PN-AAC-261
ISN-39040

Final Report to the Agency for International Development:

The Effects of Government Programs and Household Behavior on Fertility and Child Health: Analyses of Longitudinal Data from Colombia and the Philippines

in fulfillment of Contract OTR-0082-G-SS-2342-00

Mark R. Rosenzweig

University of Minnesota

Kenneth I. Wolpin

Ohio State University

November 1984

In this report we describe the results of three studies concerned with 1) the estimation of the causal effects on child health of resource allocations made by parents and 2) the evaluation of the child health effects of family planning and health programs. The central problem addressed in all studies is how estimates of the consequences of within-household resources or programs are biased when a) household members differ in characteristics ("endowments") unobserved by researchers and b) households or policy-makers distribute resources according to those characteristics. Thus the studies are concerned with both specifying and testing models of household and governmental behavior as well as econometric issues.

All three studies utilize longitudinal, micro data. In all three, the longitudinal dimension of the data is necessary but not always sufficient to obtain policy-relevant estimates. In all three studies, moreover, evidence is presented which shows that almost all prior studies evaluating programs or the effects of parental behavior which have ignored unobserved differences among households or areas are likely to have produced biased and potentially misleading estimates.

A. Study Summaries

The first study, in Part I, is concerned with how differences in characteristics among children (heterogeneity) affects the within-household and across-household distribution of resources allocated by parents to children. A dynamic model is specified which incorporates two essential realistic features: children are born sequentially and characteristics of children, prior to their birth, are uncertain. The implications of the dynamic model for the estimation of the effects of parental resource

allocations on children's health are described. An econometric methodology is developed to obtain (asymptotically) unbiased estimates of the effects on child health of such choice variables as age of the mother at the birth of the child, birth intervals, family size, breastfeeding and innoculations. Data on these variables and on age-standardized weight at birth and 6 months after birth for two or more children in the same family, obtained from the 1968-1974 Candelaria (Colombia) Promotora Survey, are used. The major results of the study are:

1. The positive effects of delayed childbearing and of broader birth intervals on birthweight are significantly underestimated when heterogeneity across households and across children within households are ignored in estimation.
2. The negative effect of birth order on birthweight is also significantly underestimated when heterogeneity is ignored.
3. The positive effects of breastfeeding on a child's weight-for-age may be overestimated in prior studies, as mother's appear more likely to breastfeed inherently healthier children.
4. Mothers with inherently healthier children are more likely to bear children at younger ages; hence the understatement in conventional estimates of the child health effects of maternal age.
5. A healthy child is more likely to be followed by a closely-spaced child than is a less healthy child within a family; however, families where children are <u>on average</u> inherently healthier exhibit wider average birth intervals. Thus, estimates of birth interval effects differ according to whether across-household and/or within-household heterogeneity is taken into account in estimation.

Both need to be taken into account if bias is to be eliminated and true, causal effects of parental decisions on children's health are to be estimated.

The study reported in Part II is concerned with the evaluation of the Promotora health program in Candelaria, Colombia. It differs from prior analyses (Heller and Drake) of the program in that the selective migration of households to Candelaria is taken into account. The study is therefore concerned with three related questions: 1) who migrates to an area with a newly-established health program--low-health households? high-fertility households? low-income households? 2) how are estimates of the effects of the single-site program biased due to selective migration? and 3) how can the program be evaluated when there is selective migration? The major results of the study are:

1. Migrants attracted to the Promotora health subsidy program who were of the same income, education and age as Candelaria residents had a) lower pre-program family size and b) children with lower pre-program levels of health compared to households residing in Candelaria prior to the program introduction.
2. As a consequence of the selective migration, policy-relevant estimates of the true effects of exposure to the program are significantly overstated when both migrants and residents are included in the sample used to evaluate the program "effect."
3. Exposure to the program, when migration selection is taken into account, significantly increased the age-standardized weight of children (the only health measure that could be used in the study) in households whose income was below the sample mean. The program

had no significant effect on child weight in households with incomes above the Candelaria mean.

4. Not taking into account within-family heterogeneity and birth spacing decisions of parents, highlighted in the first study described above, also produced misleading estimates of program exposure.

5. If the Promotora Program were continued, because of selective in-migration by low-health households, average (over all children) health levels of children could decline in Candelaria even though the Promotora program increases child health.

The study contained in Part III examines the consequences for program evaluation of programs being distributed non-randomly across areas. Almost all prior studies evaluating the effects of family planning and health programs have used cross-area information on program effects and outcomes of interest (fertility, health). However, the reasons for inter-area variability in program intensity were not addressed. In this study, a model of endogenous governmental program placement is specified in an environment in which localities differ in inherent healthiness, health-related attributes known to the public decision-makers but unobserved by the researcher. The study is concerned with the questions: 1) is there a theoretical rationale for family planning subsidies if population growth has no negative externalities (imposes no social costs beyond those borne by individual households)? 2) will family planning and/or health subsidies be distributed disproportionately to low or high-health households? 3) what are the biases in the estimates of the child health consequences of family planning and health programs when placement of the

programs is related to unobserved health-related area characteristics? The study uses longitudinal data on child health and the timing of the initiation of family planning and health programs across twenty barrios in the lowland, rice-producing areas of Laguna Province, the Philippines.

The principal results of the study are:

1. When there are health externalities (e.g., infectiousness), subsidies to fertility control may effectively substitute for health subsidies (services) if either a) reductions in contraceptive costs induce households to increase health investments in children or b) decreased family size increases child health biologically (as our findings in Part I suggest). Thus, if either condition is met, subsidization of family planning is warranted on efficiency grounds even in the absence of any population externalities.
2. If the above conditions are met, family planning services are more likely to be provided disproportionally to low-health households. Such compensatory distribution is efficient.
3. A <u>combination</u> of health and family planning services (when administratively feasible) is more efficient than the provision of either set of services alone or the provision of no services even when there are no population externalities and even if reductions in contraceptive costs do not increase health investments by households. A necessary condition is a health externality.
4. In accordance with the modeling, most of the Laguna barrios with programs had both family planning and health clinics by 1979.

5. Estimates that take into account endogenous program placement indicate that both the family planning and health programs significantly increased child health, as measured either by height-for-age or weight-for-age.

6. Both the family planning and health clinics were introduced earliest in the low-health barrios. The timing of program placement was therefore health-equalizing.

7. As a consequence of the negative relationship between pre-program health conditions and the dates of introduction of the programs, cross-section estimates of the health effects of the programs, as in prior studies, completely obscured the positive child health consequences of both the family planning and health clinics.

B. Implications for Data Collection

The studies have implications for the kinds of data required for policy-relevant estimation of program effects and for estimation of the causal consequences of household resource allocations. In general, when data, or more specifically, the survey sampling frame is less than ideal, it is necessary for the researcher to impose structure on the estimation procedures with regard to functional forms and error terms if account is to be taken of the problems enumerated in the studies here. The following points are elaborated in Parts I through III.

a. Program Evaluation: Child Health Effects

1. The ideal "data set" is created when programs are randomly allocated across areas. Cross-sectional comparisons of children born before the program (to residents) among selected and non-selected sites after a few years yield the policy-relevant estimates. No modeling is required.

2. When one site is randomly chosen for a program, longitudinal data are required on the outcomes of interest for <u>resident</u> household children born prior to the program, and methods used in the study described in Part II may be necessary because of selective migration.

3. If the single site where a program is implemented is not representative or not randomly-selected, then the best estimate of the program's effect in that site reveals little about the effectiveness of the program if implemented nationally.

4. If programs occur in multiple sites but the sites are non-randomly selected, longitudinal data are required and techniques as used in the study described in Part III may be necessary.

b. <u>Estimates of the Effects of Parental Behavior on Child Health</u>

The basic problems in estimating the direct effects of household resource allocations are a) such allocations are functions of unobserved traits of the children; i.e., the allocations are endogenous and b) current health status is a "stock" variable that is influenced not only by current resources but also by past resource allocation. Thus,

1. A sample of households in a single site surveyed at one point in time cannot be used to estimate how parental behaviors causally affect child health because a) no exogenous instruments (prices, programs) vary in the data and b) past resources allocated to each child would need to be ascertained retrospectively, and thus may be highly inaccurate.

2. If there are no available exogenous instruments--variables that influence resource allocations but are uncorrelated with health

characteristics (endowments)--because, for example, data are collected in one site, then longitudinal data on both child-specific health outcomes and on the life-cycle allocation of resources to <u>each</u> and <u>every</u> child in the family may be needed in order to obtain causal estimates of within-household resource allocations on child health. See Part I.

PART I. HETEROGENEITY, INTRAFAMILY DISTRIBUTION AND CHILD HEALTH: ESTIMATING THE HEALTH EFFECTS OF PARENTAL RESOURCE ALLOCATIONS

The estimation of the effects of household resources on the survival, health and well-being of children has been a central concern in the demographic, economic and medical literatures, (e.g. Heller and Drake, 1979; Olsen and Wolpin, 1983; DaVanzo, Butz and Habricht, 1983). One of the potential problems in obtaining estimates of the effects of such behavioral inputs as maternal age of childbearing, breastfeeding, and use of medical services on measures of child health is the existence of health-related factors known to or affecting parental decision makers but unobserved by the researcher. Variations in such unobserved factors (heterogeneity) in the sample population may provide misleading estimates of the causal relationships among parental choices and observed health outcomes. Yet few studies have been attentive to this problem.

There are two distinct sources of heterogeneity, with different implications for statistical treatment. First, there may be across-household variation in the health environment in which allocative decisions are made — mosquito infestation, sanitary conditions -- or in the inherent healthiness of parents, some of which is transmitted genetically to offspring. If parents take into consideration these household factors in their allocative decisions; for example, if households in healthier environments choose to have fewer children or to space them more widely, then the observed association between variations in such variables and measures of child health will overstate their consequences for child health. Use of information on siblings and a household fixed effect procedure circumvents this problem, given the invariance of these household health unobservables. However, only one study of the behavioral determinants of

child health has used this procedure (Olsen and Wolpin, 1983), where the importance of this type of heterogeneity is demonstrated.

A second source of heterogeneity arises from variations in the inherent qualities of children born within a family. Differences among children in healthiness or skills may affect how parents allocate resources across their offspring as well as parental fertility decisions. For example, it is well-known that an infant's intake of breastmilk depends on its ability to suckle; immature or ill infants may thus be breastfed less or not at all, leading to an upward bias in the estimation of the effects of breast-feeding on infant survival or nutritional status. The death of an infant may lead to a more closely-spaced subsequent child (the so-called replacement effect), with deleterious consequences for that child's health.

No studies of child health have attempted to deal with both intra and inter household heterogeneity. Rosenzweig and Schultz (1983) use an instrumental procedure to obtain estimates of the behavioral determinants of birthweight; however, their study assumes that parental schooling levels and husband's income are orthogonal to the unobserved factors associated with child health. If more educated and wealthier parents are also healthier and thus have inherently healthier children, however, their estimates will be inconsistent. The Olsen and Wolpin study ignores any responsiveness of parental allocations to variations in the healthiness of individual children.

Little empirical evidence exists on how resources are allocated across family members as a function of their "endowments," (Rosenzweig and Schultz, 1982). The existing theoretical literature on intrahousehold allocations (Becker and Tomes, 1976; Behrman et al., 1982; Sheshinski and Weiss, 1982) is deficient in providing insights into how parents respond to exogenous

variations in the inherent qualities of children, and thus on the direction of bias, if any, in studies ignoring such behavior and/or heterogeneity, chiefly because they assume that the qualities of all children are known by parents in advance, prior to their birth. However, early and important decisions about resource allocations to children cannot be fully informed about the characteristics of children yet unborn; such decisions are inherently dynamic and sequential (Wolpin, 1984).

In this paper, we formulate an illustrative dynamic model of intra-household allocative behavior incorporating variations in and uncertainty about individual child characteristics. The model is used to show how both the timing of childbearing and child-specific allocations vary with both household and child-specific health endowments. In part 2, we discuss the implications of the model for estimation of the behavioral determinants of child health and we use the information restrictions in the model associated with the sequencing of births to develop an estimation procedure which takes into account both intra and inter household heterogeneity. In part 3, longitudinal data on children and households from a village in Colombia are described and used to compare estimates of the effects of birth order, birth spacing and timing, per-capita food consumption, innoculations, and the incidence of breastfeeding on the age-standardized weight of children at two life-cycle points, at birth and within six months after birth. The estimates, obtained using ordinary least squares, a family fixed procedure, and the new procedure suggest the sensitivity of estimates to assumptions about heterogeneity and parental behavior. In particular, those procedures which ignore heterogeneity understate importantly the effects of birth order and birth spacing but overstate

the effects of breastfeeding. The consistent estimates obtained in part 3 are used in part 4 to compute estimates of the health endowments of individual children and of households and to estimate the effects of variations in such endowments on the behavioral variables. These estimates indicate that healthier households, for given income, have more children and more closely spaced children but consume no more food per-capita than do less well-endowed households. These results imply that households tend to reduce interfamily inequalities in child health. However, while the estimates suggest that parents are more likely to have a subsequent child quickly the more healthy is the prior (surviving) child, they are more likely to breastfeed an inherently healthier child.

1. The Model

a. Heterogeneity, the Health Technology and Information Restrictions

Assume that the health at birth h^0 of a child born to a particular family depends on its birth order, the timing of preceding births, the age of the mother at its birth, and prenatal child-specific resources. For child of order i, the (log) of health at birth is assumed to be given by

$$(1) \quad \log h^0_{i,t_i} = \gamma^0_1 t_i + \sum_{k=1}^{i-1} \gamma^0_{2,k} n_{i-k} \cdot (t_i - t_{i-k}) + \gamma^0_3 i + Z^P_i \gamma^0_4 + \mu + \varepsilon_i + V^0_i$$

where h^0_{i,t_i} is the health at birth of the child or order i born to a mother at age t_i, n_{i-k} is equal to one if a child of order i-k is born at mother's age t_{i-k}, and Z^P_i are a vector of prenatal inputs to child of order i. The randomness in observed initial health is due to a family health endowment common to all children within a family (μ), a child-specific health endowment common to all ages of a particular child (ε_i), and a purely random

serially uncorrelated draw (V_i^0). The following assumptions are made about these error components: $E(\mu_j\mu_k) = \sigma_\mu^2$ $(j=k) = 0$ $(j \neq k)$; $E(\varepsilon_{ij}\varepsilon_{k\ell}) = \sigma_\varepsilon^2$ $(i=k,\ j=\ell) = 0$ $(i\neq k,\ j\neq\ell)$; $E(V_{ij}^0\ V_{k\ell}^0) = \sigma_v^2$ $(i=k, j=\ell) = 0$ $(i\neq k, j\neq\ell)$; $E(\mu_j\varepsilon_{k\ell}) = E(\mu_j V_{k\ell}^0) = E(\varepsilon_{j\ell}V_{km}^0) = 0\ \forall_{j,k,\ell,m}$.

The health of child i at any age, a, after its birth may depend on the timing of the births of subsequent children (if any) and on post-natal resources Z_i^a allocated to it; thus, for child i

$$(2)\quad \log h_{i,t_i}^a = \gamma_1^a t_i + \sum_{k=1}^{i-1} \gamma_{2,k}^a n_{i-k}\cdot (t_i - t_{i-k}) + \sum_{k=1}^{\tau} \theta_{2,k}^a n_{t_{i+k}}\cdot(t_{i+k} - t_i)$$
$$+ \gamma_3^a i + Z_i^P \gamma_4^a + \sum_{k=1}^{a} Z^k \gamma_{5,k}^a + \mu + \varepsilon_i + V_i^a$$

Notice that prior inputs are assumed to potentially affect the stock of health at any age and that such inputs may not have uniform effects at all ages. Note also that due to the logarithmic specification the effects of all inputs on the level of a child's health depend on the magnitude of the child's health endowment, composed of the elements μ, ε_i and V_i.

Equations (1) and (2) describe the production technology relating the timing of births in the household and child-specific resources to a child's health at its birth and later in its life. Use of least squares or other single equation procedures to estimate the health technology parameters in (1) and (2) will yield unbiased estimates of these parameters only if the "inputs" are uncorrelated with both the household and child-specific endowments unobserved by the econometrician. The direction of the biases will in turn depend on whether the parents, when making their decisions about each of the inputs, observe the endowments, or components of them, and how such decisions are affected by such knowledge.

The sequential ordering of births places some important restrictions on parental information. The decision concerning when to have child i cannot, for example, depend on its child-specific endowment ε_i, which only becomes known after its birth, but may be informed by the household's health environment μ and may also depend on the perceived healthiness of prior children (ε_{i-1}). However, decisions about the level of post-birth resources Z_i^a allocated to child i may depend on (and will certainly be informed by) the initial healthiness of the child as well as on the healthiness of all prior children.

b. Parental Resource Allocations to Children

Given the existence of parental perceptions about the health environment in which they reside and about the individual, inherent traits of their children, little can be said *a priori* about how such information affects parental resource allocations to children without specifying parental objectives and constraints. Such a behavioral theory should also incorporate the biological characteristics describing the consequences of allocative decisions, as in (1) and (2), and the information constraints associated with the sequencing of births. To obtain some insights into how differences in healthiness across households and how differences in healthiness across children within households affect household allocative decisions, and thus how single equation estimates of biological relationships involving endogenous parental decisions in the presence of heterogeneity may be biased, we formulate a simple dynamic model.

Assume that the parents in each life-cycle period maximize the expected value of an intertemporally separable utility function that has as arguments

the mean H_t of the "final" child health outcomes of children in the household, the number of children in the household and a commodity X. Final child health is achieved at some arbitrary age A of the child, i.e., it is h^A_{i,t_i} for child of order i. Parents thus care both about the health and number M_t of their children, where $M_t = M_{t-1} + n_t$. The parents' problem is described by

$$(3) \quad \max_{Z_t, n_t} E[\sum_{\ell=t}^{T} \delta^{\ell-1} U(H_\ell, M_\ell, X_\ell)]$$

subject to the

per-period income constraint, which must be satisfied in each period,

$$(4) \quad F_t = w_t Z_t + X_t + P_t n_t,$$

where F_t = income, w_t = cost of a unit of resource Z, P_t = price of having a child; and to the "final" health equation (2) at a = A. Parents thus choose whether to have a child in each period and how much Z to allocate to that child after it is born and to all other children who have not yet reached their "final" health stock based on the information set Ω they have at the beginning of the period. Thus at the onset of period t, for example, parents know the household endowment μ, all their past decisions, the health technology (2), and the individual endowments (and thus health outcomes) of all prior children; they do not know the child-specific endowment ε_i of children to be born in t or after period t.

To simplify the model, assume that the decision horizon has four periods; children can be born at the beginning of period two, three, or four and health inputs are required only for one (the first) period of the child's life. Thus in the last period (four) only the level of Z for a child to be born in period four needs to be determined if the household had decided

during period three to have a child in period four. The technology of final health production is described by equation (2), except that, for simplicity, we will ignore all prenatal inputs except those associated with the spacing of births. The information sets associated with the beginning of each successive period are thus: $\Omega_1 = \{\mu; \Gamma\}$ $\Omega_2 = \{\mu, \varepsilon_1, n_1; \Gamma\}$, $\Omega_3 = \{\mu, \varepsilon_1, \varepsilon_2, n_1, n_2, Z_1,; \Gamma\}$, $\Omega_4 = \{\mu, \varepsilon_1, \varepsilon_2, \varepsilon_3, n_1, n_2, n_3, Z_1, Z_2; \Gamma\}$, where Γ represents the technology parameters.

To further simplify, assume that utility in each period is linear quadratic; thus in period four

$$(5) \quad U_4 = \alpha_1 H_4 - \alpha_2 (H_4)^2 + \beta_1 X_4 - \beta_2 (X_4)^2 + \delta M_4$$

Also assume for (innocuous) simplicity, that Z is a dichotomous variable, e.g., breastfeeding, taking on the value of 1 if Z is allocated to child i and the value of zero if it is not.

In such dynamic, forward-looking problems, it is not generally feasible to derive analytically the parental decisions rules for n_t and Z_t in any period (Wolpin (1984)). However, comparative statics can be performed readily for the fourth (final) period decision, when, in this case parents have full information about endowments. That is, the effects of the endowments of the children on the allocation of resources to the last child can be discerned in terms of the structural technological and preference parameters of the model.

Assume that it is optimal to have a child born in period four. Then at the beginning of the fourth period, the parents compare expected utility with Z = 1 to expected utility with Z = 0, given their information set Ω_4;

the difference in expected utilities J_4 is:

$$(6) \quad J_4 = E_4(U_4|Z = 1; \Omega_4) - E_4(U_4|Z = 0; \Omega_4).$$

Only if $J_4 > 0$ will Z be provided to this child.

In order to calculate J_4 explicitly it is necessary to make a distributional assumption about the random term V_i. If V_i is assumed to be normal with mean zero and variance σ_v^2, then the expected value of the health of child i to be born and breastfed in period four is given by:

$$(7) \quad E_4(h^A_{i,4}|Z = 1;\Omega_4)=\exp(\gamma^A_1 4 + \sum_{j=1}^{2} \gamma^A_2 n_j(4-t_j) + \gamma^A_3 i + \gamma^A_5 + \mu + \varepsilon_i + 1/2\ \sigma_v^2)$$

where γ^A_5 vanishes if child i is not breastfed. Let that part of (7) which contains all health determinants (inclusive of endowments) except Z be given by h^*_i; algebraic manipulation yields the following expression for J_4:

$$(8) \quad J_4 = \frac{h^*_{i,4}}{M_4} [\alpha_1(e^{\gamma_5} - 1) - \frac{\alpha_2}{M_4} \{\frac{h^*_{i,4}}{M_4} (e^{2\gamma_5} - 1) + 2(h^A_1 n_1 + h^A_2 n_2) (e^{\gamma_5} - 1)\}]$$

$$+ w_4 (\beta_2(2F_4 - w_4) - \beta_1)$$

The effect of a change in the child-specific endowment of child i born in period four on the value of J_4 for a family with any given prior allocations of n and Z is thus given by:

$$(9) \quad \frac{\partial J_4}{\partial \varepsilon_i} = \frac{h^*_{i,4}}{M_4} [\alpha_1 (e^{\gamma_5}-1) - \frac{\alpha_2}{M_4} \{2(e^{\gamma_5}-1)(h^A_1 n_1 + h^A_2 n_2) + \frac{h^*_{i,4}}{M_4} (e^{2\gamma_5}-1)\}],$$

where it will be recalled that $n_i = 0$ if no prior child is born. Expression (9) cannot be signed, as there are two opposing forces at work -- an increase in the child's endowment, given the technology described by (2),

raises the return to the resource Z and increases J_4. This positive substitution effect is embodied in the first term in (9). On the other hand, an increase in the child's endowment raises mean health directly and, given diminishing marginal utility as embodied in the parameter α_2 in (5), induces "wealthier" parents to spend their endowment on other resources. The sign and magnitude of (9) thus depend on both technology and preferences. Indeed, if the health technology were linear (rather than loglinear) it can be easily shown that the first term in (9) would vanish. Thus, in the case where endowments do not affect the productivity of inputs, more endowed children are likely to receive fewer resources; intrafamily behavior would tend to be equalizing or compensatory. When endowments augment resource returns, as in (2), the effects of intrafamily variations in child endowments on the allocation of resources across children cannot be known a priori.[1]

The effects of endowment variations across families on the allocation of resources to children are even more complex. The effect of a change in the family endowment μ on the likelihood that a child born in the last period receives resource Z consists of two effects. The first is given by expression (9); an increase in μ increases the last child's endowment and thus, for given prior fertility and health decisions, induces the substitution and wealth effects discussed. However, families with different endowments will not in general have identical fertility patterns and will not have invested identical resources across all prior children. Prior fertility and other investment decisions affect the direction of the family endowment effect on the likelihood that the last child receives resource Z, from (8), to the extent that i) the child's own health is affected (via prior spacing

decisions) and ii) mean child health levels (h_1^A, h_2^A) are altered. If, for example, variations in the household health endowment μ affected only the allocation of Z (no fertility responses), then the effect of variation in μ on the probability that the last child receives resource Z is given by:

$$(10)\quad \frac{\partial J_4}{\partial \mu} = \frac{\partial J_4}{\partial \varepsilon_i} - 2\alpha_2 \frac{h^*_{i,4}}{M_4} \left(h_1 \frac{dh_1^A}{d\mu} + \frac{dh_2^A}{d\mu} n_2\right)$$

where $dh_i^A/d\mu$ is the total effect of a change in μ on prior children's health inclusive of resource allocations. As can be seen, if more endowed families have healthier children (even if they invest less in them), then the effects of interfamily variation in endowments on the probability that the last child receives resource Z will be algebraically less (more negative) than the effect due to intrafamily endowment variation. This is because well-endowed families, given taste homogeneity, will receive less utility from any additions to mean child health than will less-endowed families.

In general, households with different endowment levels will exhibit different patterns of fertility and resource allocative behavior. To ascertain the effects of endowment variations on the complete life-cycle behavior of families and thus on the last period decision would require enormously complex calculations even in the simple dynamic model. For example, to solve for the effects of prior children's endowments on the decisions concerning whether to have a child in the third period and whether to breastfeed the child born in the second period (if it exists) requires a computation which must take into account the probability distribution of the third period child's endowment and the optimal fourth-period parental responses just discussed. At the beginning of the third period, parents must compare expected future utilities associated with their alternative fertility

choices and with their allocations of Z for all combinations of alternative choices in periods three and four, i.e., to discern whether J_3 is positive or negative; where J_3 is:

$$(11)\quad J_3 = \max\{E_3(U|n_3 = 1, Z_2 = 1; \Omega_3), E_3(U|n_3 = 0, Z_2 = 1; \Omega_3), E_3(U|n_3 = 1, Z_2 = 0, \Omega_3), E_3(U|n_3 = 0, Z_2 = 0; \Omega_3)\}$$

where $E_3(U|n_3 = 1, Z_2 = 1; \Omega_3) = E_3(U_3|n_3 = 1, Z_2 = 1) + E_3\{\max[E_4(U_4|Z_3 = 1, n_3 = 1, Z_2 = 1), E_3(U_4|Z_3 = 0, n_3 = 1, Z_2 = 1)]\}$

and E_3 is the expectation operator, given information at the beginning of period three. While no precise predictions can be derived from (11), the results indicate that both the timing of childbearing (and thus intervals between births) and the allocation of resources across children will generally depend differentially on the household's health environment (or parental endowments) and on the individual endowed healthiness of the children.

2. Estimating the Effects of Parental Choices on Child Health Outcomes and the Effects of Endowment Heterogeneity

The principal impediment to both achieving consistent estimates of health equations such as (1) and (2) and of parental responses to endowment differences among children is the absence of direct information on endowments. With neither the family endowments nor the child-specific endowments observable to researchers, it is clear from either static or dynamic intrafamily optimizing models that the right-hand-side health inputs in (1) and (2) will be correlated with the health "residuals" containing both the unobserved μ and the child-specific endowment. Least squares estimates of the γs will thus be biased.

Two procedures have been employed to circumvent the potential biases arising from endowment heterogeneity. Olsen and Wolpin (1983) employ data

on siblings and use a family fixed effect estimation procedure (FFE) to estimate a child mortality function. However, their procedures, which demonstrate the sensitivity of results to estimation techniques, purges only the family endowment component μ from the residual; their study implicitly assumes that parents do not respond to child-specific traits (they thus rule out, for example, "replacement" effects.). Rosenzweig and Schultz (1983) employ two-stage least squares to estimate a birthweight equation. Their procedure assumes, however, that household or child health endowments are orthogonal to parental characteristics such as schooling and income, an assumption that will be tested (and rejected) below.

The information restrictions of the dynamic model associated with the sequencing of births suggest that consistent estimates of the input effects Γ can be obtained from data on siblings by using both "lagged" inputs, from older siblings, and parental characteristics as instruments in a fixed effect procedure. In particular, since the information set of parents at time t in family cannot include the child-specific attributes ε_{ij} of children yet unborn, the following covariance restrictions are implied: cov $(Z^t_{ij}, \varepsilon^\tau_{kj}) = 0$, $t < \tau$, $i < k$; cov $(Z_{ij}, \varepsilon_{kj}) \neq 0$ $i \geq k$, where the superscript refers to time period; i.e., investments in child i at time t cannot be a function of child k's endowment ε_{kj} as long as they occur prior to child k's birth; Z_{kj} can be a function of both ε_{ij} and ε_{kj}.

Since the decision concerning when to have a child must be made in the absence of information on that child's specific endowments, sequencing additionally implies that cov $(n_{ij}, \varepsilon_{kj}) = 0$ $i \leq k$. This means that to estimate health outcome equations, all prenatal variables associated with child i will be appropriate instruments for differences in spacing and other prenatal inputs across child i and child i + 1. To see this, consider the birth outcome

difference equation for children one and two, from (1), with post-birth spacing variables appropriately deleted:

$$(12)\quad \ell n\tilde{h}^{0}_{2j} = (\gamma_1 + \gamma_2)\,\tilde{t}_{2j} + \gamma_4\,\tilde{Z}^{P}_{2j} + \tilde{\varepsilon}_{2j} + \tilde{V}^{0}_{2j}$$

where $\tilde{Y}_{2j} = Y_{2j} - Y_{1j}$.

As noted, OLS estimation of (12), equivalent to the family fixed effect or "sibling" difference method, would yield biased estimates of the γs, since $\tilde{t}_{2j}$ would be correlated with $\tilde{\varepsilon}_{2j}$, containing ε_{1j}. However, since t_{1j} and Z^{P}_{1j} are not correlated with either the unforeseen child specific endowments ε_{2j} or ε_{1j}, but are likely to be correlated with $\tilde{t}_{2j}$ and $\tilde{Z}_{2j}$, these lagged level variables are suitable instruments for (12) as well as the relevant difference equations for the post-birth health production technology in (2). Moreover, since the family component of the child's health endowment (the health environment, unobserved traits passed on from parents to children) is purged from (12), parental characteristics can also be used as instruments, since such characteristics (schooling, income) are unlikely to be correlated with the <u>deviations</u> of individual child traits among the offspring.

With appropriate information on birth outcomes, measures of child health, parental characteristics, and a family birth history, consistent estimates of the effects of maternal age, birth order, birth spacing and other parental inputs on health outcomes as well as of child endowments can thus be obtained using the lagged instrumental fixed effect (LIFE) procedure from families who have as few as two children. Since the residuals from such consistently-estimated birth outcome equations contain the child and family-specific endowment components, it is also possible to estimate the responses of the timing of births and the allocation of resources to individual children to those "initial" endowment components, if there are no missing child-invariant inputs (to estimate the effects of changes in μ_j) or missing child-specific inputs (to estimate child-specific endowment responses).

3. The Data

To implement the LIFE estimation procedure and to test for the importance of inter and intrafamily endowment effects in determining the timing of fertility and the allocation of resources among children, data are needed on parental characteristics, on demographic histories, and on health outcomees for multiple children within a household. We employ a unique data set from Candelaria, Colombia. These data were collected over a seven year period, from 1968 to 1974, to evaluate the impact of a program designed to provide child health services in all households in the town in which there were any children under the age of six. The services were provided by promotoras, who, at each visit (approximately every two months), also collected demographic and medical data on the individual children and parents. The data provide longitudinal information on the weight of all children under six during the entire survey period as well as information on such health inputs as innoculations (DPT) and breastfeeding.[2] There are also annual data on monthly food expenditures and family composition as well as basic socioeconomic information on parents, collected at the onset of the program. These data were analyzed by Heller and Drake (1979), who employed procedures which did not take into account any form of heterogeneity or dynamic behavior.[3]

To estimate the birth outcome equation (1) and post-birth health equation (2), we selected a subsample of 109 households in which at least two children were born during the seven-year program. For this subsample, information is thus available on health status at birth and on early post-birth input allocations for two or more siblings. The sample size is 238 children. An advantage of the data set is that none of the information was

collected retrospectively; thus the results obtained are not subject to recall error. However, the need for two or more siblings clearly results in a choice-based sample (households with higher fertility) and a relatively small sample size.[4]

We employ as a measure of health status the child's weight standardized for his or her age (in months) observed at birth and within six months after birth (the first post-birth observation).[5] The estimating equations are:

$$(13)\quad WT^0_{ij} = \gamma^0 \varepsilon_{ij} \mu^0_j \, age^{\gamma^0_1}_{ij} \, int^{\gamma^0_2}_{ij} \, order^{\gamma^0_3}_{ij} \, food^{\gamma^0_4}_{ij} \, e^{\gamma^0_5 \, sex_{ij}}$$

$$(14)\quad WT_{ij} = \gamma \varepsilon_{ij} \mu_j \, age^{\gamma_1}_{ij} \, int^{\gamma_2}_{ij} \, order^{\gamma_3}_{ij} \, food^{\gamma_4}_{ij} \, e^{\gamma_5 \, sex_{ij} + \gamma_6 \, DPT_{ij} + \gamma_7 \, bf_{ij}}$$

where age_{ij} = maternal age at birth (of child i in family j), int_{ij} = prior interval, $order_{ij}$ = birth order, $food_{ij}$ = per-capita monthly food expenditure in household, sex_{ij} = 1 if the child is male, DPT_{ij} = 1 if child innoculated against DPT, and bf_{ij} = 1 if child breastfed.

Table 1 provides descriptive statistics on the sample children and households. The first two columns pertain to the sample of households who had at least two children during the Promotora program; the second two columns refer to the households who had one or more children born during the seven-year survey period.[6] This sample will be used to estimate the effects of endowment variations on parental decisions. All but the food expenditure variable of the set of household variables are used as instruments in obtaining the LIFE estimates of (13) and (14) along with the lagged maternal age at birth, birth order, and interval variables.

Table 1

Descriptive Statistics: Two-Child and One-Child Samples

Sample Variable	At Least Two Children Mean	At Least Two Children S.D.	At Least One Child Mean	At Least One Child S.D.
Children	All Children		First Children	
Normalized Weight	.985	.186	.988	.192
Birth Order	5.29	2.86	4.62	2.85
Maternal Age at Birth (years)	27.6	5.98	27.3	6.32
Prior Interval (months)	23.5	14.7	27.6	19.5
Number of Older Siblings < 6	2.94	.877	2.40	.877
Breastfed	.885	.320	.888	.100
Innoculated (DPT)	.219	.415	.263	.189
Sex (male = 1)	.529	.500	.520	.500
Sample Size	238		383	
Families				
Years of Schooling - Mother	2.41	1.68	2.49	1.65
No Schooling - Mother	.211	.409	.179	.385
Monthly Income (pesos)	884	226	892	254
Per-Capita Food Expenditure	31.6	13.9	33.2	18.8
Enrolled in Family Planning Program	.0361	.188	.0493	.217
Sample Size	109		223	

4. Empirical Results: Behavioral Determinants of Normalized Weight

Because the sample selection rule may introduce bias into the least squares estimates of the weight-for-age equations (13) and (14) in addition to that resulting from health heterogeneity, a selection correction procedure was employed in which the determinants of the probability that the household was selected was first estimated as a function of the household characteristics.[7] These estimates were then used to predict the probability of sample inclusion for the sub-sample from which the γ estimates are obtained (Olsen, 1983). Because the family fixed effect and LIFE procedures purge out all household-level variables, no selection-correction variable is included when these procedures are used. All estimates, of course, pertain to children who lived for at least three to six months. In addition, to exploit estimation efficiencies, the two age-specific weight equations are estimated jointly as a system.

Table 2 reports estimates of the parameters of the normalized weight equations, obtained using seemingly-unrelated-regression (SUR), the family fixed effect procedure (FFE) and the lagged instrumental fixed effect technique (LIFE). Both the FFE method, which "corrects" for interfamily heterogeneity and within-family child-invariant omitted variables, and the LIFE method, which avoids as well biases associated with intrafamily heterogeneity, yield results which differ from those obtained using SUR and from each other. In particular, the negative effect of birth order on weight at birth appears to be understated significantly by both the SUR and FFE methods compared to the instrumental method -- the LIFE birth order coefficient in absolute value is double that provided by the FFE method and almost three-fold

Table 2

Behavioral Determinants of Log of Normalized Weight: At Birth and Within 6 Months After Birth

Estimation Procedure/Input	SUR (1)	SUR (2)	FFE (1)	FFE (2)	LIFE (1)	LIFE (2)
Sex (Male = 1)	-.0407 (1.59)[a]	-.0434 (2.06)	-.0425 (1.30)	-.0291 (1.25)	-.0410 (1.25)	-.0341 (1.26)
Maternal Age at Birth[c,d]	.0665 (0.82)	.0460 (0.69)	.310 (0.69)	.147 (0.45)	.761 (1.35)	-.488 (1.03)
Prior Interval[c,d]	.0404 (2.33)	.0306 (2.14)	.0501 (2.01)	.0311 (1.73)	.0563 (2.08)	.0224 (0.92)
Birth Order[c,d]	-.0842 (2.88)	-.0726 (2.96)	-.120 (1.14)	-.0853 (1.13)	-.244 (1.83)	-.0230 (0.21)
Breastfed[d]	-	.0316 (1.12)	-	-.0106 (0.24)	-	-.0358 (0.35)
Innoculated[d]	-	.0259 (1.22)	-	.0364 (1.29)	-	.0598 (1.15)
Food Per-Capita[c,d]	.0003 (0.25)	.0284 (1.23)	.00208 (0.31)	.0119 (0.28)	.00130 (0.08)	.133 (1.69)
λ[e]	-.164 (0.88)	-.265 (1.72)	-	-	-	-
Intercept	-.215 (0.85)	-.264 (1.09)	-	-	-	-
R^2	.056[b]	.092[b]	-	-	-	-
n	238		238		238	

a. Asymptotic t-values in parentheses.

b. From OLS regression.

c. Log of variable.

d. Endogenous variable.

e. Selection-correction variable.

larger than the SUR estimate of the birth order effect. The positive effect of the size of the interval preceding a birth on birthweight, statistically significant under all procedures, is ten percent greater when the LIFE method is employed compared to using the FFE method and is almost 49 percent higher than the corresponding SUR estimate. Moreover, while the SUR estimates suggest that children who are breastfed experience (marginally significantly) greater weight gains, the breastfeeding coefficients are neither positive nor significant when estimated with either the family fixed effect or LIFE methods. While this result does not necessarily imply that breastfeeding is ineffective (since the effect of breastfeeding depends on its duration and intensity and breastfeeding may augment survival), the estimates suggest that inattention to heterogeneity overstates the effects of breastfeeding incidence on children's weight and understates the effects of interval, length and birth order.[8] Moreover, the effects of household food consumption per-capita, and to a lesser extent, of innoculations appear also to be understated using either the SUR or FFE methods; but neglect of heterogeneity across and within households appears to lead to an overestimate of the persistent effects of birth order and birth intervals on post-birth weight.

While many of the individual coefficients are not measured with much precision, application of the Wu/Hausman test indicates rejection of the hypotheses that the behavioral inputs are uncorrelated with the residuals in the equations estimated by the SUR and FFE methods at the five percent level (F-test). Heterogeneity both within and across the sample households appears to be affecting the sample variation in the inputs and thus the estimated coefficients. Moreover, the magnitudes of the consistently-estimated effects (from the LIFE estimates) of some of the variables on weight

are not trivial -- increasing the mean birth interval from two to four years increases weight at birth by 16 percent (interval plus age effect); an increase in monthly per-capita food intake by 20 percent and early innoculation against diptheria, polio or tetanus raises weight-for-age within six months after birth by 2.6 percent and six percent respectively.

5. Empirical Results: Intra and Interhousehold Heterogeneity and Household Resource Allocations

As noted, the residuals $\hat{\eta}^a_{ij}$, obtained by subtracting the predicted standardized weight values based on the consistently estimated (LIFE) parameters from actual standardized weight values, contain the child-invariant household endowment, the child specific endowment, and a random error. By averaging the $\hat{\eta}^a_{ij}$ over all children i for the two periods in a family j, a consistent estimate of the family "effect" for family j $\hat{\mu}_j$ may be obtained since $\underset{i\to\infty}{\text{plim}}\ (\mu_j + \varepsilon_{ij} + V^a_{ij}) = \mu_j$. Child-specific deviations of the $\hat{\eta}^a_{ij}$ from $\hat{\mu}_j$ averaged over two periods provide an estimate of the child-specific effects $\hat{\varepsilon}_{ij}$ for family

Interpretations of each of the two residual components μ and ε_{ij} plus random measurement error as endowments/requires different assumptions about the completeness of the set of health inputs in (13) and (14). The family effect, $\hat{\mu}_j$, will unambiguously represent the exogenous health endowment of the family only if there are no omitted child-invariant endogenous variables in (13) or (14), a strong assumption. The violation of this assumption does not, of course, mean that the FFE or LIFE estimates of the γs are inconsistent (that must be due to (optimizing) behavior with respect to the child-varying inputs based on household information about the ε_{ij}). Rather, variations in $\hat{\mu}_j$ may then be due to interfamily variations in unobserved endogenous inputs and thus may reflect interfamily heterogeneity in both preferences and endowments. The

21.

residually-estimated ε_{ij}s, however, will represent child-specific endowments plus random measurement error to the extent that there are no important inputs which vary across children <u>within</u> a family, a weaker assumption. The associations between the $\hat{\varepsilon}_{ij}$ and family allocation decisions may correspond more to endowment effects than will the associations between the $\hat{\mu}_j$ and such household behavior.[9]

To estimate how variations in the health endowments of households are related to the across-household variations in fertility and household per-capita nutritional intake, we regressed the number of children less than six years of age, children ever born and monthly per-capita food consumption at the start of the survey period (1968), and maternal age at the birth of the (first) child born during the sample period on the computed household health endowment and a set of parental socioeconomic variables including the mother's schooling attainment and predicted family income based on the father's schooling, age and occupation.[10] Because exclusion of households who had less than two children during the survey period from the sample would obviously impart bias to these fertility and consumption estimates, we employed the augmented sample of households, including as well those who had only one child in the seven-year survey period. To compute the household and child-specific health endowments for the "one-child" households, we first regressed the estimated household endowments $\hat{\mu}_j$ on the total child residuals $\hat{\eta}_{ij}$ using the two-plus child sample. The estimates were then used to predict the household and child-specific endowments based on the child residuals (or total child endowments) computed from the information on the relevant life-cycle weight and input variables for each of the children born during the survey period in the "one-child" household sample (using the LIFE estimates of Table 2).

Table 3 reports the regressions employing the computed and estimated household endowments for the augmented sample. As can be seen, differences among households in the inherent average healthiness of their children is significantly correlated with inter-household differences in the pace and magnitude of fertility--households with better-endowed children exhibit significantly higher cumulative fertility and tend to have births significantly earlier; such households do not, however, consume significantly different levels of food per capita, controlling for the schooling attainment of the mother and income. Since such fertility behavior, given the estimates of Table 2, tends unambiguously to diminish the nutritional status of children at birth, the results imply that children born in healthier households, net of family inputs and income, tend to receive less favorable inputs. Inherent across-household inequalities in children's healthiness appear to be reduced by household fertility behavior.

The endowment point estimates (which are biased to zero) indicate that in households in which children on average are ten percent heavier at birth than average children in the town population (net of parental resources), the number of children ever born is higher by about one-half child and the mother accelerated the timing of the first birth during the survey period by over one year. The LIFE estimates of Table 2 suggest that such adjustments in fertility behavior would reduce weight at birth by 5.4 percent. About one-half of the initial weight advantage is thus erased due to fertility responses to family health endowment variation; children in high-μ households retain their inherent advantage on net ($dh/d\mu > 0$ in equation (10)).

The vector of socioeconomic variables is also significantly correlated with each of the fertility and food consumption variables in Table 3. The set of parental variables is also , however, significantly correlated with the computed family health endowment (five percent significance level).[11]

Table 3

Family Endowments, Fertility and Per-Capita

Food Expenditure

Variable	Children Ever Born	Maternal Age at Birth	Children Under Six	Per-Capita Food Expenditure
Family Health Endowment	.992 (1.67)[a]	-10.8 (9.42)	.359 (1.90)	.309 (0.08)
Income ($x10^{-3}$)	-.818 (0.80)	-.681 (0.36)	-.176 (0.55)	20.5 (3.06)
Schooling of Mother	-.286 (2.28)	-.620 (2.67)	.00629 (0.16)	1.34 (1.64)
Family Planning	-.818 (0.80)	.965 (0.51)	-.146 (0.46)	-2.69 (0.40)
Intercept	6.26 (6.90)	26.82 (15.8)	2.63 (9.27)	11.9 (2.01)
R^2	.055	.362	.021	.074
d.f.	218	218	218	218

a. t-values in parentheses.

This latter result implies that "healthiness" net of parental resources is not distributed randomly across the population with respect to either the educational attainment of parents or income. Thus, estimates of income or schooling effects on fertility and other health-related variables which ignore health heterogeneity across households may also be biased.[12] Indeed, the health endowment has a stronger relationship with the fertility variables than does income. The estimates indicate that, for given health endowments, income is not significantly associated with the number or spacing of children; however, higher income families consume more food per-capita. These estimates suggest that interventions that improve the health environment may induce somewhat higher fertility levels; however, income-augmenting projects would appear to have little effect on fertility. Moreover, mothers with higher levels of schooling have significantly lower family size, although they tend to have children earlier. Since only eleven of the 223 sample households contained a mother who was enrolled in the family planning program, no precise estimates can be obtained of the effects of this program; however, the relevant coefficient signs suggest that the program may be lowering fertility.

The estimates of Table 3 suggest that observationally identical households with differing health endowments exhibit significantly different fertility behavior, such that inherently healthier children appear to receive less favorable allocations. To ascertain if within-household disparities in child health endowments are exacerbated or lessened by intra-family parental allocative behavior, we estimated the effects of variations in two child-specific endowments--the health endowment as measured by $\hat{\varepsilon}_{ij}$ and the gender of the first child born in the sample period--on the subsequent fertility behavior of the parents and on the probabilities that the child is

breastfed and/or receives the DPT vaccine. Table 4 reports the maximum likelihood logit estimates of the probabilities of a subsequent short fertility interval (within three years after the birth of the first sample child), of the child being breastfed, and of the child being provided the DPT innoculation, as functions of the two child-specific endowments, the household endowment and the socioeconomic variables. These estimates indicate that while the set of socioeconomic variables is not statistically significantly related to the dependent variable in any equation, resource allocations within the household do respond to exogenous variations in the characteristics of children, although not uniformly. In particular, children with higher-than-average health endowments within the family are significantly more likely to have a more closely-spaced younger sibling than their less well-endowed siblings, but are also more likely to be breastfed. This latter result suggests why use of the single equation procedure may have overstated the "effect" of breastfeeding incidence on child weight in Table 2; as indicated in the model, evidently the returns to breastfeeding depend positively on the inherent healthiness of the child. On the other hand, the closer spacing following the birth of a healthier-than-average (or expected) child may reflect mainly an "income" effect, with parents "spending" their additional unanticipated wealth (endowment) on additional or more rapidly-accumulated children. Finally, despite boys having a weight disadvantage at birth (Table 2), neither subsequent spacing nor the probability of a child receiving breastmilk appears to be related to gender; innoculations, however, appear to be provided to boys more often than to girls but to be orthogonal to health endowments measured by weight-for-age.

Table 4

Maximum Likelihoood Logit Estimates: Family and Child-Specific Endowment Effects on Post-Birth Interval, Breastfeeding, Innoculation

Variable	Short Interval	Breastfed	Innoculated
Child Endowment	7.91 (5.45)	2.82 (1.44)	.244 (0.55)
Family Endowment	-1.19 (2.61)	.416 (0.54)	.430 (0.30)
Sex of Child (male=1)	-.175 (0.51)	.0261 (0.01)	.767 (2.33)
Income ($\times 10^{-3}$)	-.537 (0.54)	.863 (0.74)	.448 (0.61)
Schooling of Mother	-.0684 (0.75)	-.184 (1.34)	.0505 (0.51)
Family Planning	-.981 (1.14)	.218 (0.12)	.848 (1.34)
Intercept	.605 (0.68)	2.09 (1.85)	-2.79 (2.94)
d.f.	217	217	217

a. Asymptotic t-values in parentheses.

6. Conclusion

While there is a large scientific literature concerned with the child health consequences of household decisions, interest in the determinants of household decision-making over time has just begun. Few empirical studies of health have thus taken into consideration parental dynamic behavior. In this paper, we have formulated a simple dynamic model incorporating uncertainty to demonstrate the complexity of household decision rules concerning the allocation of resources to and across children when there is both unanticipated and sequential variation in child traits within the family and variation in healthiness across households. Estimates of the effects of the timing and level of fertility, use of medical services, food consumption, and breastfeeding on early measures of childrens' nutritional status were obtained based on an estimation procedure informed by the dynamic model. These estimates were compared to estimates obtained using procedures which ignore either or both intrafamily health heterogeneity and parental adjustments to child-specific-health shocks.

The results, obtained from a longitudinal sample of households in Colombia, suggested that, consistent with the model, parental behavior appears to respond to unanticipated health outcomes among children and is also significantly associated with more persistent health factors, unrecorded in the data, that vary across households. As a consequence, estimates of the child health effects of parental decisions, or the fertility effects of child mortality, ignoring the behavioral consequences of inter and intrafamily heterogeneity would appear to be biased. In particular, our results indicated that single-equation or family fixed effect techniques underestimate the negative consequences for birthweight of high fertility and short birth intervals, but overstate them

for post-birth weight. Moreover, as an evident consequence of inherently healthier children being more likely to be breastfed, the estimates neglecting heterogeneity appear to overstate the positive effects of breastfeeding.

Estimates of the effects of within-household and across household variation in endowments also appeared to suggest that i) the "endowed" healthiness of households net of parental resources allocated to children was a more important determinant of fertility behavior than income, with healthier households evidently having more children at earlier ages, and ii) within households, healther surviving children are more likely to be followed by a closely-spaced, subsequent child and to be breastfed. These results imply that existing estimates of fertility responses to child mortality confound intra and interhousehold endowment effects.

A cost of our estimation procedure, which makes use of longitudinal information on multiple children within a household to obtain production function estimates immune to missing household-level information and the existence of dynamic adjustments by parents, is low sample size and consequent loss of estimation precision. Our results imply, however, that cross-sectional samples taken from populations with little observed variation in exogenous variables (excluding parental characteristics), no matter how large or detailed, would be inadequate for obtaining consistent estimates of the consequences of parental resource allocations or of fertility behavior for child health or mortality. Moreover, longitudinal data on single children (no siblings) may also be inadequate, to the extent that there is little intertemporal variability in exogenous variables and, net of child-specific fixed effects, serial correlation in endowments over

time for a child is important relative to serial correlation of endowments across siblings net of both family and child-specific endowments. Finally, while we have estimated directly the parameters describing the health technology, no attempt was made to estimate the parameters characterizing parental preferences, thus our estimates pertaining to parental responses to within and across household endowment variation are merely first-order approximations to family behavior rules, and are subject to the usual caveats about reduced form estimates.

References

Beck, Gerald J. and B. J. van den Berg, "The Relationship of the Rate of Intrauterine Growth of Low-Birth Weight Infants to Later Growth," Journal of Pediatrics 86: 504-511, April 1975.

Becker, Gary S. and N. Tomes, "Child Endowments and the Quantity and Quality of Children," Journal of Political Economy 84: S143-S162, August 1976.

Behrman, Jere, R. Pollak and P. Taubman, "Parental Preferences and Provisions for Progeny," Journal of Political Economy 90: 52-73, February 1982.

Chernichovsky, Dov and D. Coate, "The Choice of Diet for Young Children and its Relation to Children's Growth," Journal of Human Resources XV: 255-263, Spring 1980.

DaVanzo, Julie, W. P. Butz and J-P. Habicht, "How Biological and Behavioral Influences on Mortality in Malaysia Vary During the First Year of Life," Population Studies 37: 381-402, November 1983.

Hausman, J. A., "Specification Tests in Econometrics," Econometrica 46: 1251-1271, November 1978.

Heller, Peter S. and W. D. Drake, "Malnutrition, Child Morbidity and the Family Decision Process," Journal of Development Economics 6: 203-235, October 1979.

Olsen, Randall, "Estimating the Fertility Replacement Effect," American Economic Review 73: 34-37, May 1983.

Olsen, Randall and K. Wolpin, "The Impact of Exogenous Child Mortality on Fertility: A Waiting Time Regression with Dynamic Regressors," Econometrica: 731-749, May 1983.

Rosenzweig, Mark and T. P. Schultz, "Market Opportunities, Genetic Endowments and Intrafamily Resource Distribution: Child Survival in Rural India," American Economic Review 72: 803-815, September 1982.

Rosenzweig, Mark R. and T. P. Schultz, "Consumer Demand and Household Production: The Relationship between Fertility and Child Mortality," American Economic Review 73: 38-42, May 1983a.

Rosenzweig, Mark R. and T. P. Schultz, "Estimating a Household Production Function: Heterogeneity, the Demand for Health Inputs, and their Effects on Birthweight," Journal of Political Economy 92: 723-746, October 1983b.

Sheshinski, Eytan and Y. Weiss, "Inequality Within and Between Families," Journal of Political Economy 90: 105-217, February 1982.

Wolfe, Barbara and J. R. Behrman, "Is Income Overrated in Determining Adequate Nutrition?," Economic Development and Cultural Change 31: 525-550, April 1983.

Wolpin, Kenneth I., "An Estimable Dynamic Stochastic Model of Fertility and Child Mortality," Journal of Political Economy 92: August 1984.

Wu, De-Min, "Alternative Tests of Independence Between Stochastic Regressors and Disturbances," Econometrica 41: 733-750, July 1973.

Footnotes

1. Additional ambiguity results when parents are also directly concerned about health disparities across their children. Assume that the variance in child-specific health outcomes enters linearly in the quadratic utility function (5) with a coefficient of $-\alpha_3$. Then the following ambiguously-signed term is added to expression (9):

$$(9') \quad \frac{2\alpha_3}{M_4^2} h^*_{i,4} \left[(h_1^A n_1 + h_2^A n_2)(e^{\gamma 5}-1) - h^*_{i,4}(M_4 e^{\gamma 5}-1)\right]$$

$$= -\frac{2\alpha_3}{M_4} h^*_{i,4} \left[e^{\gamma 5}(h^*_{i,4} - H_4) + H_4 - h^*_{i,4}(2 - e^{\gamma 5})/M_4\right]$$

The sign of (9') depends in part on whether the health of the final-period child net of the effect of the Z resource exceeds the mean health of all children inclusive of prior resources. Since less-endowed prior children may have greater health outcomes than does a subsequent child, due, for example, to negative maternal age and birth order effects, with inequality-averse parents it is thus possible, even when the health technology is linear, for a better-endowed last child to receive resource Z.

2. Height information was also collected, but only after two years of the program had elapsed. Restriction of our sub-sample (described below) to children with both height and weight information would have reduced the sample size by 40 percent.

3. Indeed, their specifications yield results that are not interpretable as estimates of either technology or preferences; the usefulness of their partial correlations is unclear.

4. We test for selectivity below. Note that if we had solved the dynamic model for both resource and fertility decisions, such a correction would be automatic in a full maximum likelihood approach.

5. Weight at birth has been shown to be a significant predictor of physical growth, development and morbidity; see for example Chernichovsky and Coate (1980) and Beck and van den Berg (1975). No study of the consequences of early child health status variables has taken into account heterogeneity, however.

6. There were 640 households in the original data containing a mother of childbearing age with children less than age 7 sometime during the sample period and with no missing information on the relevant variables used in the analysis. Of these, 223 had at least one child born during the sample period for which the relevant data were recorded. Because of village immigration and outmigration during the 7-year period the mean number of years of sample exposure for households is 3.8. All but 10 of the 109 households bearing two or more children were in the sample the full 7 years.

7. The sample selection equation included all of the family-level variables listed in Table 1, excluding per-capita food expenditure but including the ages of the mother and father in 1968, when the promotora program began. Not surprisingly, maternal age in 1968 and family planning enrollment were the two most significant determinants of sample inclusion; both variables were negatively associated with the probability of meeting the sample criteria.

8. The breastfeeding results are similar to those reported in Olsen and Wolpin (1983); correction for across-household heterogeneity reduced significantly the apparent positive breastfeeding effect on child survival

9. Olsen and Wolpin (1983) and Rosenzweig and Schultz (1983a and 1983b) also employ production function residuals to estimate behavioral responses to health endowments. None of these studies distinguish between adjustments to unanticipated child-specific shocks and inter-family endowment heterogeneity. Olsen (1983) attempts to decompose the child-specific (mortality) production function residual into the relevant child and family components and to estimate the fertility response to an unanticipated child death. However, his production function estimates are obtained using the family fixed effect method, which assumes the absence of intrafamily responses. His finding of a significant "replacement" effect indicates that his estimates and those of Wolpin and Olsen are thus inconsistent.

10. The first-stage income estimates are:

$$\begin{aligned}\text{income} = {} & \underset{(5.14)}{956} - \underset{(1.10)}{10.1}\ \text{agefather} + \underset{(1.26)}{.146}\ (\text{agefather})^2 - \underset{(1.92)}{363}\ (\text{agefather missing}) \\ & + \underset{(3.44)}{45.8}\ (\text{schoolfather}) - \underset{(1.79)}{104}\ (\text{father} = \text{manual laborer}) \\ & + \underset{(6.95)}{522}\ (\text{father} = \text{clerical worker}) + \underset{(1.47)}{18.8}\ (\text{schoolmother}) \\ & - \underset{(0.30)}{2.46}\ (\text{agemother})\end{aligned}$$

11. Households with a higher health endowment had significantly <u>lower</u> income (t=2.32) but contained fathers with marginally significantly higher schooling attainment (t=1.45). The schooling attainment of the mother was not statistically significantly related to the household health fixed effect.

12. Wolfe and Behrman (1983) suggest that estimates of income effects on child health may be misleading due to the existence of other family

endowments. Their data do not permit estimates of interfamily health heterogeneity. Our results (Table 3 and fn. 8) imply that estimated income effects on fertility obtained without controlling for health endowments would be negatively biased and those for maternal age at birth positively biased; the estimated income elasticity for food is not sensitive to health heterogeneity, however.

PART II: MIGRATION SELECTIVITY AND THE CONSEQUENCES OF GOVERNMENT PROGRAMS: ESTIMATING THE HEALTH AND FERTILITY EFFECTS OF THE PROMOTORA PROGRAM IN CANDELARIA, COLOMBIA

The use of cross-sectional data to estimate models of consumer or household behavior or to evaluate public programs has a long and continuing tradition in economics (e.g., Feldstein (1978), Meyer and Wise (1983), Johnson (1983), Pencavel (1984)). The treatment of cross-sectional or area-specific variation in prices or program characteristics as exogenous would appear, however, inconsistent with the Tiebout hypothesis (1956) that agents, heterogenous in preferences or endowments, locate in response to and/or select local program levels according to those preferences. If so, cross-sectional correlations between the observed behavior of agents and relative prices or program levels will not correspond to true price or program effects for any individual agent.

While some studies have shown that local laws reflect the preferences of local populations (Landes (1980), Farber (1984)), such studies appear to assume that interregional differences in population preferences are exogenous. Heterogeneity and selective-migration imply, however, that site-specific changes in prices or programs, whatever their source, will alter endogenously the characteristics and size of the population at the site, possibly inducing consequences unanticipated by the law-makers. Todaro's classic article (1969) presents a theoretical example in which non-selective migration thwarts the intended effects of an urban job creation cum minimum wage program. A local program altering relative prices, however, may also induce countervailing changes in the population via migration selectivity; e.g., a locality initiating a program subsidizing health care might attract low-health households. Lack of attention to selective migration thus makes inferences about the effectiveness of a program to be implemented nationally based on local program initiatives potentially misleading.

Despite the importance of location-choice selectivity in local public goods theory and in the evaluation of public programs, there have been few attempts to test directly for the existence of selective migration (an exception is Schultz (1983))or to test predictions for how migration responds selectively to changes in relative prices.[1] Yet, how the characteristics of agents change across activities or locations in response to relative prices, i.e., the selectivity rules, are clearly dual to the price-theoretic implications for the observable behavior of a given agent and thus are themselves subject to verification.[2] In this paper we consider how a price change or program subsidy that is location or site-specific affects the composition of residents via selective migration and biases evaluations of the effectiveness of the local program. In particular, we assess the consequences of a site-specific program subsidizing human capital investment in terms of shifts in both population composition and a representative household's resource allocations, when optimizing households that are heterogenous in preferences and in their endowments of human capital are free to choose locations in response to changes in location-specific prices.

In Section 1 the theoretical framework is described and implications are derived for how population preferences and endowments shift within a locality in response to the human capital subsidy. We show that under plausible assumptions and under all forms of heterogeneity a program subsidizing investments in children attracts high-income households with small families; children in such households, ceteris paribus, may exhibit low or high levels of human capital, however, depending on whether the principal source of heterogeneity is in tastes or endowments. The relationships between the biases in estimates of the program effects and sources of heterogeneity that arise from selective migration are also derived. In Section 2, unique longitudinal data from

Colombia describing the consequences of a local health subsidy program are used to test the implications of migration selectivity. The findings confirm the existence of selective migration. The child health care program evidently induced in-migration by households characterized by high income but, within income groups, households with low-fertility and low child health were attracted to the program site. These migration patterns are shown to be consistent with the hypothesis that heterogeneity in health endowments dominates that in tastes within the population. We also show that as a consequence of this form of heterogeneity, the effectiveness of the program based on cross-child differences in health and program exposure is considerably overestimated when selective migration is not taken into account.

1. Modeling Migrant Selectivity

a. Heterogeneity and Migration Selectivity: Who Migrates?

Consider an economy consisting of heterogenous households in spatial equilibrium: all potential profits from migration are zero; i.e., no household, net of migration costs, can increase its income by changing location. Decisions by households are characterized by the static, lifetime optimization problem in which the ith household maximizes

$$(1) \quad U(H^i, N^i, Z^i; \alpha^i),$$

where N^i = number of children in household i, H^i = average human capital of children in i, Z^i = composite consumption good and α is a vector of household-specific taste parameters, subject to the human capital production function

$$(2) \quad H^i = H(X^i; \beta^i, \mu^i),$$

where X^i is the per-child human capital input, β^i is the household's technology parameter, and μ^i is the household's human capital endowment, and the lifetime budget constraint

$$(3) \quad Y^i = P_x X^i + P_N N^i + Z^i,$$

where Y^i is income, P_x is the input price and P_N the price of a child.[3]

Assume that all prices are identical across all locations but that at a particular site a program is initiated which pays a subsidy s per unit of the human capital purchased input. Each household not at the program site now faces a potential migration decision. If the household migrates to the site, the price of the human capital input is P_X^S and lifetime income is Y^{S^i}, where Y^{S^i} is income at the program site net of the cost of migration and the program tax, which is assumed to be lump-sum and levied on all residents. Let $V(P_X, Y^i; \alpha^i, \beta^i, \mu^i)$ be the indirect utility function derived from maximizing (1) subject to (2) and (3); the difference in utility dV between migrating and not migrating is then:

$$(4) \quad dV^i = \frac{\partial V^i}{\partial P_X} dP_X + \frac{\partial V^i}{\partial Y^i} dY,$$

Roy's identity, i.e., $- X = \frac{\partial V/\partial P_X}{\partial V/\partial Y}$, yields the migration decision rule

$$(5) \quad \text{migrate iff} = X^i(P_X^S - P_X) + Y^{S^i} - Y^i > 0,$$

where $P_X^S < P_X$, $Y^S < Y$. If the subsidy is proportional to P_X ($P_X^S = (1-s)\ P_X$) and migration cum program costs are proportional to income ($Y^{S^i} = (1-C^i)\ Y^i$) this reduces to

$$(6) \quad \text{migrate iff } \frac{P_X X^i}{Y^i} > \frac{C^i}{s},$$

namely that the household migrates if and only if the income share of the

human capital input is at least as large as the ratio of the proportional migration cost to the proportional input subsidy. Notice that a secondary condition for migration to occur is that $C < s$. If the subsidy is absolute ($P_X^S = P_X - s$) and the migration cost is also absolute ($Y^{S^i} = Y^i - C^i$), then the decision rule is

$$(7) \quad \text{migrate iff } X^i > \frac{C^i}{s} .$$

Migration rule (7) can be used to derive the rules for migrant selectivity by allowing the fundamental taste, technology and endowment parameters to vary in response to changes in relative prices while holding constant the level of the utility differential between the origin and program sites. The characteristics of the "marginal" migrant household, the household that is just indifferent between moving to the site or not migrating, must change with migration costs C or the site subsidy s according to:

$$(8) \quad d\Gamma = \left(\frac{dX}{d\Gamma}\right)^{-1} \left(\frac{1}{s}\, dC - \frac{C}{s^2}\, ds\right), \qquad \Gamma = \alpha,\ \beta,\ \mu,\ Y$$

where superscripts are dropped to indicate that (8) describes a change in the type of household rather than the response of a given household.

Clearly, from (8), any characteristics of the household that increase the demand for the human capital input X must increase as C increases or must decrease as s increases in order to maintain the indifference. To discover how the observable characteristics of migrants vary with the program subsidy and/or migration costs it is thus necessary to specify how the unobservables α, μ and β affect household decision rules. Rosenzweig and Wolpin (1982) treated the special case where the human capital endowment is additive, i.e., $H = H(X;\ \beta) + \mu$, and derived general expressions for the influence of the endowment on the human capital input and on fertility, namely

$$(9)\quad \frac{dK^i}{d\mu^i} = \frac{P_X}{H^i_X}\left[-H^i_{XX}\left(\frac{dK^i}{dP_X}\right) + \frac{dK^i}{dY^i}\right] \qquad H^i_X > 0,\ H^i_{XX} < 0$$

$$K^i = X^i,\ N^i.$$

The effects of μ on the demand for X and N thus depend on the usual Hicks-Slutsky compensated price and income effects. If the fertility and the human capital of children are Hicks-Slutsky substitutes, as has often been found (e.g., Rosenzweig and Wolpin (1980)), and income effects are small relative to price effects, then a reduction in migration costs and/or an increase in the site subsidy would attract migrants with both higher human capital endowments and larger family size. If income effects dominate price effects, and are positive, then lower migration costs or a higher site subsidy induce, within income groups, less-endowed and lower-fertility in-migrants.

To generate predictions regarding the consequences of tastes and endowment heterogeneity for migration selection due to the initiation of a site-specific program subsidy requires that additional structure be imposed on the household problem. Consider a model in which the utility function is quadratic and the technology is linear for each household i:

$$(10)\quad U^i = \alpha^i_1 N^i - \alpha^i_2 {N^i}^2 + \alpha^i_3 X^i - \alpha^i_4 {X^i}^2 + \alpha^i_5 Z^i - \alpha^i_6 {Z^i}^2$$

and

$$(11)\quad H = \beta^i X^i + \mu^i,$$

which when solved in terms of the exogenous parameters yields the demand equations:

$$(12)\quad X^i = \frac{\psi^i(\alpha^i_3 - 2\mu^i\alpha^i_4) + (\alpha^i_2(2\alpha^i_6 Y^i - \alpha^i_5) - \alpha^i_1\alpha^i_6 P_N)}{\beta^i 2(\alpha_4\psi^i + \phi^i)}$$

$$(13)\quad N^i = \frac{\alpha^i_1 + P_N\,(2\alpha^i_6 Y^i - \alpha^i_5)}{2\psi^i} - \alpha^i_6\,\frac{P_N P_X}{\psi^i}\,X^i,$$

where $\psi^i = \beta^{i^2}\,(\alpha^i_2 + \alpha^i_6\,P_N^{\,2})$, $\phi^i = \alpha^i_2\alpha^i_6 P_X^{\,2} > 0$. Note that, in accord with most empirical findings, X and N vary inversely.

Assume that there is potential (unobserved) population heterogeneity in α_1 and α_3, reflecting preferences for family size and human capital, and in the endowments μ and β. Prior to the introduction of the program, the program-site and potential migrant populations are on average identical with respect to these fundamental parameters. As the subsidy is raised the <u>ceteris paribus</u> changes in these taste and endowment parameters and in income (Y) that will characterize the marginal migrant, from (8) and (12), are given by (14):

$$(14)\quad \begin{aligned}
\frac{d\alpha_1}{ds} &= \frac{C}{s^2}\left[\frac{2(\alpha_4\psi + \phi)}{\alpha_6 P_X P_N}\right] > 0\\
\frac{d\alpha_3}{ds} &= -\frac{C}{s^2\beta}\,[2\alpha_4 + \phi/\psi] < 0\\
\frac{d\mu}{ds} &= \frac{C}{s^2\beta}\,[1 + \phi/\psi\alpha_4] > 0\\
\frac{dY}{ds} &= -\frac{CP_X}{2}\,[1 + \psi\alpha_4/\phi] < 0
\end{aligned}$$

$$\frac{d\beta}{ds} = -\beta C\psi\,[(\alpha_3 - 2\mu\alpha_4)\,(\psi + \phi) + 2\beta\alpha_4 P_X\,(\alpha_1\alpha_6 P_N + \alpha_2(\alpha_5 - 2\alpha_6 Y))]\,[2s^2\alpha_4(\alpha_4\psi + \phi)^2]^-$$

Four polar cases with respect to heterogeneity are of interest: Suppose first that the populations are heterogeneous only with respect to tastes for family size and human capital investments in children. Specifically, let only α^i_1 or α^i_3 differ across households. From (14), as the subsidy increases, the

migrant population as it becomes less selective will be composed of increasingly higher α_1 and increasingly lower α_3 households. If tastes are the only source of heterogeneity, observationally identical migrant households will thus have fewer children who on average will be of higher "quality", since $dN/d\alpha_1$, $dX/d\alpha_3 > 0$ by construction, and

$$(15)\quad \begin{aligned} \frac{dN^i}{d\alpha_3^i} &= -\frac{\alpha_5 P_N P_X}{2\beta(\alpha_4\psi + \phi)} < 0 \\ \frac{dX^i}{d\alpha_1^i} &= -\frac{\alpha_6 P_X P_N}{2(\alpha_4\psi + \phi)} < 0 \end{aligned}$$

With heterogeneity in the additive health endowment μ only, increasing the subsidy and thus reducing selectivity will draw households with higher endowments, since more endowed households, for given money income, will have a lower demand for X:

$$(16)\quad \frac{dX^i}{d\mu^i} = -[\beta(1 + \phi/\psi\alpha_4)]^{-1} < 0.$$

In contrast to the tastes heterogeneity scenario, lower-μ households, despite their higher demand for human capital inputs, will always have children characterized by lower levels of human capital, as

$$(17)\quad \frac{dH^i}{d\mu^i} = 1 + \beta\frac{dX^i}{d\mu^i} = [1 + \psi\alpha_4/\phi]^{-1} > 0.$$

When there is heterogeneity in the additive endowment, migrants attracted by the program subsidy will thus be observed to have lower levels of human capital. They will also have, as in the first case, fewer children since, as indicated in (7), high-X households always migrate and, from (13), X and N vary inversely.

When heterogeneity exists solely in β, the return to the input X, in-migrant households may have either lower or higher levels of human capital (expression (14)). However, whether or not high-β (and thus high-H) or low-β (and

thus low-H) households demand higher levels of the X input and thus migrate to the program, such migrant households will have smaller families.[4] A human capital subsidy program will thus tend to attract and to serve disproportionally households within income groups characterized by low fertility whether heterogeneity exists in tastes or in human capital endowments. If the principal source of heterogeneity is in the latter, however, the program may attract, within income groups, households with lower levels of human capital, while tastes heterogeneity implies that the program will principally serve children already characterized by higher levels of human capital.

Independent of any heterogeneity in unobserved, fundamental parameters, however, if income effects are positive (as they are in the model), migrants will have relatively high income. In the third polar case of no heterogeneity, migrants will thus tend to have larger families and children with higher levels of human capital, as long as the program contains no means test provisions, but within income groups migrant and resident households will appear identical.

b. Program Effectiveness and Program Effects

Consider now the problem of evaluating a human capital subsidy program when the program is located at a specific site and migration is potentially self-selective. The average human capital h in a population of observationally identical migrants and residents at the program site is given by:

$$(18) \quad h = f_M h_M + f_R h_R,$$

where f_M and f_R are the relative proportions of (post-program) migrants and (pre-program) residents in the population respectively and h_M and h_R

are their average human capital levels. Note that h_M is the truncated mean of the non-site and site populations when such populations have the same distribution prior to the program. A change in the subsidy, s, will induce a change in the average human capital stock according to:

$$(19) \quad \frac{dh}{ds} = \gamma_R + \frac{df_M}{ds}(h_M - h_R) + f_M(\gamma_M - \gamma_R) + f_M \frac{dh_M}{ds},$$

where γ_R and γ_M are the respective average program subsidy effects on the levels of human capital in the resident and migrant populations; i.e., $-\beta dX/dP_X$, df_M/ds is the shift in the proportion of migrants in the population due to a change in program attractiveness, and dh_M/ds is the change in the mean human capital of the migrant population due to migrant selectivity; i.e., $(d\Gamma/ds)(dh/d\Gamma)$ from (8), where $\Gamma = \alpha, \mu, \beta$.

The total effect of a change in the locally-implemented program subsidy on the average human capital in the site population, given by (19), thus depends on (i) the direct effect of the subsidy on human capital investments by the original, resident population, (ii) the magnitude of the compositional change in the population via migration that is induced, weighted by the differential in mean human capital levels between the migrant and resident populations, (iii) the magnitude and sign of the difference in mean program effects in the two sub-populations, and (iv) the changes in the mean human capital of the migrants caused by the arrival of new, selectively drawn migrants, who will differ from those migrants already present. The average "effectiveness" of a program subsidy, the effect of the program if it were not site-selective (provided in all sites ("globally")) is given only by the first term in (19), if all pre-program site populations have the same mean characteristics or if the pre-program site population is representative. It is thus clear that the bias in the estimate of program effectiveness based on the program's

site-specific effects, given by the sum of the last three terms in (19), will depend on the source and magnitude of heterogeneity in the overall population and on household decision rules.

In the case in which taste heterogeneity dominates (variation in α_3), the expression for the bias, derived from the model described by (10) and (11), is:

$$(20)\quad \frac{dh}{ds} - \gamma_R = (\alpha_{3M} - \alpha_{3R})\ (\alpha_4 + \phi/\psi)^{-1}\ [\frac{df_M}{ds} + \frac{f_M}{P_X}\ (1 + \psi\alpha_4/\phi)^{-1}] - f_M \frac{\beta C}{s^2},$$

where α_{3M} and α_{3R} are the mean human capital taste parameters for migrants and residents. The selectivity equation (14) implies, as noted, that migrants will have higher average tastes for human capital ($\alpha_{3M} > \alpha_{3R}$). Thus the first term in (20) is positive, since a higher subsidy attracts more high-α_3 and thus high-H migrants ($df_M/ds > 0$). Moreover, the (high-α_3) migrants react more positively to a subsidy than do residents.[5] The last term in (20), i.e., $d\alpha_{3M}/dS$, from (14), multiplied by $dh_M/d\alpha_{3M}$, from (11) and (12), is negative, however, reflecting the <u>marginal</u> decrease in the selectivity of the migrant population associated with the higher generosity of the subsidy.[6] Since this last term is a second-order effect, (20) implies that the estimated effect of a site-specific subsidy on human capital will represent an upper bound estimate of the effectiveness of the same program applied globally when variation in preferences for human capital is the principal source of population heterogeneity. Due to tastes heterogeneity, locally-implemented human capital subsidy programs will thus appear more efficacious than they really are for the average or representative household.

When heterogeneity is confined exclusively to endowments, however, the bias derived from the model cannot be signed even when the source of heterogeneity is the additive endowment. The bias in that case is:

$$(21)\quad \frac{dh}{ds} - \gamma_R = (\mu_M - \mu_R)\ (1 + \psi\alpha_r/\phi)^{-1}\ [\frac{df_M}{ds} - \frac{f_M}{P_X}\ (\alpha_4 + \phi/\psi)^{-1}]$$

$$+ f_M \frac{\beta C\phi}{s^2\psi},$$

where μ_M and μ_R are the mean human capital endowments in the migrant and resident populations. In this case, as was shown, the subsidy attracts low-endowment (and low-health) migrants ($\mu_M < \mu_R$); however, the human capital investments by (less-endowed) migrants respond more strongly to the subsidy than do those by residents. The net contribution of the negative compositional change and the positive differential in subsidy effects to the program effect bias cannot be predicted. When endowment variation is predominant, then, no inferences about the globally applied program effect on human capital can be made from the estimates of the site-specific program effects, unless migration selectivity is taken into account. Therefore unlike in the previous case, when there is endowment heterogeneity, a human capital subsidy program could lower the average human capital in the population at the site in which it is implemented as a result of migration selectivity even if it augments the human capital of any randomly-selected household.

Similar expressions can be derived for the selectivity biases in the estimated effects of a human capital subsidy on family size. In both the tastes and additive endowment heterogeneity cases, selective migration will lead to a negative bias--the human capital subsidy attracts low-fertility

households in both cases, as was shown, and the negative response of fertility to the subsidy is stronger in both the low-μ and high-α_3 households, who make up the migrant population. Selective migration is likely to make a site-specific human capital subsidy program appear more anti-natalist than a similar but globally-applied program.

2. Empirical Application

a. Migrant Selectivity

To test the migration selection hypotheses requires a data set that at a minimum identifies migrants and residents at a specific site or sites and provides the characteristics of both groups before and after the implementation of and/or changes in a public program. In 1968, a program providing home-based preventive and maternal child health services was initiated in a small village in Colombia, Candelaria, and detailed longitudinal information was collected from 1968 to 1974 on the characteristics of parents and on the health of children aged less than six. All households present in Candelaria at any time during the seven-year period with a child under six years of age were included in the program, in which nurse-volunteers ("promotoras") visited each household approximately every two months.[7] Since Candelaria is a small village which serves in part as an "intermediate" stopover for many migrants from outlying rural areas to Cali, information on opportunities in Candelaria is disseminated relatively rapidly in outlying areas and in-migrants make up a significant proportion of the population. While the Candelaria data are thus unique in permitting identification of in-migrants and residents and in providing pre-program, baseline data on both migrants (at time of entry) and residents, there is

no information on the characteristics of migrants at origin or of the general origin populations. Thus, the health of the children of in-migrants to Candelaria can only be compared to Candelaria residents, although comparisons of the fertility of migrants with non-Candelaria, origin populations are feasible, given the availability of the 1973 Census of Colombia.[8] Pre-migration characteristics must also be estimated.

In order to test for the existence of as well as to characterize the source of migrant selectivity, if any, arising from the incentives created by the Candelaria program, we need to compare, within observationally-identical groups, the pre-program family size and some measure of the pre-program human capital of the children of migrants (households who came to Candelaria after 1969) and residents (households residing in Candelaria when the program was initiated in 1968). We use the age-standardized weight of children as a measure of human capital, since weight is the only health outcome collected in all years of the program. Because the standardization required is one that is independent of the program and relevant to the population studied, the average weight in 1968 of (resident) Candelaria children for each age-sex group is used as the standard; that is, the age-weight distribution in effect at the initiation date of the program. Since some ages (in months) were not represented in this group and others had relatively small sample sizes, a fourth-order polynomial regression of these mean age-specific weights for each sex group was used to smooth the base. A child's weight-for-age was thus defined as the ratio of the weight at his/her own age to the standard weight at that age.

Table 1 provides the sample characteristics of resident and migrant families. As hypothesized in the previous section, migrant households have higher (age-standardized) incomes on average, although the slope of the migrant age-income profile is less steep for migrants. Migrant families also

Table 1

Sample Characteristic: Resident and Migrant Families

Variable/Statistic	Resident Families Mean	Resident Families S.D.	Migrant Families Mean	Migrant Families S.D.
Pre-program children ever born	4.49	2.75	3.22	2.46
Pre-program mean child weight-for-age, 0-6	.994	.129	.997	.151
Income	959	371	1130	597
Income when entered program	720	304	1007	594
Log of income, father aged 20 years	6.53	.134	6.88	.319
Slope of log-income profile, father aged 20 years	.0198	.00857	.00705	.0154
Curvature of log-income profile ($x10^{-3}$)	-.791	.383	-.617	.430
Years of schooling-mother	2.54	1.52	2.66	1.63
Years of schooling-father	2.76	1.58	2.98	1.54
Age of mother	28.3	6.31	26.2	6.88
Age of father	34.6	8.97	32.1	8.38
Mean program exposure of children 0-6 (months)	18.4	6.32	10.8	7.36
Mean proportion of lifetime children exposed to program	.619	.213	.475	.295
Number of families	208		280	

had fewer children than resident families prior to entering the program and lower fertility than that for all rural households in Colombia (1973 Census of Colombia). While this differential conforms to the prediction of the theoretical analysis that low-fertility households would be most attracted to a program subsidizing human capital investments, whatever the principal source of heterogeneity, migrant parents are also on average younger, more educated and wealthier than resident parents. Differences in family size may thus be due to these differences in observed characteristics and tests for selectivity must be performed within observationally identical groups. Estimation of the household demand equations for fertility and health is thus required to investigate the sources of heterogeneity and selection.

As was demonstrated, the existence of heterogeneity in tastes, technology and/or endowments implies that all of the coefficients of household demand equations will be family-specific. This suggests the following estimating relationship:

$$(22) \quad K^i = X^i g^i + \varepsilon^i$$

where K^i is either the pre-program children ever born or (log) weight-for-age variable for family i, X^i is the set of exogenous characteristics of the household conditioning these choices, g^i is the family-specific parameter vector, and ε^i is a random term.

It is assumed that:

$$E\begin{bmatrix} g^i \\ \varepsilon^i \end{bmatrix} = \begin{bmatrix} \gamma \\ 0 \end{bmatrix}$$

$$(23) \quad V\begin{bmatrix} g^i \\ \varepsilon^i \end{bmatrix} = \begin{bmatrix} \Lambda & 0 \\ 0 & \sigma^2 \end{bmatrix}$$

where $$\Lambda = \begin{bmatrix} \sigma_1^2 & & & 0 \\ & \sigma_2^2 & & \\ & & \ddots & \\ 0 & & & \sigma_k^2 \end{bmatrix}$$

This is the standard random coefficients model; the estimating equation (22) may be written as

$$(24) \quad K^i = Y^i\gamma + \varepsilon^{i*}_j$$

where $\varepsilon^{i*} = (g^i - \gamma)\, X^i + \varepsilon^i$. Since the ε^{i*}'s are heteroscedastic, a generalized least squares (GLS) estimator will yield consistent and efficient estimates of the ρ's and σ's. Selectivity implies that the means of the family-specific parameters γ will differ across the migrant and resident populations.

The demand equations (24) must be modified to take into account the life-cycle nature of fertility and child health investment decisions, not incorporated, for simplicity, in the models of the previous section. First, health is a stock that is presumably a function not only of current inputs but of all inputs applied in the past, and current family size also reflects past fertility decisions. Thus, the reduced form health and fertility demand functionswill contain the determinants of all current and past inputs. Second, in a life-cycle context with perfect foresight, input demand, health demand and fertility decisions at any point in the life-cycle will depend on future, current, and past income and prices. Log income age profiles were thus estimated for residents and migrants separately using all reported income data points over the seven-year sampling period and information on occupation, age and schooling attainment.

For migrants, as noted, income prior to migration is unavailable. To the extent that there is an important structural shift in the income profile associated with migration, the profiles of migrants may be misrepresented. To ascertain if this absent information could account for any differences in parameter estimates obtained across the migrant and resident populations, we also estimate

a child health equation based on the subsample of children born after the program was begun and/or after migration to Candelaria. For this subsample, children of the same age do not differ with respect to their exposure to the program and information on actual family income is available for every year of the child's life whether the child is from a migrant or resident household.

The household log income profile is measured by three statistics: the constant in the log income equation evaluated at husband's age = 20, the first derivative evaluated at husband's age = 20 and the second derivative. These terms differ in the sample by the level of husband's education and husband's occupation. Mother's schooling is included in (24) as an observable characteristic that may shift tastes, technology or endowment parameters. Mother's age is also included in the children ever born equation as a life-cycle standardization; it is not included in the weight-for-age equation, since that is already appropriately age-standardized.

Table 2 presents the relevant random-coefficient GLS demand equation estimates, the first column for fertility (children ever born) prior to entry into the program and the second column for the pre-program (log of) mean child weight-for-age. The third column reports OLS estimates for the (log of) standardized child weight for children born after program entry. Only OLS estimates are reported for that subsample because the GLS procedure produced a large number of negative variance estimates of household-specific parameters. The reported t-values in column 3 may therefore be biased, although the coefficient estimates are consistent.

The three specifications reported include only an intercept dummy variable taking on the value of one for resident households in order to distinguish

Table 2

Migrant Selectivity, Family Size, and Child Health

Dependent Variable: Variable/ Estimation Procedure	Children Ever Born at Program Entry Random Coefficient-GLS	Log of Mean Child Weight for-Age at Program Entry[b] Random Coefficient-GLS	Post-program Children: Log of Mean Weight-for-Age[c] OLS
Resident	.565 (2.90)[a]	.0209 (1.10)[a]	.0462 (2.46)[d]
Mother's schooling	-.416 (3.60)	-.00505 (0.41)	-.00764 (0.64)
Income level (log)	.269 (0.30)	.155 (1.81)	.150 (1.82)
Income slope	-9.25 (0.42)	1.62 (0.80)	.986 (0.51)
Income slope derivative	-.639 (1.42)	-.814 (0.02)	5.97 (0.18)
Age of mother	.409 (11.6)	- -	- -
Age of mother squared	-.00185 (3.70)	- -	- -
Intercept	-7.09 (1.16)	-1.07 (1.85)	-1.03 (1.86)
R^2	.489	.037	.035
n	456	458	303

a Asymptotic t-ratios beneath regression coefficients.
b Children born before the family entered program.
c Children with full program exposure since birth; i.e., born after the family entered the program.
d t-ratios beneath regression coefficients.

migrants from residents. Regressions which are fully interactive with respect to the residence dummy, as are indicated by the theory, were also estimated, but are not reported since the overall story is unchanged with the more parsimonious specification. Most of the estimates of the individual interaction coefficients were not measured with much precision; however, F-tests reject at the five percent significance level the hypothesis that the migrant and resident pre-program demand equations are identical. Migrant selectivity is indicated.

The set of resident dummy coefficients reported in Table 2 conform to the scenario in which endowment heterogeneity dominates tastes heterogeneity. Within income/schooling groups, migrants to Candelaria had both lower pre-program family size and children with lower age-standardized weight upon entry and after compared to residents. These findings thus suggest that the immigrants drawn to Candelaria were selected not only from the upper tail of the income distribution (Table 1), but, within income groups, were self-selected from the lower tail of the endowment (μ) distribution.[9] Evidently, the slightly higher child weight observed for migrants at entry in Table 1 is due to the higher household income of migrants; the estimated income level coefficients in columns 2 and 3 confirm that higher income households value health human capital more highly (one-tail test, five percent level).

b. Evaluation of the Effectiveness of the Candelaria Program

The single-site sample design of the Candelaria data set would appear to preclude any evaluation of the promotora program, since all households face the same subsidy. However, in a life-cycle context the total subsidy varies across children to the extent that children of the same age were exposed to the subsidy for different lengths of time, a greater number of

health inputs being subsidized for children within the same age group but exposed earlier to the program. There are thus two sources of variations in the total program subsidy: First, since the dissemination of information about the Candelaria program to outlying areas and migration itself takes time, children of migrants, while facing the same subsidy as the children of residents when they arrive, will not be exposed to the program for the same length of time, given their ages, as resident children. This differential is evident in Table 1; mean months of program exposure for migrant children is less than 60 percent that of resident children. Since our results indicated that migrant children have lower health (due to selection), use of the cross-child variation in program exposure to assess the impact of the program without attention to migrant selectivity would appear to result in an upward bias in the estimate of program effectiveness. However, we also showed that low-μ households may respond more positively to a human capital subsidy; the direction of the selectivity bias in the program exposure estimate is thus ambiguous.

The second source of variation in program exposure arises from variation in the birth dates of resident children who were born prior to the program (1968). For such children, the sample would appear to approximate an experimental design as long as the program was unanticipated. However, since children born at later dates on average are born later in their parents' life-cycle, the cross-sectional variation in program exposure among resident children may also be correlated with their health endowments or with parental preferences for health, if these characteristics also influence the timing and spacing of births.[10] A relationship between program exposure and health might thus exist in the absence of any true program effect even among children of residents.[11]

To assess the consequences of migration selectivity for estimating the effectiveness of the Candelaria program thus requires attention to both sources of potential bias arising from the use of exposure information to measure subsidy differentials. The longitudinal data on the health of individual children permits this separation. Consider the reduced form estimating equation for child-specific health:

$$(25) \quad H^{ij} = \gamma^{ij} E^{ij} + \mu^{j} + \varepsilon^{ij},$$

where H^{ij} is the health of child j in family i, E^{ij} is the length of program exposure, γ^{ij} is a random coefficient on exposure, μ^{j} is a child-specific health endowment, and ε^{ij} is a random time-varying health component. All other family characteristics are suppressed for simplicity. Population heterogeneity implies that γ^{ij} differs across children if health endowments differ, since, as we have shown, the effect of s on the demand for the human capital input X depends upon fundamental parameters. Program exposure, which depends on the child's date of birth and/or on the timing of migration, may be correlated with the unobserved health endowment as a result of timing and spacing decisions and migration selectivity. With multiple observations for each child, however, a random coefficients fixed effect estimator can be used to purge out the family and child-specific health endowment.[12] Rewriting (25) in differential form yields

$$(26) \quad \Delta H^{ij} = \gamma^{ij} \Delta E^{ij} + \Delta\varepsilon^{ij}.$$

GLS estimation of equation (24) provides a consistent estimate of the program exposure effect for the sample of <u>resident</u>-household children. However, if migrant-household children are included in the sample, the distribution of the γ_{ij}'s will be truncated when migrant households are not randomly

drawn. This leads to the standard sample selection problem, since $E[(\gamma^{ij} - \gamma)\ \Delta E^{ij} + \Delta\varepsilon^{ij} | \text{migrant}] \neq 0$.

To obtain a consistent estimate of the program exposure effect, and thus the effectiveness of the program, it is thus necessary to restrict attention to resident children born prior to the program or to attempt to correct for the sample (migration) selection. We refrain from employing one of the standard sample selection correction procedures (Heckman, 1979; Olsen, 1980) since that would entail imposing further, and arbitrary, structure on the problem. The resident subsample should be large enough to permit precise estimation of the program exposure effect and thus an assessment of bias due to selective migration.

Table 3 reports estimates of the program exposure effects for both the full sample of children (inclusive of migrants) and the sample of resident children. Estimates from two specifications are reported, a linear specification in which the exposure effect is assumed to be identical across education/income groups and an interactive specification, which allows exposure effects to differ by parental characteristics, as is consistent with our linear-quadratic example.[13] Both specifications are estimated using ordinary least squares and the GLS random-coefficient fixed effect estimator FE-RC). For both samples and both estimation procedures, however, F-tests reject the linear specifications; Table 4 reports the per-month exposure effects on standardized weight by income levels implied by the interactive estimates.

Comparisons across samples and across estimates permit an assessment of the separate roles of migrant selectivity and within-group heterogeneity. Whatever the specification or estimation procedure, however, estimates

Table 3

Migrant Selectivity and Program Exposure Effects on Log of Child Weight-for-Age

Sample:	Migrants + Residents				Residents			
Estimation Procedure	OLS	OLS	F.E.	F.E.	OLS	OLS	F.E.-R.C.	F.E.-R.C.
Exposure (months x 10^{-2})	.116 (10.58)[a]	-1.59 (1.95)[a]	.408 (37.63)[a]	-1.67 (2.02)[a]	.0736 (5.56)[a]	-4.99 (0.43)[a]	.000786 (0.11)[b]	3.57 (5.36)[b]
Log of income, father at age 20	.136 (7.77)	.0977 (3.72)	-	-	.118 (4.38)	.104 (2.45)	-	-
Income slope	1.69 (3.95)	1.30 (1.98)	-	-	2.21 (3.90)	1.58 (1.77)	-	-
Income curvature	33.7 (4.54)	41.9 (3.67)	-	-	47.9 (4.86)	68.0 (4.30)	-	-
Schooling mother: more than one standard deviation below the mean (1)	.00227 (0.34)	.00509 (0.51)	-	-	-.00184 (0.20)	-.00593 (0.40)	-	-
Schooling: within one standard deviation below the mean (2)	.00365 (0.38)	-.00711 (0.50)	-	-	-.0107 (0.80)	-.0110 (0.52)	-	-
Schooling: (3) within one standard deviation above the mean (2)	.0285 (0.21)	-.0172 (0.87)	-	-	.00376 (0.20)	-.0290 (0.99)	-	-
Income × exposure	-	.00241 (2.13)	-	.00296 (2.60)	-	.000572 (0.35)	-	-.00487 (5.24)
Slope × exposure	-	.0216 (0.76)	-	.0235 (0.83)	-	.0296 (0.87)	-	.0208 (1.07)
Curvature × exposure	-	-.549 (1.15)	-	-.915 (1.96)	-	-.991 (1.61)	-	3.23 (13.0)
Schooling (1) × exposure ($\times 10^{-3}$)	-	-.103 (0.25)	-	-.377 (0.90)	-	.150 (0.27)	-	-1.55 (4.27)
Schooling (2) × exposure ($\times 10^{-3}$)	-	.682 (1.14)	-	.485 (0.80)	-	1.02 (1.29)	-	-2.42 (4.86)
Schooling (3) × exposure ($\times 10^{-3}$)	-	.124 (1.48)	-	.230 (0.27)	-	1.55 (1.42)	-	-2.57 (4.10)
Intercept	-1.02 (8.00)	-.657 (3.48)	-	-	-.804 (4.16)	-.669 (2.20)	-	-
R^2	.036	.033	.188	.193	.021	.023	.0001	.031
n	7583	7583	6126	6126	4540	4540	1877	1877

a Absolute values of t-ratios beneath regression coefficients.
b Absolute values of asymptotic t-ratios beneath regression coefficients.

Table 4

Exposure Effects by Income Level: Percent Change in Standardized Weight per Month[a]

Exposures	Migrants + Residents		Residents	
Estimation Procedures	OLS	Fixed Effect	OLS	Fixed Effect
Two σ above the mean	.0081	.376	.015	.015
One σ above the mean	.0056	.337	.0071	.080
Mean	-.057	.298	.0005	.145
One σ below the mean	-.089	.258	-.008	.211
Two σ below the mean	-.121	.218	-.016	.276

a Evaluated at mean mother's education.

from the sample including both migrants and residents greatly overstate on average the health consequences of program exposure net of migration selection effects, with the greatest differential displayed by the fixed effect estimates.[14] Moreover, the patterns of exposure effects by income revealed by the fixed effect estimates taken from the resident and full samples are quite different (Table 4) -- the per-month exposure effects decline by income group in the full sample but increase with income in the resident sample.[15]

The fixed effect random coefficients model estimated in the residents-only sample, which is presumably free of selection and heterogeneity biases, indicates that a child exposed for 1.5 years to the Candelaria program (the sample mean for residents) and who lives in a household with an income level two standard deviations below the Candelaria mean would experience a five percent gain in weight-for-age; a similarly-exposed child from a household with income at the mean would experience a 2.6 percent increase in weight-for-age, while children from households with incomes more than two standard deviations above the mean would benefit little from the program.[16] The comparable full sample estimates imply that migrant selectivity leads to an overestimate of the program exposure effect by 2400 percent among households with incomes at least two standard deviations above the mean, and a 106 percent overestimate at the mean, while the exposure effect is understated by 21 percent among households with incomes less than two standard deviations below the mean. The program thus appears to have benefited most the children of poor residents and wealthy migrants, and to have attracted, among wealthy potential migrants, those who benefit most from the program and, among poor potential migrants, those who benefit least.

3. Conclusion

When agents are heterogeneous, a change in relative prices within a discrete geographical area or activity has two distinct effects. It alters the allocation of resources by each agent facing the price change and changes the composition of agents within the location or activity. While most empirical studies have been concerned with testing the allocative responses of a representative agent to changes in incentives, the change in the spatial distribution of differentiated agents in response to area-specific conditions, the central implication of the Tiebout hypothesis, has received little theoretical development or empirical verification. In this paper we have explored the consequences of a site-specific program subsidizing human capital investments in children for both the spatial distribution of heterogenous households and for the level of human capital investment by a representative household. We show that with plausible restrictions on the optimizing behavior of each household, such a program precipitates in-migration by high-income and low-fertility households, whether the principal form of heterogeneity is in tastes or in human capital endowments. With endowment heterogeneity dominant, however, households also characterized by low levels of human capital and/or with smaller returns to investments in human capital are attracted to a program subsidizing human capital investment.

Data from a village in Colombia that implemented a subsidized health program confirm these implications of selective migration--in-migrants were evidently drawn from the low-tail of the family size distribution, were of relatively high income and, within income groups, had children whose

nutritional status was lower than that of observationally identical members of the resident population. As a consequence, evaluations of the program inattentive to migration selectivity based on differences in program exposure across children born prior to the program were shown to significantly overestimate the impact of the program for any randomly-drawn household. Program evaluations based on comparisons of the mean nutritional status of children born after the program with that of children born prior to the program and never exposed to it, however, would seriously understate the effectiveness of the program due to the selective migration of low-endowment households. Indeed, the empirical results suggest that it is possible that the equilibrium mean health of children in the village, due to migration, will be lower after than before the health program, with mean health levels increased in areas external to the program site.

Our empirical results suggest that in a country such as Colombia or the United States where the population is highly mobile, tests of theories of the behavior of individual agents based on cross-sectional data or studies of the determinants or consequences of laws based on the exogenous spatial distribution of population characteristics may be seriously flawed. The existence of migration selectivity has implications beyond those relevant to the estimation of behavioral models from cross-sectional data or to the evaluation of location-specific programs, however. Consider a national immigration policy, for example, that does not discriminate by an immigrant's country of origin. Due to differences in migration costs (distance) and in relative prices across potential sending countries, immigration will be differentially self-selective across country-of-origin groups. Such differential selectivity will result in the observed behavior of immigrants

being correlated with their country-of-origin (Chiswick (1978)) even if the distributions of population characteristics in sending countries are identical. Moreover, our framework implies that an overall increase in barriers to immigration makes immigration more selective. How immigrants differ by country-of-origin or whether the laissez-faire selectivity arising from the decision-rules of optimizing potential immigrants is superior or inferior to selection imposed by law depend on the sources of heterogeneity and the nature of the relative price differentials across the sending and receiving countries. If immigrants are principally attracted by a country's superior opportunities for human capital investment, for example, our results imply that immigrants may be drawn from the lower tails of the human capital endowment distribution; however, the less so the smaller the direct costs of immigrating.

Finally, selective migration is a component of a broader class of problems. For example, as for migration, relative price changes as well as income growth may selectively alter fertility decisions, resulting in a change in the distribution of children across households of differing endowments and preferences for human capital investment, and thus in the endowments of the representative child. The long-term consequences of national programs may thus differ significantly from their immediate effects due to selectivity in fertility decisions. The further study of selection rules would appear warranted.

Footnotes

1. Gramlich and Rubinfeld (1982) test and confirm an implication of selective migration, that the variance in voter preferences for local public goods expenditures are smaller within than among urban localities, but do not test whether and/or how the local electoral outcomes affect residential mobility. Schultz (1983) tests for and confirms differences in behavior between migrants and residents in Colombia by origin and destination but does not incorporate migration decisions within his behavioral model or derive predictions for how interarea price differences generate selectivity rules.

2. While heterogeneity and selection, combined with information asymmetries, form the basis for many models of behavioral phenomena (e.g., Spence (1973); Guasch and Weiss (1981)), that literature has seen little empirical application. Heterogeneity and selection are also explicitly recognized in most econometric studies of labor supply behavior; however, selectivity is essentially treated as a nuisance rather than as a testable implication of the theory (an exception is Heckman (1974)).

3. The budget constraint ignores the interaction between the human capital of children and the number of children, as in Becker and Lewis (1973). Use of the non-linear budget constraint does not alter any of the testable implications of the model (Rosenzweig and Wolpin (1980)).

4. It is easy to demonstrate that households with higher returns to the human capital input X will always have higher levels of human capital. However, the higher return induces both an income effect, lowering the demand for X so as to allocate the higher wealth to the increased consumption of other goods, and a price effect, which raises the demand for

X. Note that the additive endowment only carries with it an income effect; higher μ households are wealthier and have healther children but do not obtain more human capital per unit of the input.

5. The subsidy effect on health, from (12), is

$$dh^i/ds = \beta[2\psi(\alpha^i_3 - 2\mu^i\alpha_4)\ \alpha_2\alpha_6 P_X \beta^{-1}\alpha_4^{-1} - (\phi - \beta\alpha_4\psi)\ (\alpha_2\alpha_5 - 2\alpha_2\alpha_6 Y + \alpha_1\alpha_6 P_N)]\ [2(\beta\alpha_4\psi + \phi)]^{-1},$$

from which subsidy-effect differentials can be computed when α and μ differ.

6. $dh_M/ds = (d\alpha_{3M}/ds)\ (dh_M/d\alpha_{3M}) = \beta C/s^2$.

7. The program was funded by the U.S. Agency for International Development; Candelaria residents and in-migrants thus did not incur any direct program costs.

8. Various health programs had been in operation in Candelaria before, but not after, the implementation of the "promotora" program. As a consequence, rates of malnourishment and fertility in Candelaria were lower than in the overall population in Colombia prior to 1968 (Heller and Drake (1979). Since in our sample only post-1968 migrants can be identified, recent but pre-1968 migrants attracted to the prior health programs will be counted as residents; differences between residents and post-1968 migrants will thus underestimate the selectivity induced by a health subsidy program.

9. As noted above, the existence of pre-1968 programs minimizes the estimated health differential between the post-1968 migrants and the pre-1968 residents in Candelaria since some proportion of the latter were attracted by the prior health programs. However, while the estimates in Table 2 are thus lower bound estimates of the migrant-resident health differential, it is possible that migrants are not drawn from the lower tail of the

health endowment distribution characterizing the non-Candelaria origin populations, since these populations exhibit lower mean health than do the residents in Candelaria. The Colombian Census data indicate that post-1968 migrants to Candelaria do exhibit lower fertility than in the rural population as a whole as well as in the Candelaria resident population.

10. That spacing patterns are related to household health endowments in Candelaria households is shown in Rosenzweig and Wolpin (1984).

11. Heller and Drake (1979) exploit differences in program exposure among children to evaluate the effectiveness of the Candelaria program. They ignore the selectivity associated with both migration and parental spacing decisions. Their specifications estimate program exposure effects conditional on such endogenous variables as parental breast-feeding, use of medical services, and food expenditures; their findings thus cannot be compared with our reduced-form estimates.

12. A within-family (cross-child) fixed effect estimator would not provide consistent estimates if endowments differ among children and the spacing of children responds to realizations of child-specific endowments. In that case, a child's health outcome will affect the interval to the next child so that the difference in program exposure between children within the same family will be related to their relative health as part of the family's optimization process and regardless of the program. Evidence on these dynamic spacing patterns is presented in Rosenzweig and Wolpin (1984).

13. The subsidy effect, given in footnote 5, depends on fundamental parameters as well as on income.

14. Similar results are obtained when the fraction of the child's life during which he or she is exposed to the program is used to measure exposure. With either measure, identification of the program effect arises from the nonlinear relationships among age, exposure and weight induced by the age/sex standardization.

15. These differential program effects by income level are not due to program design but reflect the nonlinearity in income of the input demand equation.

16. Part of the impact of the program appears to work via encouraging greater and/or more rapid investments in children. Fixed effect (logit) estimates indicate that children of the same age but exposed longer to the program were more likely to be receiving breastmilk and to have received innoculations against diptheria, polio or tetanus.

References

Becker, Gary S. and H. Gregg Lewis, "On the Interaction Between Quantity and Quality of Children," Journal of Political Economy 82, April/May, 1979, S279-S288.

Chiswick, Barry R., "The Effects of Americanization on the Earnings of Foreign-Born Men," Journal of Political Economy 86, October 1978, 897-921.

Farber, Henry S., "Right-to-Work Laws and the Extent of Unionization," Journal of Labor Economics 2, July 1984, 319-352.

Feldstein, Martin, "The Effect of Unemployment Insurance on Temporary Layoff Unemployment," American Economic Review 68, December 1978, 834-846.

Guasch, J. L. and Andrew Weiss, "Self Selection in the Labor Market," American Economic Review 71, June 1981, 275-284.

Gramlich, Edward M. and Daniel L. Rubinfeld, "Micro Estimates of Public Spending Demand Functions and Tests of the Tiebout and Median-Voter Hypotheses," Journal of Political Economy 90, June 1982, 536-560.

Heckman, James J., "Sample Selection Bias as a Specification Error," Econometrica 47, April 1979, 153-161.

_____, "Shadow Prices, Market Wages and Labor Supply," Econometrica 42, July 1974, 679-694.

Heller, Peter S. and William D. Drake, "Malnutrition, Child Morbidity and the Family Decision Process," Journal of Development Economics 6, June 1979, 203-235.

Johnson, Janet L., "Sex Differentials in Unemployment Rates: A Case for No Concern," Journal of Political Economy 91, April 1983, 293-303.

Landes, Elisabeth M., "The Effect of State Maximum-Hours Laws on the Employment of Women," Journal of Political Economy 28, June 1980, 476-484.

Meyer, Robert H. and David A. Wise, "The Effects of the Minimum Wage on the Employment and Earnings of Youth," Journal of Labor Economics 1, January 1983, 66-100.

Olsen, Randall J., "A Least Squares Correction for Selectivity Bias," Econometrica 48, November 1980, 1815-1820.

Pencavel, John H., "The Tradeoff Between Wages and Employment in Trade Union Objectives," Quarterly Journal of Economics 99, May 1984, 215-232.

Rosenzweig, Mark R. and Kenneth I. Wolpin, "Heterogeneity, Intrafamily Distribution and Child Health," mimeograph, University of Minnesota, April, 1984.

_____, "Governmental Interventions and Household Behavior in a Developing Country," Journal of Development Economics 10, April 1982, 209-225.

_____, "Testing the Quantity-Quality Fertility Model: The Use of Twins as a Natural Experiment," Econometrica 48, February 1980, 227-240.

Schultz, T. Paul, "Migrant Behavior and the Effects of Regional Prices: Aspects of Migrant Selection in Colombia," Economic Growth Center Discussion Paper No. 433, August, 1983.

Spence, A. M., Market Signalling: Information Transfer in Hiring and Related Processes, Cambridge: Harvard University Press, 1975.

Tiebout, Charles M., "A Pure Theory of Local Expenditures," Journal of Political Economy 65, October, 1965, 416-424.

Todaro, Michael, "A Model of Labor Migration and Urban Unemployment in Less Developed Countries," American Economic Review 59, March 1969, 138-148.

PART III. OPTIMAL GOVERNMENTAL INTERVENTIONS AND THEIR CONSEQUENCES: EVALUATING THE HEALTH EFFECTS OF FAMILY PLANNING AND HEALTH CLINICS IN LAGUNA, PHILIPPINES

Considerable public resources in both high and low-income countries are devoted to the subsidization of fertility control and health investments. The effects of these programs are thus of some concern, and social scientists have devoted attention to the evaluation of these programs. Most evaluation studies (e.g., Hermalin (1972), Khan and Sirageldin (1979), Rosenzweig and Schultz (1981), Rosenzweig and Wolpin (1982)) have essentially compared the variation in measures of the intensity of proram effort across localities with the corresponding inter-area variation in fertility and health. Little or no attention has been paid to the causes of the cross-area variability in the levels of such programs. Yet, if the allocation of public health and family planning services or subventions across localities is systematically related to factors determining fertility and health outcomes that are known to subsidy providers but unobserved by researchers, such cross-sectional estimates will produce misleading conclusions about program effectiveness.

Interest has grown recently in incorporating endogenous public resource allocations within models concerned with private agent behavior. Empirical applications or tests of such models, however, have been scarce and have been principally concentrated in the area of agricultural policy (Guttman (1978), Huffman and Miranowski (1981), Huffman and McNulty (forthcoming)). Moreover, existing general economic theories of public allocations do not provide much guidance for predicting how publicly-financed human capital subsidies and, in particular, "family planning" subsidies are distributed among heterogenous recipients. Altruism theories of public transfers (Hochman and Rodgers (1969), Roberts (1984)) would appear to suggest that the least-endowed receive the highest transfers, but such models

provide no rationale for the use of subsidies to particular goods such as contraceptives. Pressure-group theory (Becker (1982)) suggests that groups that are (1) relatively small in number, (ii) have and can command resources for lobbying and (iii) derive the greatest benefits from public transfers or interventions will receive the highest transfers. This model would appear to imply that the rich--small in number and with greater resources--rather than the poor would receive the highest fertility control subsidies, since, assuming that the poor have the largest families, they avert less births than the rich and thus benefit least from subsidies to fertility control.

While the existence of externalities from health (infection) might provide a Pigovian (and pressure-group) rationale for the subvention of health investments among the low-health poor, the empirical and theoretical rationale for fertility control subsidization based on the existence of direct, population externalities is less clear (Eckstein and Wolpin (1984)). Moreover, since a birth from any source contributes equally to population growth, the existence of population externalities (e.g., congestion) do not obviously provide a basis for selective subsidization of households by income or human capital endowments. A model of health and family planning subsidies is needed.

In this paper we formulate and test an optimizing model determining the distribution of family planning and health subsidies across heterogenous households and assess the biases in cross-area estimates of the health effects of such subsidies due to public resource optimization. The model incorporates different features of the general theories of public allocations: the welfare of "donor" households are directly

but asymmetrically affected by the behavior of recipient households, as in altruism models; thus, recipients of subventions have an intrinsic advantage in obtaining resource transfers, as implied by competitive interest group models, but the distribution of subsidies is a function of price effects as well as recipient characteristics, as in optimal taxation models. The model is used to derive rules for the distribution of both health and fertility control subsidies and to discern the conditions under which subsidies to fertility control alone or in combination with health subsidies are optimal (Pareto-improving), without resort to arbitrary specifications of population externalities or altruistic concern for family size.

In section 1, the model is set out. It is shown that when there is a health externality, subsidization of fertility control can substitute pervasively for a health subsidy if and only if family size and health are gross substitutes. It is then shown that when only family planning subsidies are provided, such subsidies are likely to be distributed disproportionately to low-health households, resulting in underestimation of the health effects of family planning programs from cross-section data. In the next section, it is demonstrated that a combination of family planning and health subsidies are Pareto-improving in the presence of the health externality even when cross-price effects and fertility externalities are absent. It is then shown that the distribution of the two subsidies will be positively correlated across areas (even in the absence of administrative scale economics), but it is impossible to establish whether such subsidies, when jointly provided, will favor the least or most endowed of recipient households.

In section 2, longitudinal data describing child health and publicly-provided family planning and health programs in 20 barrios in Laguna Province in the Philippines are used to estimate the effects of such programs on child health and the relationships between the distribution of the programs and pre-program health levels, i.e., the governmental allocation rules. The results are consistent with the model: (i) dates of family planning and health program initiation across barrios are positively correlated, (ii) family size and health are gross substitutes among households and in some barrios family planning programs but not health programs are present; (iii) both programs were initiated earliest in the low-health barrios and (iv) as a consequence, the positive and significant child health effects of both the family planning and health programs are completely obscured when no account is taken of the systematic associations between program placement and areal health endowments.

1. Modeling the Distribution of Health and Family Planning Subsidies

a. Evaluating Subsidy Effects on Health Production Among Heterogeneous Households

Consider a set of T low-income households each residing in a different health environment. Each household i chooses a level of health for its children H^i, its family size N^i and its consumption Z^i solving the following problem:

$$(1) \qquad \max U^i = U(H^i, N^i, Z^i),$$

where health production is described by the function

$$(2) \qquad H^i = h(X^i, N^i) + \mu^i, \qquad h^i_x > 0, h^i_{xx}, h_{NN} < 0,$$

subject to the full income constraint

$$(3) \qquad F^i = p_N N^i + (p_c - s^i_c)(\nu^i - N^i) + (p_x - s^i_x) X^i N^i + p_z Z^i,$$

where X^i = per-child health input, μ^i = exogenous, health parameter or endowment, F^i = full income net of taxes, if any; ν^i = potential fertility in the absence of fertility control; p_K = price of good K, K = N, X, Z; p_c = cost of fertility control or averted births, and s^i_c and s^i_x are per-unit subsidies to fertility control and health inputs, respectively, provided in each health environment by a central "government" or donor.

The solution for each household's average per-child health net of the environmental effect in terms of the exogenous variables unique to it is

$$(4) \qquad H^i = H(s_x^i, s_c^i, F^i, \mu^i).$$

Estimation of (4) to obtain the average effect of the subsidy s_j^i on child health when μ^i is unobserved yields the estimate:

$$(5) \qquad \frac{dH^i}{ds_j^i} = h_x^i \frac{dX^i}{ds_j^i} + h_N^i \frac{dN^i}{ds_j^i} + \left(\frac{ds_j^i}{d\mu^i}\right)^{-1}, \qquad j = x, c,$$

The true effect of a change in the subsidy s_j on the health outcome is given by the first two terms in (5): the subsidy (price) effects on the health input provided to each child and on family size weighted by their respective marginal health effects, from (2). The third term in (5) is the bias which arises when the μ^i are unobserved by the researcher and vary with the subsidy rates. Only if the subsidies are distributed independently of the μ^i, or, more generally, of any of the parameters unique to each area which influence health investments, will the association between the subsidies and health net of observed variables provide unbiased estimates of subsidy effects.

The sign of the bias in (5) will obviously depend on the allocation rules used by the agents who distribute the subsidies. If such agents follow a compensatory rule, for example, providing higher subsidies to less-endowed areas, then the subsidy effects obtained from (4) estimated, say, by least squares will understate the true consequences of increasing the subsidy for any randomly-chosen household; if such subsidies go to the better-endowed, their effect will be overestimated. Consideration of the possibly systematic association between subsidies provided to agents and the environmental or other characteristics of the agents in the estimation of subsidy effects clearly yields better (policy-relevant) estimates of those effects. Moreover, such estimates also permit the testing of models

of governmental resource allocations, which should provide the rules by which public resources are distributed among heterogenous agents or localities, as well as a rationale for the particular set of instruments used to effect resource transfers.

b. The Optimality of Family Planning Subsidies and their Distribution

To discern the rules by which subsidies to fertility control might be distributed among households behaving as described above in the absence of any arbitrarily-assumed direct population externality, consider a wealthy household having the same objective function as in (1) but facing a health externality. In particular, let the technology of health production for the well-off household be

$$(6) \qquad H = h(X, N, H^*) \qquad h_x, h_N, h_H > 0,$$

where H^* is the mean health of the children in the T low-income households, i.e.,

$$(7) \qquad H^* = \sum_{i=1}^{T} \alpha^i H^i$$

where $\alpha^i = N^i/(\sum_{i=1}^{T} N^i)^{-1}$ and max $(H^i) < H$. Thus, while the health of the well-off children depends on the mean health of the children of the poor, there is no direct fertility externality. As in altruism models, the externality is asymmetric--poor households do not consider or are not affected by the consumption set of the "donor" household.

Assume that the wealthy household can observe all the health endowments but, initially, cannot subsidize health investments and cannot tax fertility directly ($s^i \geq 0$). If each household's fertility control is differentially

subsidized, the budget constraint for the wealthy household is:

$$(8) \qquad G = P_N N + P_c (\nu - N) + P_x XN + P_z Z + \sum_i s_c^i (1 + \theta_c^i)(\nu^i - N^i),$$

where G = full income of the high-income household and θ_c^i is the loss in subsidy transfers to the ith household associated with transaction costs (waste, graft). In this setup, the transfer scheme is politically feasible, since the majority of households (the poor) and possibly all households are potentially better off. The questions are: 1) under what conditions will the wealthy household subsidize fertility control and 2) how will the subsidy, if warranted, be distributed among the poor households.

Maximization of the wealthy-household utility function subject to (8), (9) and the price-taking behavior of the poor households, as described by model (1) through (3), yields the equilibrium conditions:

$$(9) \qquad \frac{s_c^i}{P_c} = -\alpha^i \frac{P_x}{P_c} \frac{NX}{N^i} \frac{\varepsilon_{HH*}}{\varepsilon_{HX}} \left[\varepsilon_{HX}^i \frac{\eta_{XP_c}^i}{\eta_{NP_c}^i} + \varepsilon_{HH}^i + \left(\frac{H^i - H*}{H*}\right)\right] -$$

$$(\eta_{NP_c}^i)^{-1} \left(\frac{\nu^i - N^i}{N^i}\right)(1 + \theta_c^i) > 0 \qquad i = 1 \ldots T$$

$$\text{or } \frac{s_c^i}{P_c} = 0.$$

where $\varepsilon_{HH*} = (\partial H/\partial H*)(H*/H)$ and $\varepsilon_{HX} = (\partial H/\partial X)(X/H)$, from (7); $\varepsilon_{HX}^i = (\partial H^i/\partial X^i)(X^i/H^i)$ and $\varepsilon_{HN}^i = (\partial H^i/\partial N^i)(N^i/H^i)$ from (2); and the η^i are the demand price elasticities characterizing the ith household.

Condition (9) has two terms. The first contains the health gains to the rich household associated with increasing the fertility control subsidy. There are three sources of gains: the first term in brackets is the health return which occurs because of cross price effects. Raising the fertility control subsidy increases health

if fertility and child health are gross substitutes, since $\eta^{i}_{Np_c} > 0$. The second term is the return due to the direct or biological effect of family size on child health, through (2), in poor households. This term is positive if decreases in family size biologically augment child health. The third bracketed term is the "eradication" effect of fertility control subsidies--decreases in the size of families with below-average child health ($H^i < H^*$) increase the mean health of the poor households; family planning subsidies provided to the lowest-health households thus increase the health of the wealthy households via the health externality even if fertility and child health are independent (in terms of price or biological effects) in poor households.

The second term in condition (9) is the marginal cost incurred by the wealthy from increasing the fertility control subsidy to household i. Such costs are higher the greater the number of averted births (the lower is fertility), the smaller the own price elasticity of fertility, and the higher are transaction costs.

Condition (9) indicates that fertility control subsidies can be used as substitutes for health subsidies for all poor households when there is a health but not a population externality and even when no biological relationship exists between fertility and child health, as long as the fertility control cross-price elasticity for health is sufficiently strong and negative; i.e., fertility and health are gross substitutes. Condition (9) also suggests that, given the optimality of fertility control subsidization, the lowest-health (eradication effect) and the highest-fertility (cost effect) households would receive

the largest family planning subsidies. However, this does not imply that those households with the lowest health endowments receive the highest subsidies. Indeed, the distributional rules will depend on price effects. To see this, consider the effect on the subsidy rate to the ith household when that household experiences an increase in its health endowment μ^i and adjusts its behavior accordingly. For simplicity, assume that the ith household's health is (initially) at the mean of the health distribution of poor households, i.e., $H^i - H^* = 0$. Total differentiation of the system of first-order conditions describing the wealthy-household allocations, treating price effects as parameters, yields:

$$(10) \quad \frac{ds_c^i}{d\mu^i} = [(\frac{p_x N}{h_x} h_x^i \frac{(1-\alpha^i)}{\Sigma N^i} (\frac{dX}{dp_c}) - 1) (\frac{dN^i}{d\mu^i}) - \alpha^i \frac{p_x N}{h_x} h_{xx}^i (\frac{dX^i}{dp_c}) (\frac{dX^i}{d\mu^i})] (\frac{ds_c^i}{d\theta^i})_c + [\frac{p_x N}{h_x} \alpha^i - s_c^i [\frac{dN^i}{d\mu^i})] \frac{ds_c^i}{dG}$$

Expression (10) has two terms, the first corresponding to the (compensated) own price or cost effect of the ith subsidy and the second corresponding to the associated income effect on the wealthy household. The magnitudes of these terms depend in turn on the magnitudes and signs of the endowment effects on the health investments and fertility of the ith household. These are given by

$$(11) \quad \begin{aligned} \frac{dX^i}{d\mu^i} &= -\frac{p_x N^i}{h_x^i} [h_{xx}^i (\frac{dN^i}{dp_x})_c - \frac{dN^i}{dF^i}] \\ \frac{dN^i}{d\mu^i} &= -\frac{p_x N^i}{h_x^i} [h_{xx}^i (\frac{dN^i}{dp_x})_c - \frac{dN^i}{dF^i}] \end{aligned}$$

where c denotes compensated effect:

From expressions (10) and (11), sufficient (but not necessary) conditions for larger family planning subsidies to be provided to low-endowment households (compensatory subsidization) are that (i) fertility and health are gross substitutes ($dX/p_c < 0$, $dN/dp_x > 0$ (so that $dN^i/d\mu^i \geq 0$) and (ii) $dX^i/d\mu^i > 0$; i.e., more endowed households invest less in health. In that case, the returns to further health investments will be smaller in high- than in low-μ households and high-μ households will have at least as many averted births (at least as high family planning subsidy costs) as low-μ households. Thus, where fertility control subsidies but not health subsidies are prevalent, fertility and health investments will likely be gross substitutes in recipient households and such subsidies will be distributed disproportionately to the lowest endowment households.

c. Combining Health and Family Planning Subsidies

Having shown that family planning subsidies can effectively substitute for health subsidies when there is (only) a health externality, under certain conditions, we now consider whether a combination of health and fertility control subsidies is redundant, that is, we consider whether fertility control subsidies will be used in addition to health subsidies in the presence of the health externality and in the absence of a population externality.

The budget constraint for the subsidy provider when both subsidies can be used is:

$$(12) \qquad G' = G + \Sigma\, s_x^i\, (1 + \theta_x^i)\, X^i N^i$$

and the equilibrium conditions for the two subsidy rates are:

$$(13) \qquad \frac{s_c^i}{p_c} = -\alpha^i \frac{p_x NX\ \varepsilon_{HH*}}{p_x N^i\ \varepsilon_{HX}} [\varepsilon_{HX}^i \frac{\eta_{Xp_c}^i}{\eta_{Np_c}^i} + \varepsilon_{HN}^i + (\frac{H^i - H*}{H*})] -$$

$$(\eta_{Np_c}^i)^{-1} (\frac{\nu^i - N^i}{N^i}) p_c + s_x^i \frac{X^i}{p_c} [1 + \frac{\eta_{Xp_c}^i}{\eta_{Np_c}^i}]$$

$$\text{or } \frac{s_c^i}{p_c} = 0.$$

$$(14) \qquad \frac{s_x^i}{p_x} = [1 + \alpha^i \frac{NX}{N^i X^i} \frac{\varepsilon_{HH*}\ \varepsilon_{HX}^i}{\varepsilon_{HX}} (\eta_{Xp_x}^i - \frac{\eta_{Np_x}^i\ \eta_{Xp_c}^i}{\eta_{Np_c}^i}$$

$$\frac{\eta_{Np_x}^i}{\eta_{Np_c}^i} \frac{p_c}{p_x} \frac{(\nu^i - N^i)}{X^i N^i}] \psi^{-1}$$

$$\text{or } \frac{s_x^i}{p_x} = 0,$$

where $\psi = (\eta_{Np_x}^i + \eta_{xp_x}^i) [1 - \frac{\eta_{Np_c}^i + (\eta_{Np_c}^i + \eta_{Xp_c}^i)}{\eta_{Np_c}^i (\eta_{Np_x}^i + \eta_{Xp_x}^i)}]$.

As before, the optimal subsidy levels depend on price effects. However, in this case both health and family planning subsidies may be used even if the objective functions for the low-income households are strongly separable, no biological relationship exists between family size and child health, and all low-income households invest equally in child health. In that case, the equilibrium conditions are:

$$(15) \qquad \frac{s_x^i}{p_x} = [1 + \alpha^i \frac{NX}{N^i X^i} \frac{\varepsilon_{HH*}\ \varepsilon_{HX}^i}{\varepsilon_{HX}} \eta_{Xp_x}^i] (\eta_{Xp_x}^i)^{-1}$$

$$(16) \quad \frac{s_c^i}{p_c} = s_x^i \frac{X^i}{p_c} - \left(\frac{\nu^i - N^i}{N^i}\right) (\eta_{Np_c}^i)^{-1}.$$

$$\text{or } \frac{s_c^i}{p_c} = 0$$

Expression (15) indicates that the health subsidy will be used as long as there is a health externality. Expression (16) indicates that with sufficiently high health-subsidy expenditures by the subsidizing agent, positive family planning subsidies will also be optimal. Moreover, family planning and health subsidies, where both are used, will be positively correlated. The complementarity between the two subsidies, despite the single health externality, arises from the interaction between family size and per-child health expenditures in the "governmental" budget constraint (12)--an increase in the family planning subsidy to household i which lowers family size in i by one child saves the subsidizing agent the amount $s_x^i X^i$; the cost of that increase depends positively on the number of births averted by the ith household and inversely on the magnitude of its (own) fertility control elasticity.

When strictly positive health and family planning subsidies are jointly optimal, the magnitudes of the subsidies will also generally depend on the differing health endowments of the recipient households. Moreover, the direction of the endowment-subsidy association is likely to be identical for both the family planning and health subsidy. However, unlike in the single-subsidy case, no simple sufficient condition regarding household demand relationships determines the sign of the associations between the two subsidies and the health endowments.

3. Empirical Application: Laguna Province, the Philippines

a. The Data and the Distribution of Government Facilities

We have shown that the effects of government interventions on per-child health within a family are incorrectly estimated if the distribution of those interventions are influenced by the health predispositions of households, associated with endowments or tastes, that are unobserved by the researcher. In order to correctly assess the impact of government programs designed to influence health outcomes and to discover the government placement rules, it is thus necessary either to estimate or to measure pre-program heterogeneity in health outcomes. We will attempt to obtain consistent estimates of both the health effects of governmental family planning and health facilities and of facility placement rules based on longitudinal data describing the distribution of such public programs and child health in 20 barrios (villages) in the lowland rice-producing areas of Laguna Province in the Philippines. Information from surveys of 240 randomly-selected households residing in these barrios on the age, height and weight of every family member was collected in 1975 and 1979. Information was also obtained in the 1979 survey round on the dates of introduction of rural health clinics, family planning clinics, and primary schools financed by the national government for each of the barrios.

The distribution of the public facilities across barrios is reported in Table 1. While all but two of the twenty barrios have a public primary school, with such schools having been in existence for at least fifteen years prior to 1979 for each of the other eighteen barrios, health and family planning facilities were more recently introduced and are less prevalent.

Table 1

Distribution of Public Facilities in Twenty Laguna Barrios by Number of Years Instituted Prior to 1979

Years in Barrio	Family Planning Clinic	Rural Health Clinic	Primary School
0	8	7	2
0 - 4	4	3	0
5 - 9	5	0	0
10 - 14	2	4	0
15 - 19	0	2	4
20 +	1	4	14
Total	20	20	20

Seven barrios had no public health clinic and eight barrios had no family planning facility by 1979, with seven of the thirteen existing health facilities and eleven of the twelve family planning facilities introduced less than fifteen years prior to 1979.

The joint distribution of the family planning and health clinics appears in conformity with the health externality model, as such facilities appear to be placed in a complementary pattern--the Spearmen rank correlation of establishment dates for the family planning and health clinics is .62. Moreover, of the seven barrios that had no health clinic, five also did not have a family planning clinic and of the eight barrios without a family planning clinic, five also did not have a health clinic. Five barrios had neither facility as of 1979. The existence of two barrios without a health clinic but with a family planning clinic suggests, as noted, that if direct population externalities are ruled out, child health and family size should appear to be gross substitutes among the Laguna households. This is confirmed below.

b. Estimation Framework

To exploit the longitudinal data on health and the information on the dates of program initiation, we modify the above framework to accommodate the realities that governmemt programs are initiated at different times and that observed child health in any period is a stock variable influenced by resources allocated in the current and prior periods. The impact of a program on the current health status of a particular child will thus depend upon the length of its previous exposure to the program. We will exploit the variability in program exposure across children to estimate the effects of the health and family planning programs and to estimate the barrio-specific health endowments. Variation in program exposure across children, however,

occurs both because barrios differ in the timing of program introduction and because children within the barrio differ in their dates of birth. If child health investments differ systematically by the birthdate of the child due to health-related factors about which the researcher is unaware, a spurious relationship between child health and program exposure is generated even if the timing of government programs across barrios is unrelated to family or barrio endowments.

Let t_s represent the year of the survey, t_p the year the program was instituted, and t_b the year of birth of a child. The program will have been in effect $t_s - t_p$ years and for children born prior to t_p, i.e., $t_b < t_p$, $t_s - t_p$ will be the number of years each such child will have been exposed. Yet, a child born one year prior to the program will likely be more strongly impacted by the program than a child born five years prior to the program. We thus adopt as a measure of program exposure the fraction of a child's lifetime during which the child was exposed to the program. Let $p^a_{i\ell}$ be the program exposure of child i residing in barrio ℓ who is of age a at the survey date, where $a = t_s - t_b$. Thus,

$$p^a_{i\ell} = 0 \text{ if the program does not exist in the barrio as of the survey date}$$

$$= \frac{t_s - t_p}{a} \text{ if } t_s \geq t_p > t_b \text{ in barrio } \ell$$

$$= 1 \text{ if } t_s > t_b \geq t_p \text{ in barrio } \ell.$$

Consider the following child health demand equation for a child i aged a in barrio ℓ observed at t_s:

$$(17)\quad H^a_{i\ell t_s} = p^a_{i\ell}\,\beta + u_i + \mu_\ell + \varepsilon_{i\ell t_s},$$

where H is an age-standardized measure of health, u_i is a time-invariant, child-specific health endowment, the μ_ℓ are location-specific health factors and ε is a random error term. Least squares estimation of (17) when μ_ℓ is unobserved leads to a biased estimate of β, the program exposure effect, if t_p, the date the program was introduced, is related to the area's endowments, as would be the case with non-random program placement.

Within-family or barrio estimators of β, which purge out, respectively, household and locational characteristics, are also biased, however, even if program placement is uncorrelated with child or family-specific endowments u if child-specific health endowments (within-family) or household endowments (within-barrio) influence the spacing of children. In differenced form, for a family with at least one child born prior to the program's introduction, the within-family estimator is:

$$(18) \quad H^a_{j\ell t_s} - H^a_{i\ell t_s} = \left[\frac{(t_s - t_p)\ (t_{b'} - t_b)}{(t_s - t_{b'})\ (t_s - t_b)}\right]\beta + (u_j - u_i) + (\varepsilon_{jt_s} - \varepsilon_{it_s})$$

where $a' = t_s - t_{b'} > a$. As can be seen, even if the dates of program introduction t_p are independent of the child-specific error u, if child j's birth date $t_{b'}$ is related to his/her older siblings' health status u_i the within-family estimate of the program exposure effect is also biased. In Rosenzweig and Wolpin (1984) and Rosenzweig (forthcoming) it is shown that birth spacing and other child-specific inputs are significantly correlated with <u>prior</u> sibling and family-specific endowments, leading to biased estimates of child-specific resource allocations. Thus, as long as program placement is not responsive to purely random disturbances (or perturbations with little persistence), only within-child estimators will yield consistent estimates of the effect of program exposure, given systematic program placement and endowment-conditioned birth spacing behavior. Longitudinal information on child health outcomes is required.

c. Program Assessment: Comparisons of Cross-Sectional and Panel Estimates

To estimate the effects of the facilities on child health, we selected a sample of children (defined to be under eighteen as of 1979) for whom height and weight information exists in both years of the Laguna survey, yielding a working sample of 274 children in eighty-five households. Table 2 provides descriptive statistics for the sample children at each of the two survey dates. Height and weight are standardized by age and sex according to a national schedule. The average child in this sample in each of the two survey years is somewhat over ninety percent as tall as the average Filipino child of the same age and sex but only a little over eighty percent as heavy. However, the average child in the sample has evidently grown slightly in both dimensions relative to the standard between the two surveys.

Three separate specifications were estimated corresponding to alternative assumptions about unobservables in the determination of height and weight. In the first two colunns of Table 3 ordinary least squares regressions are reported using the 1979 cross-section of 274 children. The second pair of columns repeats the cross-sectional regressions but adds barrio dummy variables. The third set of columns reports first-differenced regressions using the 1979 and 1975 (matched) samples. The first column of each set includes the child's exposure to primary schools in addition to exposure to the health and family planning clinics. In the upper half of the table the dependent variable is the log of standardized height; in the lower half the dependent variable is the log of standardized weight.

The differences in estimated program exposure effects across the specifications are striking for either health measure. In the height regressions, both the cross-section and barrio fixed-effect health and

Table 2

Sample Statistics

Variable	Mean	S.D.
Natural logarithm of height normalized by Philippines age standard, 1975	4.525	.0715
Natural logarithm of height normalized by Philippines age standard, 1979	4.543	.0566
Natural logarithm of weight normalized by Philippines age standard, 1975	4.377	.147
Natural logarithm of weight normalized by Philippines age standard, 1979	4.407	.147
Exposure to public health unit, fraction of years, 1975	.456	.480
Exposure to public health unit, fraction of years, 1979	.512	.451
Exposure to family planning clinic, fraction of years, 1975	.162	.314
Exposure to family planning clinic, fraction of years, 1979	.285	.333
Exposure to public primary school, fraction of years, 1975	.871	.330
Exposure to public primary school, fraction of years, 1979	.872	.336
Number of years rural health in barrio, 1979	10.0	10.3
Number of years family planning clinic in barrio, 1979	6.45	13.67
Number of years public primary school in barrio, 1979	28.3	16.6
Number of barrios	20	
Number of children	274	

Table 3

Estimates of the Effects of Exposure to Governmental Programs on the Standardized Height and Weight of Children

Variable	OLS Cross-Section[a]		Fixed-Effect Barrio		Fixed-Effect Child	
			Log of Standardized Height			
Rural Health Unit Exposure	-.00976	-.00473	.00950	-.0205	.0507	.0511
	(1.04)	(0.53)	(0.16)	(0.40)	(1.58)	(1.21)
Family Planning Exposure	-.00605	-.0131	-.0135	-.00913	.0709	.0710
	(0.49)	(1.12)	(0.40)	(0.27)	(3.31)	(3.32)
Public Primary School Exposure	-.0193	-	-.172	-	(.0569	-
	(1.75)		(1.14)		(0.10)	
R^2	.0448	.0339	.1738	.1695	.0741[b]	.0660[b]
F	2.09	1.88	2.09	2.12	7.23	9.61
d.f.	267	268	248	249	271	272
			Log of Standardized Weight			
Rural Health Unit Exposure	-.0443	-.0313	-.0762	-.162	.0235	.0992
	(1.82)	(1.35)	(0.50)	(1.20)	(1.59)	(1.52)
Family Planning Exposure	.0446	.0263	.0677	.0803	.0990	.121
	(1.40)	(0.87)	(0.75)	(0.90)	(1.99)	(2.76)
Public Primary School Exposure	-.0503	-	-.494	-	.240	-
	(1.76)		(1.24)		(0.20)	
R^2	.0447	.0337	.1453	.1401	.0514[b]	.0500[b]
F	2.08	1.87	1.69	1.69	4.89	7.16

[a]Equation also includes the age and educational attainment of each parent.
[b]From first-differenced equation.

family planning clinic "effects" are generally negative with standard errors that are at least as large as the point estimates. The child fixed-effect (longitudinal) estimates, however, indicate that exposure to health and family planning clinics increases height, with the family planning effect statistically significant at the usual confidence levels and the health clinic effect marginally significant. The point estimates indicate that the height of a child for whom no health clinic existed would be five percent below that for a child always exposed to a clinic, while exposure to a family planning clinic increases height by seven percent. The same pattern emerges for public primary schools, although in that case the child fixed-effect point estimate has a very large standard error, due most likely to the small variance in exposure associated with the longevity of public schools displayed in Table 1.

The weight regressions tell a very similar story: the cross-section and within-barrio associations between health clinic exposure and age-standardized weight are negative, while the child fixed effect estimates, measured relatively precisely, indicate that exposure to either the health or family planning programs increases the weight-for-age of children. Here, however, the family planning effect is somewhat more robust to specification, although the effect of this program on child weight is underestimated by more than 100 percent when only the cross-sectional variation in program placement is utilized. The point estimates (last column) indicate that unit increases in health and family planning clinic exposure increase age-standardized child weight by nine and twelve percent respectively.

d. Program Placement Rules

Whether child health status is measured by age-standardized height or

weight, the estimates of the child health effects of the family planning program purged of contamination by the endogeneity of program placement or birth-spacing in Table 3 indicate that child health and family size are substitutes--subsidies to fertility control evidently augment resource allocations to child health investment among Laguna households. Thus, as we have shown, family planning clinics may substitute for health clinics in the presence of health externalities and/or may effectively complement health clinics even in the absence of other externalities, due to the interaction between family size and health investments in the "governmental" budget constraint.

In this section we seek to discern whether the dates of introduction of both the health and family planning clinics are systematically related to the average child health endowment within a barrio, i.e., we estimate the governmental program allocation rules. The child-specific effects that are estimated from (17) contain the elements u_i, μ_ℓ, a constant, and the effects of all time-invariant determinants of height and weight, e.g., mother's schooling, but net out the effects of the programs. However, since there are only two observations on each child, the estimated fixed effect measures the true pre-program child effect with error. Averaging child-specific effects within each barrio thus yields a measure (gross of time-invariant factors and random errors) of pre-program barrio level health presumably observed by the government, though only indirectly by us, and used by it to plan the timing of public program introduction. We have two such measures, corresponding to height and to weight.

Table 4 reports the estimates of the impact of the average barrio-level health endowment as measured by (the ℓn of) child height on the length of

Table 4

Estimates of the Effects of Barrio Child Health Conditions

on the Placement of Governmental Programs[a]

Endowment Measure	Public Program: Rural Health Unit (1)	Rural Health Unit (2)	Rural Health Unit (3)	Family Planning Clinic (1)	Family Planning Clinic (2)	Family Planning Clinic (3)	Primary Public School (1)	Primary Public School (2)	Primary Public School (3)
ℓn Height, standardized[b]	-102 (0.92)	-	-	13.2 (0.09)	-	-	-81.9 (0.46)	-	-
ℓn Height Effect	-	-145 (4.30)	-	-	-129 (2.30)	-	-	-58.3 (0.77)	-
Predicted ℓn Height Effect	-	-	-199 (4.34)	-	-	-151 (1.91)	-	-	-17.3 (0.16)
R^2	.0452	.507	.512	.0004	.228	.168	.0114	.0316	.0015

[a]Dependent variable = years since program was initiated.
[b]OLS coefficient.

time in years that each of three programs--health clinics, family planning clinics, primary schools--have been in existence in the barrio. There are thus twenty observations. Parental education levels were intially included as discussed above, but were jointly insignificant at conventional levels and so are excluded from the results actually presented. The first row uses actual mean height in the barrio and would only be correct if the programs themselves had no impact on height. The second row uses the barrio-average fixed-effect computed from the child fixed-effect height regression reported in column 6 of Table 3. The third row uses the predicted height fixed-effect obtained from a first stage regression in which the (ℓn) height fixed-effect is regressed on the (ℓn) weight fixed-effect, computed from the last column estimates of Table 3. The purpose of this latter procedure is to purge the estimate of the height fixed effect of measurement error under the assumption that height and weight are both measures of the same underlying health indicator.

While the timing of program initiation for all three programs appears unrelated to average child height in the barrios (row one), when the height effects of the three programs are removed, as in rows two and three, the estimates indicate that the health and family planning clinics were distributed systematically over time and, as expected, were allocated in a similar manner. Moreover, the statistically significant, fixed effect estimates imply a compensatory government allocation rule for both of the evidently complementary health programs. Barrios with lower pre-program health "endowments" evidently received both types of health-augmenting programs earlier. The timing of primary public school placement is not significantly related to the health endowments, however, a result consistent with our finding that such schools do not appear to significantly affect child health.

The point estimates based on the predicted height measure indicate that where pre-program standardized height was one percent higher (about one-fourth of the standard deviation) the introduction of a health unit was retarded by about two years. The distribution of family planning clinics was almost as responsive to health endowment variation; their introduction was delayed by about one and one-half years for every percentage increase in standardized height. The compensatory program placement rule followed by the governmental authorities for the complementary health and family planning programs thus appears to have been responsible for the significant negative biases observed in the cross-sectional estimates of the effects of the two programs in Table 3.