through roles 1, 2, 3, and 5 outweighs any effect arising from role 4.

7/ For example, if it is found that the effects of health programs on child survival are greater in families with less educated mothers, the joint hypotheses that the programs reduce input costs and that more educated mothers are more able and willing to take advantage of such subsidized health services (B.2) can be rejected. This is, incidentally, a widely accepted view of the process by which public health services benefit disproportionately the more educated (Caldwell, 1979). It may be a valid hypothesis at very low levels of income and average education, such as in sub-Saharan Africa, but may also be rejected by data from higher income populations, such as in Latin America.

8/ In this context, it is not the size of an R^2 (or the portion of variation in individual behavior which one explains) that is the appropriate criterion for judging the "importance" of behavioral demands in the study of child survival. The importance of this relation is better evaluated in terms of the magnitude of "slope" coefficients in conjunction with the range of changes that might reasonably occur in the conditioning variables. Also the size of the slope coefficients relative to their standard error (i.e. t ratio) provides a useful indicator of the precision of the slope coefficients.

9/ Ideally, the theoretical structure will also imply the direction of some effects (one-tailed tests may then be applied) and restrictions across coefficients in a single or a group of demand equations. This stronger level of confirmation of theory, however, is uncommon where unconditional demand relationships are being estimated, because the demand and production parameters

cannot generally be separately identified. If the estimated slope of the relationship does not exceed 2.4 times its standard error (two tailed test, sample size exceeding 60), one cannot be confident that in even 99 percent of the random samples would the slope coefficient be different from zero. In this case, where that ratio is less than 2.4, the relationship can be viewed as suggestive, but not confirmed with much confidence. In the case of mutually exclusive and exhaustive discrete variables conditioning an outcome (e.g. racial groups), the joint F test may be employed, as the suppressed category is arbitrary and hence the magnitude of the individual class coefficient and associated t ratios are meaningful only relative to the (arbitrarily) suppressed class.

10/ For instance, this framework is not designed to help interpret the inverse partial correlation between ownership of consumer durables (toilets) and child survival, even though child survival may be affected by sanitation facilities. Neither can one interpret causally the simple association between women's employment behavior and their fertility. Both outcomes are jointly determined choices, just as the quantities of apples and oranges purchased by the consumer are joint decisions, to some degree, linked through the budget constraint and probably related in the utility function. We would not expect a regression of apples on oranges to be interpreted as suggesting that apple purchases have the associated effect on the purchases of oranges. They are both outcomes of a system of demand equations determined by prices, endowments, and preferences.

11/ The shadow prices of the final commodities produced in the home are not

identical to market prices, for they are an amalgam of the prices of time and goods inputs, and because the individual's time may be valued differently depending on what the individual does, the shadow prices reflect consumption and labor supply choices of the individual. Thus, the commodity prices discussed by Becker (1965) and later Willis (1974) with regard to the household's time allocation or choice of fertility do not satisfy the criteria for market prices that are appropriately treated as exogenous determinants of demand.

12/ This assumes explicitly that people do not migrate to a community where the available mix of public services fits their tastes. Once it is admitted that interregional migration is heavily determined by people's preferences for child survival (or closely related household commodities such as fertility), then regional characteristics are not a suitable basis for purging the input demand equations of the effect of health heterogeneity. Conversely, if governments locate health programs in regions that have an unusual average level of health endowment, say areas where malaria is still uncontrolled, then care must be exercised in interpreting the "effect" of such programs on child survival. In this case, the regional health program becomes itself a response to "endowments" and must be explained by the more basic governmental priorities in conjunction with regional characteristics. See Schultz, 1983.

13/ For example, Schultz (1971) estimated across 361 regions of Taiwan the cross-substitution effects on fertility of different types of field workers in the family planning program. Thus, in a locality that had already received much attention by "pregnancy health workers" the "village health education

workers" were notably less effective in reducing birth rates among older women. This conclusion was derived from a reduced-form specification of a fertility equation that assumed child mortality as predetermined.

14/ Categorical or higher order polynomials in the explanatory variables might be introduced. Box-Cox (1964) power transformations might also discriminate among alternative functional forms for the dependent variable, as well as those for the explanatory variables.

References

- Becker, G.S. (1965) "A Theory of the Allocation of Time," Economic Journal, 75, 299: 493-517, December.
- Box, G.E.P. and D.R. Cox (1964) "An Analysis of Transformations," Journal of the Royal Statistical Society, Ser. B 26:211-243.
- Butz, W.P., J. DaVanzo and J.P. Habicht (1981) Improving Infant Nutrition, Health, and Survival, R-2924-AID, The Rand Corp., Santa Monica, CA, June.
- Caldwell, J.C. (1979) "Education as a Factor in Mortality Decline," Population Studies, Vol. 33, No. 3: 395-413, November.
- Cochrane, S.H., J. Leslie and D.J. O'Hara, (1981) "Parental Education and Child Health," paper presented at Symposium on "Literacy, Education and Health Development," Ann Arbor, Michigan, March.
- Eisner, V., J.V. Brazie, M.W. Pratt, and A.C. Hexter (1979) "The Risk of Low Birth Weight," American Journal of Public Health", Vol. 68: 887-93.
- Farah, A.A. and S.H. Preston (1982) "Child Mortality Differentials in Sudan," Population and Development Review, Vol. 8, No. 2: 365-383.
- Griliches, Z. (1957) "Specification Bias in Estimates of Production Functions," Journal of Farm Economics, Vol. 39: 8-20.
- Grossman, M. (1972) "On the Concept of Health Capital and the Demand for Health," Journal of Political Economy, Vol. 80: 223-255, March/April.
- Heckman, J., (1979) "Sample Selection as a Specification Error," Econometrica 47, No. 1: 153-161, January.
- Manski, C.F. and D. McFadden (1981) Structural Analysis and Discrete Data with Econometric Applications, Cambridge, MA: MIT Press.

- Mincer, J. (1963) "Market Prices, Opportunity Costs and Income Effect," in Measurement in Economics (ed.) C. Christ, et al., Stanford University Press, Stanford, CA.
- Olsen, R.J. and K.I. Wolpin, (1983) "The Impact of Exogenous Child Mortality on Fertility," Econometrica, Vol. 51, No. 2: 731-750, May.
- Rosenzweig, M.R. and T.P. Schultz (1982a) "The Behavior of Mother as Inputs to Child Health," in Economic Aspects of Health, (ed.) V. Fuchs, Chicago: University of Chicago Press.
- _____ and _____ (1982b) "Child Mortality and Fertility in Colombia: Individual and Community Effect," Health Policy and Education, Vol. 2: 305-348.
- Schultz, T.P. (1971) Evaluation of Population Policies: A Framework for Analysis and Its Application to Taiwan's Family Planning Programs, R-643-AID, Santa Monica, CA: The Rand Corporation, June.
- _____ (1980) "Interpretation of Relations Among Mortality, Economics of the Household, and the Health Environment," in Proceedings of the Meeting on Socioeconomic Determinants and Consequences of Mortality, New York-Geneva, United Nations and World Health Organization, u.d.
- _____ (1983) "Migrant Behavior and the Effects of Regional Prices: Aspects of Migrant Selection in Colombia," Yale University Economic Growth Center Discussion Paper No. 443, New Haven, CT, August.
- Willis, R.J. (1974) "Economic Theory of Fertility Behavior," in The Economics of the Family (ed.) T.W. Schultz, University of Chicago Press, Chicago, IL.