ECONOMIC DEVELOPMENT CENTER

THE SELECTIVITY OF FERTILITY AND THE DETERMINANTS OF HUMAN CAPITAL INVESTMENTS: PARAMETRIC AND SEMI-PARAMETRIC ESTIMATES

Mark M. Pitt
Mark R. Rosenzweig

ECONOMIC DEVELOPMENT CENTER

Department of Economics, Minneapolis

Department of Agricultural and Applied Economics, St. Paul

UNIVERSITY OF MINNESOTA

The Selectivity of Fertility and the Determinants of Human Capital

Investments: Parametric and Semi-Parametric Estimates

Mark M. Pitt

Brown University

and

Mark R. Rosenzweig

University of Minnesota

October 1989

Abstract

The Selectivity of Fertility and the Determinants of Human Capital

Investments: Parametric and Semi-Parametric Estimates

Mark M. Pitt and Mark R. Rosenzweig

In this paper we assess the importance of heterogeneity and selective

fertility in altering estimates and interpretations of the determinants of

the human capital of children. We set out a sequential model of human

capital investments in children incorporating endogenous fertility and

heterogeneity in human capital endowments to

illustrate the fertility selection problem and issues of identification.

Empirical results based on parametric and semi-parametric estimates of

selectivity models applied to data on birthweight and schooling in Malaysia

indicate that the hypothesis of no fertility selection is strongly rejected,

with mothers having higher birthweight children tending to have

substantially lower birth probabilities (negative birth selectivity). As a

consequence, the positive association between mother's schooling and

birthweight is substantially underestimated and the positive effects of

delaying childbearing overestimated when birth selectivity is not taken into

account. The schooling results indicate strong rejection of the "efficient

schooling" model, in which schooling is allocated efficiently across

children, but only when the selectivity of fertility is taken into account.

Considerable attention has been paid in recent years to the influence of public programs and the characteristics of parents on both the health and schooling of children (Behrman, forthcoming; Behrman and Deolalikar, 1988; Pitt and Rosenzweig, 1985; Schultz, 1988; Strauss, 1988). Inferences are commonly drawn about the relative effectiveness of a variety of interventions, including family planning initiatives, in influencing such measures of the human capital of children as frequency of illness, height and weight or school enrollment. A pervasive finding is the importance of mother's schooling attainment in determining the health status and schooling of children, with some support as well for the hypothesis that family planning programs augment child health.

A related set of studies, concerned with how parental behaviors directly influence child health (Grossman and Joyce, 1988; Olsen and Wolpin, 1983; Rosenzweig, 1986, Rosenzweig and Schultz, 1983), has demonstrated that there is considerable heterogeneity in the health of children, net of parental investments. Such studies have also shown that parental investments appear to respond to the exogenous health-related characteristics (health endowments) of children. No studies, however, have considered how the expected health or human capital endowments of children (inclusive of exogenous individual-specific components of human capital and exogenous environmental influences on human capital) might have influenced the presence of the child whose health is being measured. Yet children of a given age j at the time of a sample survey, the units of analysis in all studies of child health or schooling, represent the outcomes of fertility decisions taken approximately j+1 or more years before the survey date. And births occurring in a particular time period are not likely to be randomly drawn from the potential population of households or parents; that is,

1

fertility rates are likely to differ systematically across households heterogeneous in human capital endowments, genetic or environmental. If parents differ in inherent healthiness, for example, then parental characteristics or fertility-related programs may be found to be related to the health outcomes of children, the "survivors" of birth processes, solely because they alter the composition (defined by healthiness) of households or parents that bear children.

A large literature also exists that demonstrates a strong relationship between measured parental characteristics and fertility, again most notably mother's schooling attainment and age (Birdsall, 1988; Hotz and Miller, 1988). The possibility that unobserved human capital endowments also influence fertility decisions, that fertility is selective, does not necessarily alter the validity of inferences obtained from fertility studies. However, the selectivity of fertility implies that information obtained from reduced-form child health studies, even in conjunction with fertility studies, does not provide all of the information required for understanding the effects of an intervention on child health or for testing models of parental behavior. Knowledge of the direction and magnitude of the selectivity of fertility is also required. For example, if an intervention lowers fertility, then it may appear to augment the average health of the population of children born if fertility is positively selective or may appear to lower average child health if fertility is negatively selective, even if there is no direct influence on the allocation of resources to those children who are born. Policy conclusions about the value of expanding family planning interventions or opportunities for the schooling of women may thus be quite different if their sole effect on the measured human capital of children results from a reduction in the number of

2

low-endowment children who are born or results solely from a shift in births from households that care little about human capital to those that care a great deal, for given resource levels.

Despite pervasive findings indicating socioeconomic differentials in fertility and heterogeneity in human capital endowments, no studies of birth outcomes, child health, or schooling investment, inclusive of studies that claim that the number of children and maternal age at birth are "endogenous" (Rosenzweig and Schultz, 1983), have taken into account the choice-based nature of samples of children due to the possible selectivity of fertility. This is so despite the availability of econometric techniques for obtaining estimates corrected for non-random sample censoring, which have been applied to a variety of topics in the economics (and sociology) literature in the last decade.

There may be two reasons for the absence of studies of human capital investment that account for fertility selection. First, consideration of the selectivity of births in estimating the determinants of the human capital of children imposes severe data requirements. Information on children must be obtained from a probability sample of all households. Yet, for example, most health surveys are based on samples of children, births or pregnancies (e.g., the U.S. National Health Examination Surveys, the U.S. National Nativity Followback Surveys) so there is no information on women (potential mothers) who did not bear children. Moreover, most fertility or general purpose surveys, based appropriately on samples of (fecund) women or households, do not collect information on child health or other indicators of children's human capital. Few data sets are comprehensive in their coverage of fertility, the human capital of children and the socioeconomic characteristics of parents.

A second problem in applying standard selectivity models to the determinants of the human capital of children is that identification of interesting parameters, when there is fertility selection, is difficult. The inherent sequencing of births and human capital investments in children make it theoretically implausible, except in some special cases, discussed below, to employ standard exclusion restrictions applied to regressors to achieve identification. It is difficult to justify the existence of variables that influence fertility that do not also affect human capital investments in the children born. As a result, the choice of a distributional assumption for errors also serves as the critical identification restriction. The results obtained may be quite sensitive to such assumptions, none of which are justified by economic theory. Standard models of selectivity have considered only a limited number of distributions (normal (Heckman, 1979), logit (Lee, 1982), uniform (Olsen, 1982)). Recently, however, methods of estimation for selectivity models have been developed (Ichimura, 1988; Ichimura and Lee, 1988) that can yield consistent estimates of behavioral parameters without imposing any distributional assumptions. Such procedures permit tests of the validity of the distributional assumptions commonly employed in selection models and, as we show, this is true even when all behavioral parameters are not identified without such assumptions.

In this paper we assess the importance of heterogeneity and the selectivity of fertility in altering estimates and interpretations of the determinants of two measures of child human capital frequently studied by economists--weight at birth and schooling--based on household data from Malaysia. In part 1 of the paper we briefly set out a sequential model of human capital investments in children incorporating endogenous fertility and

4

heterogeneity in human capital endowments and parental preferences. The model is used to illustrate the fertility selection problem and issues of identification. In section 2 of the paper we set out the model to be estimated and discuss estimation procedures, identification and tests of distributional assumptions using semi-parametric estimators. In section 3, we describe the data used and report our findings. The results indicate that (i) the hypothesis of no fertility selection is strongly rejected, with mothers having higher birthweight children tending to have lower birth probabilities (negative birth selectivity), (ii) as a consequence, the positive association between mother's schooling and birthweight is substantially underestimated and the positive effects of delaying childbearing overestimated when birth selectivity is not taken into account, (iii) birth selectivity appears to be far more important in biasing estimates of the determinants of birthweight than is the selectivity of mortality, at least in Malaysia, and (iv) the assumption of the joint normality of the birthweight and fertility disturbances, employed in estimation, could not be rejected for the birthweight model. The schooling results indicate strong rejection of the "efficient schooling" model, in which schooling is allocated efficiently across children, a result robust to fertility selection and obtained without distributional assumptions about unobservables. The normal maximum-likelihood estimates of the (validated) non-efficient schooling model, while indicating fertility selectivity similar to that for birthweight, could not, however, pass the test of the distributional assumption. Nor could we find another set of parametric distributional assumptions that both allowed identification and could not be statistically rejected versus a non-parametric alternative.

5

1. Theory

a. Selective fertility and the allocation of human capital

To illustrate the relationship between human capital investment behavior and the potentially selective censoring of data associated with the process of fertility in a world of heterogeneous agents, we employ a simple three-period decision model. We assume that parents can have a child in each of the first two periods and can invest in the human capital of each child only in the period after it is born. The production of human capital h_j^i for the ith child in the jth family, if it is born, is described by the technology

$$(1) \quad h_j^i = h(y_j^i, \mu_j) \ ,$$

where y_j^i = resources provided to child i in family j, μ_j = the human capital endowment of the child, which we have, for simplicity, assumed to be the same for all children, and $\partial h_j^i / \partial y_j^i > 0$.

Assume that it is optimal for all parents to have a child in their first period. The decision to have a second child is made by comparing maximal lifetime utility with and without that child. The former is given by

$$(2) \quad v_j^1 = \max_{y_j^i, x_j^i} \{U^0(1, x_j^0) + \delta U^1(h_j^1, 2, x_j^1) + \delta^2 U^2(h_j^2, 2, x_j^2)\}$$

$$\text{s.t.} \quad F_j = p_x \sum_{i=0}^{2} x_j^i \psi^i + p_y \sum_{i=1}^{2} y_j^i \psi^i + c(1 + \sum_{i=1}^{2} \psi^i) \ ,$$

where U^i = utility in the ith period of the parents' life, F_j = lifetime income, δ = subjective discount rate, ψ = market discount rate, p_x = price of the consumption good x, p_y = price of the human capital investment good y, and c is the per-period cost of a child, which is borne by the household

6

for the first two periods of its life.[1] If the family foregoes having a
second child its maximal lifetime utility is:

$$(3) \quad v_j^0 = \max_{y_j^1, x_j^i} \{ U^0(1, x_j^0) + \delta U^1(h_j^1, 1, x_j^1) + \delta^2 U^2(0, 1, x_j^2) \}$$

$$\text{s.t. } F_j = p_x \sum_{i=0}^{2} x_j^i \psi^i + p_y y_j^1 \psi^1 + c(1 + \sum_{i=1}^{2} \psi^i) .$$

Whether or not the parents have the second child depends on whether the
difference between (2) and (3), ∇_j, given by (4),

$$(4) \quad \nabla_j = v_j^1 - v_j^0 ,$$

is positive.[2] The usual restrictions on preferences imply that higher
levels of the endowment μ always induce a higher level of h (even if
investments y are compensatory). Fertility will thus be negatively
selective with respect to human capital outcomes h if $d\nabla_j/d\mu_j < 0$ and
positively selective if $d\nabla_j/d\mu_j > 0$.

Consider the household m from the population of households
characterized by identical levels of income and facing identical price
vectors that is just indifferent between having the (second) child or not;
i.e., for whom $\nabla_m = 0$. All households differ in their endowment μ; the mth
household, with endowment μ_m, is thus the "marginal" fertility household in
the population such that a change in any price must move the household to
become a member of either the population that does or does not bear a child.
We can ascertain how the endowment of the marginal household differs across
populations facing different prices by totally differentiating (4). The
total differential is:

7

$$(5) \quad d\nabla_j = \frac{\partial \nabla_j}{\partial \mu_j} d\mu_j + \frac{\partial \nabla_j}{\partial c} dc + \frac{\partial \nabla_j}{\partial p_y} dp_y \;,$$

which vanishes for the marginal household (j=m). The change in the endowment μ_m of the marginal household as the cost c of a child increases that maintains the indifference ($d\nabla_m = 0$) is thus given by

$$(6) \quad \frac{d\mu_m}{dc} = - \frac{\partial \nabla_m / \partial c}{\partial \nabla_m / \partial \mu_m} \;.$$

Because a rise in c must decrease ∇, and thus lead to lower fertility, if fertility is positively (negatively) selective the endowment of the marginal household will be higher (lower) in a population in which the direct cost of children is higher. Therefore, the sub-population of households giving birth is characterized by a higher (lower) average endowment in environments with higher costs of fertility--only higher(lower)-endowment households choose to have a second child when faced by high costs of fertility.

That endowments influence fertility is the reason that the composition of births (households giving births), characterized by endowments, changes in response to a price change.[3] Price changes also have direct effects on the allocation of resources y to children that are born, which may offset or reinforce the changes in the average endowments of the population of births. The first-order necessary condition for the allocation of y^i to the first or second child, conditional on its birth, is given by (7):

$$(7) \quad u_h^i h_y^i = \lambda p_y \, (\psi^i / \delta^i) \;,$$

where λ = marginal utility of income.

The decision rules for y^i for parents who bear a child can be written as

(8) $y^i = y(p_y, c, p_x, F, \mu, i)$,

which has the same arguments as the decision rule for whether or not to have
a child, given by (4). While (4) and (8) are unlikely to have the same
functional form, as can be seen, the model presents no obvious exclusion
restrictions in the sense that the rule for any particular decision is
influenced by some exogenous variable not influencing another decision.[4]
Even if there was uncertainty about endowments and decisions were sequential
and myopic, because the birth decision always precedes the human capital
investment decision, there cannot be less observed or known (to parents)
influences on human capital investments than on the birth decision
(information always accumulates).

A special case of the fertility cum human capital model which has been
given considerable attention in the literature (Becker and Tomes, 1979;
Behrman et al., 1985) is the "efficient schooling" model. This model does
give rise to exclusion restrictions, which, as we show below, aid in the
identification of the selectivity effects of price changes. In the
efficient schooling variant of the model, parents do not care about the
human capital (schooling) of their children per se, but about their
children's incomes. Moreover, parents can directly transfer to or extract
resources from children. The income I_j^i of child i in family j is thus

(9) $I_j^i = \alpha h_j^i + b_j^i$,

where α = the rental rate per unit of human capital and b_j^i = parental income
transfer.

The efficient schooling model is thus

9

(10) $v_j^{1e} = \max_{y_j^i, x_j^i, b_j^i} \{U^0(1, x_j^0) + \delta U^1(I_j^1, 2, x_j^1) + \delta^2 U^2(I_j^2, 2, x_j^2)\}$

$\quad$ s.t. $\quad F_j = p_x \sum_{i=0}^{2} x_j^i \psi^i + p_y \sum_{i=0}^{2} y_j^i \psi^i + c(1 + \sum_{i=1}^{2} \psi^i) + \sum_{i=1}^{2} b^i \psi^i$

and

(11) $v_j^{0e} = \max_{y_j^1, x_j^i, b_j^1} \{U^0(1, x_j^0) + \delta U^1(I_j^1, 1, x_j^1) + \delta^2 U^2(I_j^1, 0, 1, x_j^2)\}$

$\quad$ s.t. $\quad F_j = p_x \sum_{i=0}^{2} x_j^i \psi^i + p_y y_j^1 \psi^1 + c(1 + \sum_{i=1}^{2} \psi^i) + b^1 \psi^1$,

where the decision rule for fertility is the same as (4), with v_j^k replaced by v_j^{ke}, $k = 0,1$. In this model, it is easy to show that the necessary first-order conditions for the y (schooling) are

(12) $\alpha h_y^i = p_y$,

which do not depend on the utility function of the parents. The decision rule for the allocation of schooling, when allocated efficiently, is thus

(13) $y^i = y^e(p_y, \mu)$.

Parental income, the timing (order) of the birth, and the direct cost of children do not influence the allocation of schooling in this model, although these variables clearly still influence the decision whether to have a child and thus the endowments of children who are born. Since endowments influence schooling, even when allocated efficiently, as in (13), changes in, say, the cost of children will thus alter the average level of schooling allocated to children via the change in the average endowments of the children born, as in (6). Testing the exclusion restrictions of the efficient schooling model may thus yield misleading results when birth selectivity is not taken into account.

b. Statistical considerations

The behavioral models outlined indicate that changes in income or prices can alter the human capital characteristics of the population of children born by altering both who is born, or which parents bear children, and the resources allocated to children who are born. A linear representation of estimating equations corresponding to the solved-out (reduced-form) decision rules of models such as described by (1) through (8) or (9) through (13) is

$$(14) \quad f^* = X_f\beta_f + \beta_{fh}\mu_h + v_f = X_f\beta_f + \epsilon_f$$

$$(15) \quad h = X_h\beta_h + \mu_h + v_h = X_h\beta_h + \epsilon_h.$$

where f^* corresponds to the differential ∇, h is a measure of human capital, and the X_j, $j=f,h$, are vectors of exogenous variables corresponding to prices and income; the compound error terms in each equation contain stochastic variables unknown to the data analyst but known to the parents (the endowments μ_j), and an error term summarizing all non-systematic shocks (v_f and v_h). The β's are the behavioral responses by parents to the exogenous variables inclusive of the endowments.

If the error terms have a zero mean, then the covariance between the compound error terms ϵ_f and ϵ_h is

$$(16) \quad \text{cov}(\epsilon_f, \epsilon_h) = \beta_{fh}\text{var}(\mu)$$

if we assume that $\text{cov}(v_f, v_h) = 0$. Covariation between the disturbances in the (latent) fertility and health equations arises as long as fertility responds to human capital endowments, i.e., $\beta_{fh} \neq 0$ (as indicated by the model) and there is unmeasured population variability in human capital

11

endowments $(\text{var}(\mu) \neq 0)$.

In the case in which all errors are jointly normally distributed, it is straightforward to estimate the bias in the estimates of the reduced-form human capital equation (15) when selection associated with fertility is not accounted for. The regression function (Heckman, 1979) for the "population" of births, suppressing subscripts, would be

$$(17) \quad E(h*|X_h, f*>0) = X_h\beta_h + E(\epsilon_h|\epsilon_f > -X_f\beta_f) = X_h\beta_h + \text{cov}(\epsilon_h,\epsilon_f)\lambda,$$

where $\lambda =$ the ratio of the density and distribution functions for the standard normal variable $X_f\beta_f$ with $\text{var}(\epsilon_f)$ normalized to unity. Estimating (17) based on the sample of births without taking into account birth selection is equivalent to omitting the λ term in (10). The estimated effect of a change in an X on h in the choice-based births sample is then, from (16),

$$(18) \quad \frac{dE(h*|X, f*>0)}{dX} = \beta_h + \text{cov}(\epsilon_h, \epsilon_f)\frac{\partial\lambda}{\partial X} = \beta_h - \beta_f A\beta_{fh}\text{var}(\mu)$$

where $A = \lambda^2 + X_f\beta_f\lambda$. The second (negative) term in (18) is the bias arising from selection (in the normal case).

It is easy to see from (18) and the model why the pervasive finding that more educated mothers have healthier children could be solely the result of birth selection. If women with higher endowments are more likely to have children, <u>ceteris</u> <u>paribus</u>, $\beta_{fh} > 0$ (positive birth selectivity), then the covariance term will be positive. Since $\lambda > 0$ and $X_f\beta_f > 0$, the sign of the bias will depend on the sign of $\hat{\beta}_f$, the effect of mother's schooling on the (latent variable associated with the) probability of a birth. The negative effect of schooling on fertility is a common finding in

the literature; as a result of birth selectivity, therefore, the schooling effect on health is likely to be biased upward if fertility is positively selective. Alternatively, if fertility is negatively selective, then least-squares estimates of schooling effects are biased downward.

As is well known, and as expression (18) also indicates, the magnitude of the bias in choice-based birth (child) samples not only depends on the covariation between error terms and on the values of population coefficients, but on the degree of censoring. For example, for a representative woman in a population where one-half of women do not give birth in a given calendar interval ($X_f \beta_f = 0$), the A term in the bias component in (18) is 0.64; where 75 percent do not give birth the A-weight is 0.76. However, even in a high-fertility population, such as Bangladesh where only 15 percent of women of child-bearing age do <u>not</u> have a child in a pre-specified five-year age range as of a given date, the weight in the bias term is still 0.36.[5] The magnitude of the typical bias induced by fertility selectivity will thus vary across environments, being stronger in low-fertility compared to high-fertility environments. However, even in the latter settings, no more than 85 percent of the population of fecund women will have given birth to a child in any five-year time interval.

2. Estimation of the Fertility Selection Model

    a. <u>Standard normal-likelihood model</u>

    Because ∇ is unobserved, the full econometric model corresponding to (1) through (4) takes the conventional form

$$I_i^* = X_{fi}\beta_f + \epsilon_{fi}$$

(19) $\quad y_i^* = X_{hi}\beta_h + \epsilon_{hi}$

$$y_i = y_i^* \quad \text{if} \quad I_i^* > 0$$

$$= 0 \text{ if} \quad I_i^* \leq 0, \quad i=1,2,\ldots,n,$$

where I_i^* is a continuous latent variable underlying a dichotomous birth

realization indicator ($I_i=1$), y_i^* is a continuous measure of human capital,

X_{fi} and X_{hi} are sets of regressors associated with fertility and human

capital production, and ϵ_{fi} and ϵ_{hi} are the compound errors. If, as is the

standard assumption, the ϵ_{hi} and ϵ_{fi} are jointly i.i.d. drawings from a

bivariate normal distribution having zero means and covariance Σ

$$(20) \quad \Sigma = \begin{pmatrix} \sigma_f^2 & \sigma_{fh} \\ \sigma_{fh} & \sigma_h^2 \end{pmatrix}$$

The likelihood of this model is given by

$$(21) \quad L = \prod_{I=0} \text{Prob}(I_i^* \leq 0) \prod_{I=1} f(y_i | I_i^* > 0) \, \text{Prob}(I_i^* > 0)$$

$$= \prod_{I=0} \text{Prob}(I_i^* \leq 0) \prod_{I=1} \int_{-X_{fi}\beta_F}^{\infty} f(\epsilon_{fi}, y_i) d\epsilon_{fi}$$

where $f(.,.)$ is the joint density of ϵ_{fi} and y_i. Computation is

simplified by writing this joint density as the product of a conditional and

a marginal density, resulting in the likelihood

$$(22) \quad L = \prod_{I=0} (1-\Phi(X_{fi}\beta_f/\sigma_f)) \prod_{I=1} \Phi\left(X_{fi}\beta_f/\sigma_f + \frac{\sigma_{fh}}{\sigma_f\sigma_h}(y_i - X_{hi}\beta_h)\right)$$

$$\times \left(1 - \frac{\sigma_{fh}^2}{\sigma_f^2\sigma_h^2}\right)^{-\frac{1}{2}} \frac{1}{\sigma_h} \phi((y_i - X_{hi}\beta_h)/\sigma_h) \,,$$

14

where $\Phi(\ )$ and $\phi(\ )$ are the normal distribution and density functions, respectively. Only β_f/σ_f is identifiable and therefore σ_f is normalized to unity without loss of generality, but all other parameters are identifiable even if the sets of regressors X_{fi} and X_{hi} overlap completely; as demonstrated below, the assumption of normality "identifies" the β_h vector. A test of the null hypothesis of a zero correlation ρ between ϵ_{fi} and ϵ_{hi}, $\rho = \sigma_{fh}/\sigma_h$, is a test of the absence of birth selection bias.

b. **Semiparametric estimation**

The reliance on a distributional assumption for identification when there is birth selection, which appears necessary for estimating the determinants of health, a variable likely to be of direct concern to parents, is not fully satisfactory. Newly-developed semiparametric procedures (Ichimura, 1987 and Ichimura and Lee, 1988) permit, however, tests of distributional assumptions in selection models. The important advantage of semiparametric methods is that they yield consistent estimates of a model's parameters (or a combination of these parameters) even when the error distribution is not known to have any specific parametric form. There is usually no theoretical justification for choosing a particular parametric distribution. Our choice of normally-distributed errors in the likelihood derived above simply reflects common practice and computational considerations. Imposing an incorrect parametric distribution, which the normal distribution may well be, results in inconsistent estimates, which may be more biased than those which do not take into account selection at all. The disadvantage of the semiparametric method is that identification of the parameters β_h requires placing at least one exclusionary restriction on β_h; that is, at least one regressor in the vector X_f does not appear in the vector X_h. As noted above, such restrictions are not theoretically

15

justifiable in estimating the determinants of health, and are justifiable in the case of schooling only for the highly-restrictive efficient schooling model.

The identification problem is well illustrated by examining the least squares estimator (Heckman, 1976) for the selection model with normally distributed errors (described above)

$$(23) \quad E(y_i | I_i^* > 0) = X_{hi}\beta_h + E(\epsilon_{hi} | I_i^* > 0)$$
$$= X_{hi}\beta_h + g(X_{fi}\beta_f) \ ,$$

where $g(X_{fi}\beta_f) = \sigma_{fh}\phi(X_{fi}\beta_f)/\Phi(X_{fi}\beta_f)$. If X_f is a subset of X_h then equation (23) makes clear that only the nonlinearity of the function $g(\cdot)$ identifies the model. Not all distributional assumptions result in nonlinear functions $g(\cdot)$. Olsen (1980) suggested the use of the (0,1) uniform distribution for ϵ_{fi}. This distributional assumption results in a function $g(X_{fi}\beta_f)$ linear in the regressors X_{fi}. At least one exclusion restriction is required for identification in this case.

The Ichimura-Lee estimator places no prior restriction on the form of the function $g(X_{fi}\beta_f)$. Identification of β_h (except for an intercept) requires that at least one variable in X_{fi} is not included in X_{hi}. Consider the case in which X_{fi} and X_{hi} completely overlap. One cannot distinguish between the "true" model given by equation (23), and another model given by

$$(24) \quad E(y_i | I_i^* > 0) = X_{hi}\tilde{\beta}_h + h(X_{fi}\beta_f) \ ,$$

where $\tilde{\beta}_h = \beta_h + c\beta_f$, c is an arbitrary constant, and $h(X_{fi}\beta_f) = g(X_{fi}\beta_f) - X_{fi}c\beta_f$. Now consider fixing (normalizing) one of the coefficients in β_h associated with a regressor which appears in both X_{hi} and X_{fi}. To simplify the exposition, assume, innocuously, that the first k_f ($k_f + 1$ inclusive of

16

an intercept) regressors in X_{hi} are those that also appear in X_{fi}, that the first of these regressors is the one to be normalized, and that the remaining k_z ($k_z > 0$) regressors in X_{hi} do not appear in X_{fi}. In models concerned with birth outcomes, characteristics of children unforeseen by parents prior to conception, e.g., a child's gender or whether he or she was born a twin, are included in the k_z subset of regressors as they are likely to affect human capital investment but not the probability of conception. The normalization takes the form

(25) $\quad \delta_{h1} = \beta_{h1} + c\beta_{f1} = b_0$,

where b_0 is a normalizing value (with unity a convenient value for b_0). This normalization fixes the (otherwise arbitrary) constant c that precluded identification in the example above

(25') $\quad c = (b_0 - \beta_{h1}) / \beta_{f1}$.

Among the set of overlapping regressors, the jth normalized coefficient is then

(26) $\quad \delta_{hj} = \beta_{hj} + c\beta_{fj} = \beta_{hj} + (b_0 - \beta_{h1})\beta_{fj} / \beta_{f1}.\quad j=2,3,\ldots k_f.$

The coefficients β_{hj}, associated with the non-overlapping regressors ($j > k_f$), are identified with or without this normalization ($\delta_{hj} = \beta_{hj}$). The normalization $\delta_{h1} = b_0$ thus identifies a nonlinear function of the parameters β_{fj} and β_{hj}, all of which are individually identifiable if a parametric distribution is specified.

The benefit of estimating the normalized parameters δ_h is that they permit a test of the null hypothesis of normally-distributed errors versus a nonparametric alternative even in the absence of exclusion restrictions.

17

Denote the maximum likelihood estimates of the full set of parameters of the selection model (19) under the assumption of normality as $\hat{\beta}_f^N$ and $\hat{\beta}_h^N$. One can construct comparable estimates of δ_h for a given b_0 under the assumption of normality, $\hat{\delta}^N$, using equation (26), as all the relevant parameter values are identified under normality. Denoting $\hat{V}_\delta^N$ as the asymptotic covariance matrix of $\hat{\delta}^N$ under the null hypothesis of normality, and $\hat{V}_\delta^S$ as the asymptotic covariance matrix of $\hat{\delta}^S$ estimated semiparametrically, a test of normality is given by the Hausman-like test statistic

$$(27) \quad (\hat{\delta}^N - \hat{\delta}^S)'(\hat{V}_\delta^S - \hat{V}_\delta^N)^{-1}(\hat{\delta}^N - \hat{\delta}^S) \; ,$$

which is distributed as chi-square with degrees of freedom equal to the dimension of the vector δ minus one, ($= k_f + k_z - 1$). An estimate of the covariance matrix $\hat{V}_\delta^N$ can be derived by the delta method.

Two other tests are possible. First, the set of k_z child characteristics that appear among the X_h but not among the X_f because they are revealed to parents after the birth, are individually identified without imposing a parametric distribution. A test of the parametric distribution with less power thus would involve only a comparison of the k_z coefficients estimated from both models. A second test would impose empirically-justified exclusion restrictions, based on the normal (or other parametric distribution) maximum-likelihood estimates, in obtaining the semi-parametric estimates. That is, those variables found to be not jointly statistically significant based on the estimates obtained from the parametric model could be excluded and the joint null hypothesis of the exclusion restriction and the parameterization of the distribution could be tested.

In the case of normally-distributed errors, probit maximum-likelihood estimation consistently estimates the parameters β_f up to a scalar proportion. These parameters can also be estimated up to a scalar proportion semiparametrically. This binary choice model pertains to the class of models known as <u>single index models</u>, which in the case of a latent variable linear in the X's, take the form

$$(28) \quad I_i = \psi(X_{fi}\beta_f) + \nu_i \ ,$$

where the disturbances ν_i have zero mean, $E(\nu_i|X_{fi}) = 0$, and the transformation function ψ is not known. The conditional expectation $E(I_i|X_{fi}\beta_f)$ can be estimated from a random sample of size n by

$$(29) \quad E_{ni} = \frac{1}{(n-1)a_n} \sum_{j \neq i} I_j \ K(\frac{X_{fi}\beta_f - X_{fj}\beta_f}{a_n}) \ / \ \frac{1}{(n-1)a_n} \sum_{j \neq i} K(\frac{X_{fi}\beta_f - X_{fj}\beta_f}{a_n}) \ ,$$

where a_n is the bandwidth (window width) and K is a kernel function. Semiparametric least squares (SLS) estimation of the parameters β_f (up to a scalar proportion) is achieved by minimizing

$$(30) \quad \min_{\beta_f} \sum_{i=1}^{n} (I_i - E_{ni})^2 \ .$$

Semiparametric least squares estimation of the selection model (19) is accomplished by noting that $g(X_{fi}\beta_f) = E(\epsilon_h|I_i=1, X_{fi})$ can be estimated nonparametrically by

$$(31) \quad \xi_{ni} = \frac{1}{(n-1)a_n} \sum_{j \neq i}^{n} I_j (y_j - X_{hj}\beta_h) K(\frac{X_{fi}\beta_f - X_{fj}\beta_f}{a_n})$$

$$/ \frac{1}{(n-1)a_n} \sum_{j \neq i}^{n} I_j K(\frac{X_{fi}\beta_f - X_{fj}\beta_f}{a_n}) \ .$$

The joint SLS estimator of the system of equations (23) and (28) is

$$(32) \quad \min_{\beta_f, \beta_h} \ \frac{1}{n} \sum_{i=1}^{n} \left[I_i - E_{ni}\right]^2 + I_i \left[y_i - X_{hi}\beta_h - \xi_{ni}\right]^2 \ .$$

Ichimura and Lee demonstrate that this SLS estimator is $\sqrt{n}$ - consistent and asymptotically normal. The asymptotic covariance matrix of the limiting distribution of $\sqrt{n}(\hat{\theta} - \theta)$ is $B^{-1} \Omega B^{-1}$, where B is consistently estimated by

$$(33) \quad B_n = \frac{1}{n} \sum_{i=1}^{n} \left[\frac{\partial E_n'}{\partial \theta} \frac{\partial E_n}{\partial \theta} + I_i \left(\frac{\partial \beta_h'}{\partial \theta} X_{fi} + \frac{\partial \xi_n}{\partial \theta}\right) \left(\frac{\partial \beta_n'}{\partial \theta} X_{fi} + \frac{\partial \xi_n}{\partial \theta}\right)'\right] \ ,$$

where $\theta = \{\beta_f, \beta_h\}$, and Ω is consistently estimated by

$$(34) \quad \Omega_n = \frac{1}{n} \sum_{i=1}^{n} \left[\frac{\partial E_n}{\partial \theta} \hat{e}^2 \frac{\partial E_n}{\partial \theta'} + I_i \left(\frac{\partial \beta_h}{\partial \theta} X_{fi}' + \frac{\partial \xi_n}{\partial \theta}\right) \hat{v}^2 \left(\frac{\partial \beta_h}{\partial \theta} X_{fi}' + \frac{\partial \xi_n}{\partial \theta}\right)'\right] \ ,$$

where $\hat{e} = I_i - E_n$ and $\hat{v} = y_i - X_{hi}\hat{\beta}_h - \xi_{ni}$.

The kernel function chosen is a form of a biweight kernel function (Silverman 1986, page 43).

$$(35) \quad K(S) = \left[-D_1\left((1 - \frac{s^2}{2})^3 / \sqrt{2}\right) + D_2\left(2(1-s^2)^3 - (1 - \frac{s^2}{2})^3 / \sqrt{2}\right)\right]$$

and $D_1 = 1$ if $1 < S \leq 2$, zero otherwise, and $D_2 = 1$ if $0 < S \leq 1$, zero otherwise.

3. Application

    a. Data and sample criteria

    The essence of the birth selectivity model applied to the analysis of
the determinants of children's human capital is that a child of age i at the
time of a survey is the outcome of fertility decisions taken approximately
i+1 periods prior to the survey by a population of potential parents. In
order to characterize the selectivity that yields a sample of children
surveyed at a given date and to obtain correct estimates of the determinants
of the human capital of the children it is thus necessary to have
information on variables influencing both the birth probabilities of the
cohort of potential parents (defined by their date of birth) and the human
capital measure for the cohort of children born (defined by their date of
birth). In particular, it is necessary to have a well-defined probability
sample of women "at risk" with respect to fertility combined with a complete
pregnancy history describing the outcomes and timing of each of their
pregnancies along with information on the relevant determinants of those
outcomes. Until recently most surveys have specialized in data acquisition
--demographic surveys have obtained pregnancy histories of women, but usually
with little information on the health of children or on the socioeconomic
environment of the household; health-oriented surveys have generally
collected information based on samples of children, births or pregnancies,
so it has not been possible to ascertain from them the degree of censoring
associated with birth selectivity in assessing the impact of health programs
or parental variables.

    The 1976-77 Malaysia Family Life Survey is one of the few surveys to
provide information on the human capital of children based on a sample of
(ever-married) women. Information is provided on the weight at birth for

all live births and, for children aged six and above, on parental expectations of <u>completed</u> schooling attainment. Because weight at birth is available in the data set whether or not the child subsequently dies, we can distinguish the effects of birth selectivity from those of mortality selectivity, if any. Because there is information on child deaths, we can compare estimates obtained with and without the population of children born who died and thus compare the effects of both birth and mortality selection on the determinants of birthweight.

The availability of schooling expectation information permits the isolation of the influence of the selectivity of fertility on schooling decisions because it allows estimates of schooling decisions for young children, for whom the probability of having left home is small, while also avoiding sample censoring due to schooling not being completed. The expected schooling attainment estimates will be afflicted, however, by mortality selection, since the expectation information is available only for children alive at the survey date. However, the infant mortality rate in Malaysia is not very high, especially compared to many low-income countries--less than four percent of children born within five years of the survey had died.[6] And we can assess the effects of mortality selection based on the results from the birthweight analysis, as noted.

We constructed two samples for estimation from the data set. For the analysis of birthweight, we selected a sample of ever-married women aged 15-50 who reported information on their husband's income. For these women we obtained information on the birthweight of their latest child born within the last five years prior to the survey. Thus, we examine the determinants of the weight of birth of children aged less than four years at the time of the (first round) of the survey. We chose this age group of children to

22

minimize recall error in both birthweight and mortality. The first three columns of Table 1 report the characteristics of the parents and children in the samples defined by sample selection rules based on birth and/or survival criteria. Note that it is the sample of children defined by the criteria in the third column (children born and surviving) that is most typically used to examine the determinants of child health. Such children represent only 62 percent of all households that could have had a child of the predefined age in the Malaysia setting.

To analyze the determinants of schooling, we selected a sample of women aged 25 to 50 with husband present and obtained information on schooling expectations (or actual schooling if schooling was completed) for all of their surviving children, if any, born six to ten years prior to the survey.[7] Less than three percent of the sample of women who had had a birth six to ten years before the survey did not also have a surviving child, so that sample censoring due to mortality does not appear to be severe.[8] However, as shown in the last two columns of Table 1, the potential for birth selectivity is high--over 27 percent of the women "at risk" of a pregnancy did not have a child of the pre-selected age range.

Table 2 reports the normal maximum-likelihood estimates of the determinants of having a child of the relevant age among the sample of married, spouse-present women in the birthweight and schooling-attainment samples. Consistent with most other fertility studies, the results indicate that fertility is significantly related to the observable characteristics of parents. In particular, more educated women are significantly less likely to be represented among the children born, as are older women. The results also suggest that households with bathing facilities are significantly more likely to be represented among the population of children born in the

Table 1

Mean Characteristics of Malaysian Women, Mothers and Children

by Sample Selection Criteria: Married, Spouse Present Women in 1976

Characteristic	Woman Aged 15-50 (1)	Sample Selection Criteria (1) and Live Birth in Last 5 Years (2)	(2) and Surviving Children (3)	Woman Aged 25-50 (4)	(4) and Surviving Child Born 6-10 Years Ago
Women:					
Schooling attainment (years)	3.74 (3.64)[a]	3.80 (3.43)	3.90 (3.42)	3.35 (3.54)	3.01 (3.25)
Age	33.3 (8.3)	31.1 (6.8)	31.1 (6.8)	35.7 (6.80)	35.5 (6.16)
Husband's earnings in last month	793 (860)	760 (749)	776 (758)	828 (875)	763 (767)
Chinese (percent)	12.1	11.2	10.5	11.9	12.8
Indian (percent)	38.3	37.4	37.6	40.5	38.5
Child (Latest pregnancy):					
Birthweight (ounces)	-	109 (19.1)	110 (18.4)	-	-
Expected schooling attainment (years)	-	-	-	-	11.4 (4.30)
Gender (female = 1)	-	.464	.468	-	.453
Number of women	901	563	541	730	530
Number of children	563	563	541	530	530

Source: 1976-77 Malaysia Family Life Survey.

a. Standard deviation in parentheses.

Table 2
Maximum Likelihood Probit Estimates: Determinants of
Probability of Birth in 5-Year Period

	Sample	
Variable	Women Aged 15-50, Birth in Last 5 Years	Women Aged 25-50, Birth 6-10 Years Ago
Woman's age[a]	48.7	78.1
	(10.4)	(7.68)
Woman's age squared[a]	-7.39	-11.1
	(10.8)	(7.73)
Woman's schooling[a]	-.182	-.172
	(3.57)	(3.09)
Husband's earnings ($\times 10^{-3}$)[a]	.0904	.00320
	(1.73)	(0.05)
Woman Chinese	-	-.0351
		(0.30)
Woman Indian	-.245	.206
	(1.67)	(1.13)
Distance to family planning center[a]	-.0268	-.0295
	(1.01)	(1.01)
No family planning clinic in village	-.317	-.191
	(1.12)	(0.61)
House has no bathing facilities	-.635	-.385
	(3.66)	(2.03)
Constant	-78.8	-136.5
	(9.97)	(7.57)
χ^2	239.3	86.4
Number of women	901	730

a. Variable in log form.
b. Absolute value of asymptotic t-ratio in parentheses.

specified calendar interval, but the proximity of family planning clinics does not appear to significantly influence the probability of a birth in either Malaysia sample. The children in Malaysia who were aged 0-4 or 6-10 in 1976-77 were thus not randomly selected from the households of women of child-bearing age, at least with respect to the commonly-measured characteristics of these households.

b. <u>Results: Birthweight</u>

The least squares estimates of the determinants of birthweight for the sample of surviving children aged 0-4 are reported in the first column of Table 3. In this sample, based on selection criteria commonly used in studies of child health, the set of coefficients associated with parental characteristics (mother's age and schooling and father's monthly earnings) is not jointly statistically significant. In the second column, children who died subsequent to their birth are added to the sample of surviving children and the equation is reestimated using least squares. In this sample of live births, used in all studies of birthweight, mother's schooling has a positive effect on birthweight that is significant at the 0.10 level.

In columns three and four, we report the parameter estimates and test statistics from the normal maximum-likelihood selectivity model applied to surviving children and to all live births, respectively. Application of the test of normality using the semi-parametric kernel estimates leads to non-rejection of the normality assumption.[9] The last column is thus appropriately "corrected" for both birth and mortality selectivity and the third column is only afflicted by mortality selectivity. Comparisons of the estimates in columns one and three thus indicate the effects of taking into account the selectivity of fertility on the determinants of birthweight

Table 3
Determinants of Log of Birthweight: Corrected and Uncorrected
for Birth and Mortality Selectivity

Variable/ Estimation procedure	Birth and Mortality Selected OLS	Birth Selected only OLS	Mortality Selected only Normal ML	No Selectivity Normal ML
Mother's age[a]	.965 (0.99)[b]	.871 (0.86)[b]	-3.47 (3.44)[c]	-4.30 (4.01)[c]
Mother's age squared[a]	-.137 (0.95)	-.118 (0.79)	.538 (3.58)	.668 (4.20)
Mother's schooling[a]	.00880 (1.09)	.0140 (1.66)	.0221 (2.32)	.0300 (3.01)
Husband's earnings $(x10^{-3})$[a]	.00174 (0.19)	.00340 (0.35)	-.00709 (0.66)	-.00501 (0.44)
Child female	-.0358 (2.35)	-.0311 (1.94)	-.0287 (1.77)	-.0252 (1.51)
Mother Indian	-.0775 (3.05)	-.119 (4.52)	-.0481 (1.80)	-.0785 (2.77)
Distance to family planning clinic $(x10^{-3})$[a]	.818 (0.17)	5.95 (1.25)	.00241 (0.37)	.00651 (0.96)
No family planning clinic in village	-.0442 (0.78)	-.113 (2.03)	-.0116 (0.21)	-.0620 (1.54)
House has no bathing facilities	-.0489 (1.45)	-.0376 (1.07)	.0129 (0.35)	.0287 (0.73)
Constant	3.04 (1.85)	3.13 (1.82)	10.3 (6.10)	11.7 (6.42)
F	2.45	4.44	-	-
-Log likelihood	-	-	1168.1	1240.2
ρ	-	-	-.797 (18.2)	-.849 (26.7)
$\chi^2(8)$-Normality test	-	-	-	0.98
Number of children	541	563	541	563
Number of mothers	901	901	901	901

a. Variable in log form.
b. Absolute value of t-ratio in parentheses in column.
c. Absolute value of asymptotic t-ratio in parentheses in column.

based on samples of surviving children; comparisons of columns two and four indicate the effects of neglecting birth selectivity based on samples of live births, and comparisons of the estimates in columns three and four indicate the effects of censoring due solely to mortality (when birth selectivity is taken into account).

Comparisons of all of the estimates across columns in Table 3 indicate that birth selection is strong, negative and statistically significant-- among observably-identical women, those more likely to have a child have less-endowed children and/or tend to allocate less resources (prenatally) to children; the estimated correlations between the fertility and birthweight disturbances are between -.80 and -.85. Evidently, as a consequence of negative fertility selection and of more educated women being less likely to have a child in any given time interval, the effect of mother's schooling on birthweight net of fertility selectivity is more than double that obtained applying least squares to an evidently non-random sample of live births (column four versus column two).

The estimated effects of mother's age on birthweight are also changed dramatically when birth selectivity is taken into account, again because of the strong effects of age on the probability of giving birth. The estimates of the effect of the mother's age on the probability of a birth occurring in the last five years in Table 2 indicate that age has a positive effect on fertility until age 27 and then declines. The finding of negative birth selectivity suggests that estimated age effects on birthweight will thus be biased downward at early ages and then biased upward due to a change in the composition of mothers by age. The least squares estimates of birthweight (columns one and two) indicate that birthweight rises with mother's age, at a declining rate. This finding conforms to the conventional notion that

25

birth postponement augments child survival. However, this result is evidently due to birth selectivity; the normal maximum-likelihood estimates corrected for such selectivity indicate that birthweight is relatively high at the youngest ages, declines with age until age 26, and then rises again. Figure 1 plots the observed relationship between the age of mothers and their children's birthweight based on the least squares estimates from the sample of births, which reflects both life-cycle and population composition effects, and the "true" life-cycle pattern of birthweight for a randomly-chosen woman as indicated by the selectivity-corrected estimates. The latter does not support the notion that postponement of a birth necessarily augments birthweight.

Comparisons of the birth-selectivity corrected estimates across the live-birth and surviving children samples (columns four and three) indicate that mortality selection also affects inferences about the determinants of birthweight, although not as importantly as does birth selectivity. The biases due to censoring by death appear to be qualitatively similar to those due to birth selection--the effect of mother's schooling is downward biased (by 26 percent) and the linear (in logs) age and squared age terms are biased upward and downward respectively. Inferences about the effects of these variables on birthweight, net of birth selectivity, are not, however, substantially altered by the presence of mortality selectivity at least at the relatively low levels of mortality in the Malaysia settings.

Of the environmental variables, the selectivity of fertility appears to be wholly responsible for the marginally statistically significant positive association between household bathing facilities and birthweight and evidently masks a marginally significant positive effect of the presence of a family planning clinic on birthweight.[10]

Figure 1. Expected Birthweight of Child by Age of Woman
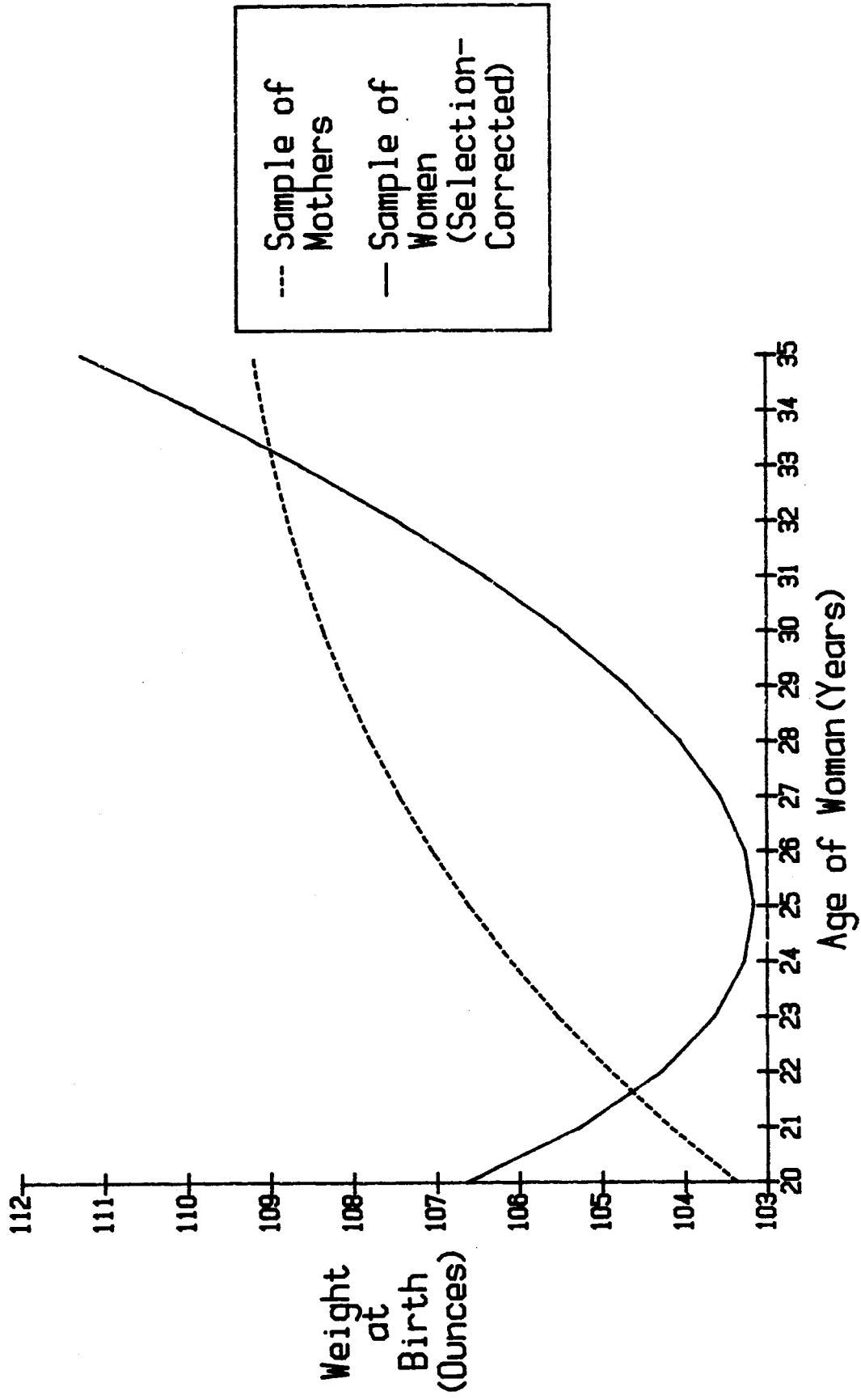
c. <u>Results: Schooling expectations</u>

Column one of Table 4 reports the least squares estimates of the determinants of parental schooling expectations based on the efficient schooling model, with the age and household variables excluded. To be fair to the model, we have included the schooling attainment of the mother and husband's earnings because these variables could reflect interhousehold differences in endowments and thus may be associated with the schooling of children even if schooling is allocated efficiently; i.e., according only to rate of return criteria. The least squares estimates of the full (inefficient) schooling specification is reported in column two. The F-statistic on the set of variables excluded under the efficient schooling model indicates that we cannot reject that model at conventional significance levels based on the least-squares estimates.

The normal maximum-likelihood estimates of the efficient schooling specification corrected for birth selection are reported in the third column of Table 4. These results, consistent with those for birthweight, indicate a strong degree of negative fertility selectivity; the estimated correlation in the errors across the fertility and schooling expectation equations is -0.99.[11] Moreover, our kernel estimates of this model from the semi-parametric specification indicate non-rejection of the normal distribution assumption. However, the normal maximum-likelihood estimates and the semi-parametric estimates of the full schooling specification, corrected for birth selection, indicate, respectively, rejection of the efficient schooling model and of the normal distribution assumption. These results thus highlight the extent to which lack of attention to the selectivity of fertility can lead to misleading conclusions, in this case the false acceptance of the efficient schooling model (and the normality of the joint

Table 4
Determinants of the Log of Expected Children's Schooling Attainment:
Corrected and Uncorrected for Birth Selectivity

Variable/ Estimation procedure	Efficient Schooling Model, Birth Selected OLS	Full Schooling Model, Birth Selected OLS	Efficient Schooling Model, Selectivity Corrected Normal ML	Full Schooling Model, Selectivity Corrected Normal ML
Mother's age[a]	-	-1.36 (0.24)[b]	-	-23.5 (3.90)[c]
Mother's age squared[a]	-	.242 (0.30)	-	3.37 (3.98)
Mother's schooling[a]	.117 (4.83)[b]	.143 (5.44)	.146 (5.58)[c]	.180 (6.01)
Husband's earnings $(\times 10^{-3})$[a]	.0119 (0.42)	.0202 (0.69)	.0198 (0.60)	.0217 (0.62)
Child female	-.0216 (0.43)	-.0205 (0.40)	-.0105 (0.27)	-.00549 (0.13)
Mother Chinese	-.272 (4.76)	-.261 (4.47)	-.204 (3.14)	-.213 (3.20)
Mother Indian	-.119 (1.49)	-.123 (1.51)	-.144 (1.56)	-.156 (1.53)
Distance to family planning clinic[a]	-	.0100 (0.69)	-	.0170 (0.91)
No family planning clinic in village	-	.0193 (0.11)	-	.0616 (0.39)
House has no bathing facilities	-	-.154 (1.36)	-	.0686 (0.71)
Constant	2.40 (27.1)	4.16 (0.42)	2.63 (29.7)	43.4 (4.07)
F	10.0	5.69	-	-
-Log likelihood	-	-	1033.0	1014.5
F(5,519)-Efficient schooling test	-	1.31	-	-
$\chi^2(5)$-Efficient schooling test	-	-	-	38.1
χ^2(d.f.)-Normality test	-	-	0.003(5)	35.5(9)
ρ	-	-	-.991 (181.2)	-.993 (185.3)
Number of children	530	530	530	530
Number of mothers	730	730	730	730

a. Variable in log form.
b. Absolute value of t-ratio in parentheses in column.
c. Absolute value of asymptotic t-ratio in parentheses in column.

error distribution).

The rejection of the assumption of the normality of the distribution of disturbances for schooling, and of the exclusion restrictions permitting identification of the determinants of schooling implied by the efficient schooling model without an explicit parameterization of the disturbance distribution, led us to search for additional parametric distributional assumptions that would provide identification and also would not be rejected. We followed the approach of Lee (1982, 1983) by allowing the distributions of the errors ϵ_{hi} and ϵ_{fi} to be correlated but specifying only their marginal distributions, $G(\epsilon_h)$ and $F(\epsilon_f)$. Each of these distributions can be transformed into a standard normal variable by applying an inverse normal transformation

$$(36) \quad \epsilon_f^* = J_1(\epsilon_f) = \Phi^{-1}(F(\epsilon_f))$$

$$(37) \quad \epsilon_h^* = J_2(\epsilon_h) = \Phi^{-1}(G(\epsilon_h)) \; .$$

By assuming that the transformed variables are jointly normal with zero means, unit variances and correlation coefficient ρ, a bivariate distribution having these distributions is then specified as

$$(38) \quad H(\epsilon_f, \epsilon_h, \rho) = B[J_1(\epsilon_h), J_2(\epsilon_f), \rho]$$

where $B(.,.,\rho)$ is the standard normal bivariate distribution. The log likelihood function based on this specification is

$$(39) \quad \ell nL = \sum_{i=1}^{N} \left\{ I_i \ell n\Phi((J_1(X_{fi}\beta_f) - \rho J_2(X_{hi}\beta_h)) / \sqrt{1-\rho^2}) \right.$$

$$+ I_i \ell n[g((Y_i - X_{hi}\beta_h)/\sigma_h)]$$

28

$$- I_i \ell n \sigma_h + (1-I_i) \ell n(1-F(X_{fi}\beta_f)) \Big\} \, ,$$

where g( ) is the density of ϵ_h. This approach was used to estimate, by maximum likelihood, the schooling model under various assumptions on the distribution of the errors. Among the distributions we tried were the Student t distributions with degrees of freedom equal to 3 and 5 and Chi-square distributions with degrees of freedom equal to 3, 5 and 100. The Chi-square test statistics were 27.6 and 28.7 for the t-distributions and 26.6, 28.8 and 41.3, respectively, for the Chi-square distributions, thus indicating rejection of all of these fully parameterized models in favor of an unknown semi-parametric alternative.

4. Conclusion

In this paper we have formulated a simple life-cycle model to illustrate how in a population characterized by heterogeneity in human capital endowments and deliberate control of fertility, levels of human capital will change in response to alterations in the economic environment. The model suggests that there are two mechanisms by which human capital levels are altered. First, the composition of households, classified by human capital endowments, who bear a child in a pre-specified time period will change and, second, among those self-selected households having children, resources allocated to human capital investments will be altered. The latter effect has been the primary concern of studies of the determinants of human capital investment; however, inattention to the first mechanism, the selectivity of fertility, can result in misleading inferences about allocative responses and to inappropriate conclusions about the consequences of policy interventions.

29

Empirical results based on data from a comprehensive survey of ever-married women in Malaysia indicates that fertility is highly selective. In particular, we found that among observationally-identical women, those who tend to have children with higher birthweight, an important predictor of subsequent child development, and who report that they expect their children to obtain higher levels of schooling are significantly less likely to have a child in any given life-cycle period. As a consequence, we found that estimates of the effects of maternal schooling attainment on birthweight obtained without attention to the selectivity of fertility are underestimated by more than 50 percent. Moreover, the common finding that later childbearing (up to a certain point) augments birthweight appears also to be solely a consequence of the negative selectivity of birth rates. The magnitude of the bias in these variables arises in part because both age, in a non-linear manner, and maternal schooling attainment are strong correlates of birth probabilities. However, we also found that the presence of family planning clinics, although evidently not effective in altering birth rates, appeared to augment weight at birth once compositional effects associated with selective fertility were taken into account. Moreover, inattention to birth selectivity led to a false acceptance of the hypothesis that schooling is allocated efficiently by parents.

As is well known, estimates of interesting behavioral parameters that accommodate sample selectivity typically require strong assumptions about the distribution of the unobservable factors characterizing the population that is studied. Following standard practice, we obtained our results corrected for the selectivity of fertility based on an assumption of normality. However, we also used newly-developed selectivity models that do not require any parametric assumptions about the unobservables to test this

30

assumption. We found that we could not reject the hypothesis of normality for the birthweight model or for the efficient schooling model. However, the latter model, in its parametric normal form, was also rejected and the tests employing the semi-parametric estimates indicated as well the rejection of normality for the full schooling model in which parental preferences also shape schooling decisions. A number of other parametric alternatives to normality were also rejected.

Thus, our selectivity-corrected results for birthweight appear to be robust, but our inferences about the determinants of schooling investments in the full model cannot be relied upon with confidence. Our findings suggest, however, that inferences about how levels of health or schooling are related to parental characteristics, parental resource allocations, or to policy interventions must be attentive to the selectivity of fertility. In environments where fertility rates are extremely low, such as in Europe and the United States, caution in drawing inferences about the determinants of human capital investments may be particularly warranted.

Footnotes

1. We have assumed, for simplicity, that parental utility is only affected by the human capital of the child in the second period of its life. This assumption does not alter the main conclusions of the model.

2. We have assumed that parents know the endowment of the child in advance of its birth. It is straightforward to recast the model to incorporate uncertainty with respect to endowments, as in Rosenzweig (1986).

3. Heterogeneity in preferences could also be added to the model, although there is less direct evidence on its existence compared to endowment heterogeneity, which has been found in a number of health production function studies. While we provide an endowment interpretation to the model, we do not (and cannot) interpret our empirical findings as indicating a particular source of heterogeneity. Estimation of the human capital production function in a selectivity framework could provide a means of identifying whether or not endowments influence birth rates.

4. The lack of exclusion restrictions is not confined to models incorporating the selectivity of fertility. Life-cycle labor supply models incorporating endogenous savings, fertility and/or human capital investments do not deliver the result that non-earnings income or fertility are exogenous variables that can identify wage equations corrected for the selectivity of labor force participation. Yet such variables are commonly employed as identifying instruments in studies of the determinants of female wage rates.

5. We obtained this finding from the 1981-82 Nutrition Survey of Bangladesh, a national probability sample of 385 households located in 15 villages. In the sub-sample of married women aged 21-29, 16.7

percent had not given birth in the five years preceding the survey. In the Malaysia survey that we use to study the determinants of human capital, described below, 22.5 percent of ever-married women in the same age group had not given birth in the pre-survey quinquennium. For women aged 15-50, the proportions were 35.5 and 37.9, respectively, in the two populations.

6. The infant mortality rate in Malaysia in 1984 was 38 (per 1000 births), as compared to 124 in Bangladesh and 11 in the United States (1987 World Bank Report).

7. Less than ten percent of the children in the age group had completed their schooling.

8. However, the lower death rate for the older compared to the younger cohort of children suggests the presence of recall error in reports of mortality; the expected cumulative death rate for the children aged 5-10 should be just over four percent based on Malaysia vital statistics.

9. Because the sole identified coefficient in the semi-parametric model, the gender of the child, is only marginally significant (t=1.51) in the normal maximum-likelihood model, it is not surprising that the test of normality based solely on the normal and semi-parametric estimates of this parameter also indicates non-rejection of normality ($\chi^2(1)$=0.01).

10. Family planning clinics in Malaysia are known to disseminate information on child health and prenatal care. Our estimates indicate that such facilities are more successful in improving child health than in reducing fertility.

11. To insure that our maximum-likelihood estimates were not local maxima, we tried a number of different starting values. The estimates converged to the results reported in all cases. Inefficient two-step

33

estimates also indicated values for ρ similar to the maximum-likelihood

estimates reported.

References

Becker, Gary S. and Tomes, Nigel, "An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility," _Journal of Political Economy_ 87, December 1979, 1153-89.

Behrman, Jere R., "Intrahousehold Allocation of Nutrients in Rural India: Are Boys Favored?" Do Households Exhibit Inequality Aversion?" _Oxford Economic Papers_, forthcoming.

Behrman, Jere R. and Deolalikar, Anil, "Health and Nutrition," in H. Chenery and T. N. Srinivasan, eds., _Handbook of Development Economics_, Amsterdam: Elsevier Science, 1988, 633-711.

Behrman, Jere R., Pollak, Robert A. and Taubman, Paul, "Parental Preferences and Provision of Progeny," _Journal of Political Economy_ 90, February 1982, 52-73.

Birdsall, Nancy, "Economic Approaches to Population Growth," in H. Chenery and T. N. Srinivasan, eds., _Handbook of Development Economics_, Amsterdam: Elsevier Science, 1988, 478-542.

Grossman, Michael and Joyce, Theodore J., "Unobservables, Pregnancy Resolutions, and Birthweight Production Functions in New York City," NBER Working Paper No. 276, October 1988.

Heckman, James J., "The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models," _Annals of Economic and Social Measurement_ 5, 1976, 475-92.

Heckman, James J., "Sample Bias as a Specification Error," _Econometrica_ 47, January 1979, 153-162.

Hotz, V. Joseph and Miller, Robert, "An Empirical Analysis of Life-Cycle Fertility and Female Labor Supply," _Econometrica_ 56, January 1988, 91-118.

Ichimura, Hidehiko, <u>Estimation of Single Index Models</u>, Ph.D. Thesis, Department of Economics, MIT, 1987.

Ichimura, Hidehiko and Lee, Lung-Fei, "Semiparametric Estimation of Multiple Index Models: Single Equation Estimation," <u>International Symposia in Economic Theory and Econometrics</u>, Cambridge University Press, forthcoming.

Lee, Lung-Fei, "Some Approaches to the Correction of Selectivity Bias," <u>Review of Economic Studies</u> 49, 1982, 355-372.

_____, "Generalized Econometric Models with Selectivity," <u>Econometrica</u> 51, March 1983, 507-512.

Olsen, Randall J., "A Least Squares Correction for Selectivity Bias," <u>Econometrica</u> 48, May 1980, 1815-20

Olsen, Randall and Wolpin, Kenneth I., "The Impact of Exogenous Child Mortality on Fertility: A Waiting Time Regression with Dynamic Regressors," <u>Econometrica</u> 51, May 1983, 731-749.

Pitt, Mark M. and Rosenzweig, Mark R., "Health and Nutrient Consumption Across and Within Farm Households," <u>Review of Economics and Statistics</u>, November 1985, 212-223.

Rosenzweig, Mark R., "Birth Spacing and Sibling Inequality: Asymmetric Information within the Family," <u>International Economic Review</u> 27, February 1986, 55-76.

Rosenzweig, Mark R. and Schultz, T. Paul, "Estimating a Household Production Function: Heterogeneity, the Demand for Health Inputs, and their Effects on Birthweight," <u>Journal of Political Economy</u> 91, October 1983, 723-746.

Schultz, T. Paul, "Education Investments and Returns," in H. Chenery and T.

  N. Srinivasan, eds., <u>Handbook of Development Economics</u>, Amsterdam:

  Elsevier Science, 1988,

Silverman, B. W., <u>Density Estimation for Statistics and Data Analysis</u>,

  London: Chapman and Hall, 1986.

Strauss, John, "Households, Communities and Preschool Children's Nutrition

  Outcomes: Evidence from Rural Cote d'Ivoire," 1988, <u>Economic

  Development and Cultural Change</u>, forthcoming.