ELSEVIER Journal of Econometrics 77 (1997) 209–235

JOURNAL OF Econometrics

The effects of improved nutrition, sanitation, and water quality on child health in high-mortality populations

Lung-fei Lee[a], Mark R. Rosenzweig[*,b], Mark M. Pitt[c]

[a] *Department of Economics, The Hong Kong University of Science and Technology*
[b] *Department of Economics, University of Pennsylvania, Philadelphia, PA 19104, USA*
[c] *Department of Economics, Brown University, Providence, RI 02912, USA*

Abstract

A framework is set out for estimating the effects of interventions on child health that considers changes in the allocation of family resources, who among children survive (survival selectivity), and changes in the health of surviving children net of family resources. Estimates based on structural-equations semi-parametric models applied to data describing households from rural areas of two low-income countries indicate that conventional reduced-form estimates understate the effectiveness of improving sanitation facilities. This is due to the reduced allocation of household resources to children in households with better facilities but not to mortality selection, which is negligible.

Key words: Survival selection; Structural equations; Semi-parametric estimators; Anthropometric measures; Rural populations
JEL classification: I12; I18; C14; O15

1. Introduction

Considerable attention has been paid in recent years to the influence of public programs, the socioeconomic status of parents, and parental behaviors on both the health and mortality of children (Behrman and Deolalikar, 1988a). Studies by Behrman and Deolalikar (1988b), Strauss and Mehra (1988), Thomas et al. (1988), Pitt, Rosenzweig, and Hassan (1990a, 1990b) are recent examples of studies of child health in low-income countries in which the health outcomes of children in the sampled populations are related to the availability of health

* Corresponding author.

This research was supported in part by NICHD grant HD-26384 and NSF grant SBR-9223325. We are grateful for the comments of three referees and the Editor.

PII S0304-4076(96)01813-1

programs and facilities in the area in which they reside, the prices of food and other goods, and observed exogenous characteristics of the individual and his or her parents. From these reduced-form equations, inferences are commonly drawn about the relative effectiveness of a variety of interventions, including investments in health infrastructure, in influencing measured health outcomes, such as frequency of illness or height and weight. Many studies have also been carried out that examine the impact of similar variables on the survival of children (e.g., DaVanzo, 1988; Olsen and Wolpin, 1983). These latter studies as well as those concerned with health status indicate the importance of mother's schooling, but few general conclusions can be made about the effects of particular interventions on health. However, studies of the determinants of health generally do not consider mortality; they are based on samples of surviving children. To the extent that children who survive differ systematically from those who do not, as suggested by the literature on child survival, inferences about the effects of raising adult schooling levels or of particular program interventions based on samples of surviving children, as is conventionally done, may be misleading, particularly in high-mortality populations. Indeed, it is likely that such estimates will lead to understatements of the effectiveness of health-augmenting interventions.

The average health of a surviving population may change in response to an intervention that improves health facilities for two distinct reasons: (i) the intervention may alter the investments made by parents in children and (ii) changes in mortality induced by the intervention may alter the average health of the surviving population by allowing those individuals with higher inherent probabilities of illness (those more frail or less healthy) to survive. Reduced-form estimates of the determinants of health from samples of surviving children provide the effects of interventions on the average health of surviving children inclusive of both of these effects. To assess the effectiveness of an intervention thus requires that both responses, of parents and of survival, be taken into account. It is recognized that individuals have different levels of inherent healthiness or frailty (see Vaupel, 1988). Such endowments of health, which include individual-specific inherent healthiness and also environmental influences that affect health regionally but cannot be influenced by households, might be expected to be importantly related to the likelihood of infant death.

In this paper, we set out a framework for estimating the effects of health interventions that improve the health infrastructure that takes into account three responses: a) changes in the allocation of family resources (i.e., nutrition) to children, b) changes in who among children survive, and c) changes in the health of the children who survive net of family resource allocation. We pay particular attention to the endogenous allocation of household resources to children and to the selectivity effects of alterations in the health infrastructure via their effects on child survival. We obtain estimates of the effects of improved nutrition and of water sources and waste disposal facilities on anthropometric measures of child health using data sets from two countries – Bangladesh and the Philippines – that provide the requisite detailed fertility, mortality, nutrition, and health information necessary to implement the full model. Changes in water and waste facilities can potentially affect the spread of disease, and thus improvements in such facilities can potentially reduce illness and mortality and lead to better health among survivors. Yet most studies, based on reduced-form estimates applied to surviving children, do not yield significant health effects.

Econometric techniques for obtaining estimates corrected for nonrandom sample censoring have been applied to a variety of topics in the economics and sociological literatures in the last decade, but have never been applied comprehensively to child health, where problems of selectivity would appear to be potentially severe, particularly in populations with high rates of child mortality. Moreover, the results obtained may be sensitive to the assumptions made about error distributions, which do not come from economic (or any other) theory. Standard models of selectivity have considered only a limited number of distributions (e.g., normal (Heckman, 1979)). Recently, however, methods of estimation for selectivity models have been developed (Robinson, 1988; Powell, 1987; Ichimura and Lee, 1991) that yield consistent estimates of the behavioral parameters and permit classical hypothesis tests without imposing any distributional assumptions on disturbances. Moreover, such procedures allow tests of the distributional assumptions commonly employed in selection models. We employ newly-developed semi-parametric estimators in the context of a structural equations system to estimate (i) the effects of water and sanitation facilities on child survival and (ii) the effects of increased calorie consumption and improvements in the health infrastructure on measures of children's nutritional status net and gross of the effects of these interventions on child survival.

2. Theory

To fix ideas about the roles of heterogeneity in health, health investments, and mortality selectivity in assessing the effects of improvements in the health infrastructure, it is useful to examine a simple model of health determination. The central feature is the health technology, given by $h_j^* = h(z_j, y, v_j)$, in which $v_j = \mu + e_j$, h_j^* is the health of a person in period j of his/her life, z_j represents the resources or inputs (nutrition) allocated to health in that period, y represents the health infrastructure (water quality, sanitation), μ represents the health endowment, that component of an individual's health that is not subject to intervention and is not observable to the data analyst, and e_j represents a random health shock (illness) at period j.

The 'effectiveness' of a change in y is given by $\partial h/\partial y$ when no other changes in the health inputs occur. The observed or 'reduced-form' effect of a change in y on an individual's health is obtained by dh_j^*/dy $(= \partial h/\partial y + (\partial h/\partial z_j)(dz_j/dy))$. This indicates that the reduced-form relationship between y and h_j^* depends not only on the effectiveness of y but also on the response of the allocated inputs z_j to the change in the health infrastructure. If, for example, $\mathrm{d}z/\mathrm{d}y < 0$, less resources are allocated to health where the infrastructure is more favorable to health, then the 'reduced-form' relationship between y and h^* will understate the effectiveness of the intervention. Inferences about the effectiveness of an intervention on the health of any individual, given by direct estimates of the health technology, are also affected by the possibility of selective censoring due to mortality. Mortality censoring is nonrandom if survival depends on health, which in part depends therefore on μ. It can be shown as long as the intervention positively affects first-period health, survival censoring will always lead to a reduction in the average health endowment of the surviving population, as the marginal survivor's health endowment will always be lower. Thus the reduced-form association between the *average* health of survivors and any health-improving intervention depends on both how the intervention affects resource allocations to health and on the association between health endowments and resource allocations.

To understand the potential biases in commonly-used regression estimates of the determinant of child health, it is necessary to embed the health technology and survival in a behavioral model. Consider a simple two-period dynamic model in which health is by the production function and children may die when $h_2^* < 0$. Parents have only one child and choose the amount of z to allocate to the child's health. Maximal expected 'lifetime' utility in the first period is

$$V = \max_{z_1} \mathrm{E}\{U(h_1(z_1, y, v_1), c_1) + \delta[\theta(h_1)V^s + (1 - \theta(h_1))U(0, I_2)]\}, \quad (2.1)$$

where $V^s = \max_{z_2} U(h_2(z_2, y, v_2), c_2)$, $c_j = I_j - p_j z_j$, and $c_j =$ period-j parental consumption, $I_j =$ period-j income, $\delta =$ discount rate, and $\theta =$ probability of survival. In this model, first-period allocations are made to the child who is born based on survival expectations, given all available information, including the health endowment, the price of z and the levels of y and of income in both periods. In the second period, however, uncertainty is resolved, and the allocation problem is static. Consider linear representations of the reduced-form health equation in the first period determining survival, of the health production technology for period-2 health, and of the demand equation for the z inputs in the second period:

$$h_{1i}^* = x_{1i}\bar{\eta} + \gamma_1\mu_i + e_{1i} = x_{1i}\bar{\eta} + \varepsilon_i, \quad (2.2)$$

$$h_{2i}^* = z_{2i}\beta_1 + x_{2i}\beta_2 + \mu_i + e_{2i} = z_{2i}\beta_1 + x_{2i}\beta_2 + v_i, \quad (2.3)$$

$$z_{2i} = x_{3i}\pi + \delta\mu_i + w_i = x_{3i}\pi + u_i, \quad (2.4)$$

where i indexes any individual who is born, x_{1i} is a vector of all the observed exogenous variables in the model (y, p_1, p_2, I_1, and I_2), $x_{2i} = y$, x_{3i} is a vector including y, p_2, and I_2, and we have written the health endowment in terms of second(final)-period health. The first equation is obtained by solving for the optimal z_{1i} from (2.1) in terms of the exogenous variables known to the family and replacing z_{1i} in the production technology with the optimal z_{1i} decision rule. The β_2 correspond to the effectiveness of the health intervention variables, the β_1 measure the effect of the endogenous health input z, and $\bar{\eta}$ reflects the effects of x_{1i} on first-period health and thus survival into the second period. Eq. (2.4) is the second-period decision rule for z, to which we have added an error term w that reflects the possibility of measurement or other errors.

This model suggests that the health input z_{2i} will be correlated with v_i ($\delta \neq 0$ in (2.4)), given parents' knowledge or anticipation of μ_i, so that estimation of (2.3) by least squares provides biased estimates of the β. The censoring associated with health being observed (by the researcher) only when a child who is born survives also induces a bias. It is easiest to show this assuming all stochastic terms are jointly normally distributed. If z_{2i} is replaced by the fitted value of z_{2i} ($\hat{z}_{2i} = x_{3i}\pi$) from (2.4). The reduced-form health regression function (Heckman, 1979) for the surviving population in the second period is

$$\mathrm{E}(h_{2i}^*|\hat{z}_{2i}, x_{1i}, h_{1i}^* > 0) = \hat{z}_{2i}\beta_1 + x_{2i}\beta_2 + \mathrm{cov}(\gamma_2\mu_i + w_i\beta_1 + e_{2i}, \varepsilon_i)\lambda_i, \quad (2.5)$$

where $\gamma_2 = \delta\beta_1 + 1$ and λ_i is the inverse Mills ratio evaluated at $x_{1i}\bar{\eta}$ with $\mathrm{var}(\varepsilon_i)$ normalized to unity. Estimating (2.5) based on the surviving population without taking into account mortality selection is equivalent to omitting the λ-term. The effect of a change in an x_2 on h_2^* in the survivors sample is given by

$$\frac{\partial \mathrm{E}(h_{2i}^*|\hat{z}_{2i}, x_{1i}, h_{1i}^* > 0)}{\partial x_{2i}} = \beta_2 - \bar{\eta}_2 A_i[\gamma_1\gamma_2 \mathrm{var}(\mu)], \quad (2.6)$$

where $A_i = \lambda_i^2 + x_{1i}\bar{\eta}\lambda_i$ and $\bar{\eta}_2$ is the coefficient of x_{2i} in $\bar{\eta}$ in (2.2), and we have assumed that w, e, and μ are mutually independent. The second term in (2.6) is the bias arising from mortality selection (in the normal case). Because $\lambda_i > 0$ and $A_i > 0$, the bias in the effectiveness parameter is, as expected, negative as long as the intervention represented by x_{2i} both increases survival and augments health, i.e., $\bar{\eta}_2 > 0$ (under the assumption that γ_1 and γ_2 have the same sign). The magnitude of the bias depends positively on i) the amount of endowment heterogeneity ($\mathrm{var}(\mu)$), ii) the effect of x_{2i} in augmenting the probability of survival, and iii) the fraction of births that do not survive, which affects the magnitude of the λ-term and thus the size of the A-term in (2.6). Thus, the bias will be potentially more serious in high-mortality populations, where health interventions are most likely to be thought efficacious. These general features of the bias due to selection will pertain also in cases in which errors are not normally distributed, but the precise conditions will be different.

3. Econometric procedures

Eqs. (2.2)–(2.4) describe a simultaneous equation model with selectivity. The individual i is a survivor if $h_{1i}^* > 0$; otherwise, the individual is deceased. Let d_{1i} be the surviving indicator and $d_{2i} = 1 - d_{1i}$. The disturbances ε_i, u_i, and v_i are assumed to be independent of x_{1i}. While the values of h_2 and z can be observed only for the survivors, x_1 is observed for both the survivors and the deceased. Neither economic theory nor the medical literature provides any guidance on the population distribution of the disturbances (inclusive of health endowments). Yet the results obtained may be quite sensitive to arbitrary distributional assumptions although there are few examples of tests of robustness to distributional assumptions (see, e.g., Mroz, 1987). We thus estimate the model using a semiparametric method, without imposing parametric distributional assumptions. We take into account that the data that we use will contain multiple children from the same households so that the error terms will not necessarily be independent.

3.1. Estimation

The identification of the latent survival equation (2.2) without imposing a parametric assumption on its disturbance requires the presence of a relevant continuous exogenous variables in x_1 for which the coefficient is nonzero and has a known sign (Ichimura, 1993; Klein and Spady, 1993). We adopt the convention that $\bar{\eta} = (1, \eta')'$, where the first variable in x_1 corresponds to the variable for normalization and η is the remaining coefficients (Ichimura, 1993). The identification of the simultaneous equations (2.3) and (2.4) requires in general more stringent conditions than the usual rank identification condition of the classical simultaneous equation model (Lee, 1994). The identification conditions can be simplified under some circumstances. When there exists a relevant exogenous variable which appears only in the survival equation (2.2) but not in the outcome equations (2.3), the identification of the outcome equations requires a similar rank identification condition of the classical linear simultaneous equations model. In the model constraints such as income or wealth do not appear in (2.3), so (2.3) is identified under the classical rank condition.

The survival equation (2.2) can be estimated by the semiparametric maximum likelihood (SML) method originated by Klein and Spady (1993). Choice probabilities $\mathrm{E}(d|x_{1i}\bar{\eta})$, where $d = d_1$ or d_2, given any possible value of η and the 'index' $x_1\bar{\eta}$ can be estimated nonparametrically by the nonparametric regression:

$$\mathrm{E}_n(d \mid x_{1i}\bar{\eta}) = \sum_{j\neq i}^{n} d_j K\left(\frac{x_{1i}\bar{\eta} - x_{1j}\bar{\eta}}{a_n}\right) \Bigg/ \sum_{j\neq i}^{n} K\left(\frac{x_{1i}\bar{\eta} - x_{1j}\bar{\eta}}{a_n}\right), \tag{3.1}$$

where n is the sample size, $K(\cdot)$ is a kernel function, and a_n is a bandwidth parameter. The logarithmic semiparametric likelihood function for a binary response model is

$$\ln L_n(\eta) = \sum_{i=1}^{n} I_X(x_i)\{d_{1i} \ln \mathrm{E}_n(d_1|x_{1i}\bar{\eta}) + d_{2i} \ln \mathrm{E}_n(d_2|x_{1i}\bar{\eta})\}, \tag{3.2}$$

where I_X is the set indicator of a set X of x_1. The set X is a subset contained in the support of the regressors in x_1 which is used to trim the tails of the distribution of continuous variables in x_1. The trimming of regressors is a mathematical trick to overcome some theoretical difficulties in estimating the tails of nonparametric regression functions in (3.1). The proportion of trimming can be arbitrarily small for samples with large size. However, because of the trimming of the regressors, the asymptotic variance of the estimator becomes relatively complicated and there are difficulties in constructing an estimate of the variance. Such difficulties can be eliminated if trimming is applied to the index $x_1\bar{\eta}$. For this reason, we use a two-step estimation procedure to derive an estimator for η. A consistent estimator $\tilde{\eta}$ can first be derived from (3.2). With $\tilde{\eta}$, we trim the tails of the distribution of the estimated index $x_1\bar{\tilde{\eta}}$ and then use a Newton one-step procedure:

$$\hat{\eta} = \tilde{\eta} + D_n^{-1}\frac{1}{n}\sum_{i=1}^{n} I_T(x_{1i}\bar{\tilde{\eta}}) \sum_{l=1}^{2} d_{li}\,\frac{\partial \ln \mathrm{E}_n(d_l|x_{1i}\bar{\tilde{\eta}})}{\partial \eta}\,, \tag{3.2'}$$

where $I_T(x_{1i}\bar{\tilde{\eta}})$ is an indicator which is zero if $x_{1i}\bar{\tilde{\eta}}$ is less than some small lower quantile or greater than some upper quantile of the estimated index, and

$$D_n = \frac{1}{n}\sum_{i=1}^{n} I_T(x_{1i}\bar{\tilde{\eta}})\sum_{l=1}^{2} d_{li}\frac{\partial \ln \mathrm{E}_n(d_l|x_{1i}\bar{\tilde{\eta}})}{\partial \eta}\,\frac{\partial \ln \mathrm{E}_n(d_l|x_{1i}\bar{\tilde{\eta}})}{\partial \eta'}\,.$$

In our empirical study, we have arbitrarily trimmed an upper 4-percentile and a lower 4-percentile of the distribution of $x_1\bar{\tilde{\eta}}$.[1] Details on the computation of asymptotic variances of semiparametric estimates are contained in the appendices.

Each of the input equations in (2.4) for survivors forms a sample selection model. As a convention, sample observations are arranged such that the first n_1 observations, where $n_1 < n$, correspond to sample observations of survivors. These equations can be estimated by a semiparametric least squares (SLS) method (Robinson, 1988; Powell, 1987; Ichimura and Lee, 1991):

$$\begin{aligned}&\min_{\pi} \sum_{i=1}^{n_1} I_T(x_{1i}\bar{\hat{\eta}})(z_i - x_{3i}\pi - \mathrm{E}_n(z - x_3\pi\,|\,x_{1i}\bar{\hat{\eta}}))'\\ &\times(z_i - x_{3i}\pi - \mathrm{E}_n(z - x_3\pi\,|\,x_{1i}\bar{\hat{\eta}})),\end{aligned} \tag{3.3}$$

where

$$\mathrm{E}_n(z - x_3\pi|x_{1i}\bar{\eta}) = \sum_{j\neq i}^{n_1}(z_j - x_{3j}\pi)K\left(\frac{x_{1i}\bar{\eta} - x_{1j}\bar{\eta}}{a_n}\right)\Big/\sum_{j\neq i}^{n_1}K\left(\frac{x_{1i}\bar{\eta} - x_{1j}\bar{\eta}}{a_n}\right).$$

[1] In a Monte Carlo study in Lee (1995), estimates are not sensitive to trimming. Justification for trimming can be found in Lee (1990,1995).

The SLS estimator from (3.3) has a closed-form expression:

$$\hat{\pi} = \left[\sum_{i=1}^{n_1} I_T(x_{1i}\hat{\hat{\eta}})(x_{3i} - \mathrm{E}_n(x_3|x_{1i}\hat{\hat{\eta}}))'(x_{3i} - \mathrm{E}_n(x_3|x_{1i}\hat{\hat{\eta}}))\right]^{-1} \times \sum_{i=1}^{n_1} I_T(x_{1i}\hat{\hat{\eta}})(x_{3i} - \mathrm{E}_n(x_3|x_{1i}\hat{\hat{\eta}}))'(z_i - \mathrm{E}_n(z|x_{1i}\hat{\hat{\eta}})). \tag{3.4}$$

The health equation (2.3) differs from the input equations (2.4) in that the vector z of endogenous variables appears on the right-hand side of (2.3). Let $w = (z, x_2)$. Define the following matrices:

$$\hat{X}_3 = \begin{pmatrix} I_T(x_{11}\hat{\hat{\eta}})(x_{31} - \mathrm{E}_n(x_3|x_{11}\hat{\hat{\eta}})) \\ \vdots \\ I_T(x_{1n_1}\hat{\hat{\eta}})(x_{3n_1} - \mathrm{E}_n(x_3|x_{1n_1}\hat{\hat{\eta}})) \end{pmatrix},$$

$$\hat{W} = \begin{pmatrix} I_T(x_{11}\hat{\hat{\eta}})(w_1 - \mathrm{E}_n(w_1|x_{11}\hat{\hat{\eta}})) \\ \vdots \\ I_T(x_{1n_1}\hat{\hat{\eta}})(w_{n_1} - \mathrm{E}_n(w_{n_1}|x_{1n_1}\hat{\hat{\eta}})) \end{pmatrix},$$

and

$$\hat{H}_2^* = \begin{pmatrix} I_T(x_{11}\hat{\hat{\eta}})(h_{21}^* - \mathrm{E}_n(h_{21}^*|x_{11}\hat{\hat{\eta}})) \\ \vdots \\ I_T(x_{1n_1}\hat{\hat{\eta}})(h_{2n_1}^* - \mathrm{E}(h_{2n_1}^*|x_{1n_1}\hat{\hat{\eta}})) \end{pmatrix}.$$

The $\beta = (\beta_1', \beta_2')'$ can be estimated by

$$\hat{\beta} = [\hat{W}'\hat{X}_3(\hat{X}_3'\hat{X}_3)^{-1}\hat{X}_3'\hat{W}]^{-1}\hat{W}'\hat{X}_3'(\hat{X}_3'\hat{X}_3)^{-1}\hat{X}_3'\hat{H}_2^*. \tag{3.5}$$

This estimator can be interpreted as being derived from a two-stage semiparametric least squares (TSLS) procedure. The reduced-form equations z in (2.4) are first estimated by SLS in (3.3). Predicted values of $z - \mathrm{E}_n(z|x_1\hat{\hat{\eta}})$ are used as instrumental variables for $z - \mathrm{E}_n(z|x_1\hat{\hat{\eta}})$ to derive the estimator $\hat{\beta}$ in (3.5) (Lee, 1994).

The semiparametric estimators in (3.2′), (3.4), and (3.5) are $\sqrt{n}$-consistent and asymptotically normal under some requirements on the bandwidth parameter a_n and the order of bias of the kernel function. The kernel function used in our estimation procedure is $K_4(t)$ where $K_4(t) = 2K(t) - K(t/\sqrt{2})/\sqrt{2}$ with $K(t) = \frac{35}{32}(1-t^2)^3$ for $|t| < 1$, and $K(t) = 0$, elsewhere. This kernel function is computationally simple and its design has the effect of eliminating the undesirable biases of nonparametric regression functions (Bierens, 1985). The bandwidth is chosen as $a_n = c/n^{1/5.5}$ where c is a constant factor independent of n. This bandwidth sequence satisfies rate requirements of convergence for semiparametric estimators. [2]

3.2. *Tests for the normality of errors in the survival equation, simultaneity, and selectivity*

If the disturbance ε of (2.1) is assumed to be normally distributed, the survival equation is a probit model and can be estimated by the classical probit maximum likelihood (probit ML) method. The normality assumption of the survival equation can be tested by comparing the SML estimate with the probit ML estimate in a Hausman-type testing procedure. The test statistic, taking into account the family correlation structure, is derived in Appendix B. To set up this Hausman-type test statistic, one has to be careful about different normalization rules. Estimates are renormalized as ratios for compatibility. [3] It is also possible to derive generalized Hausman-type statistics to test (i) whether the z variables in (2.3) are endogenous and (ii) whether there is sample selection based on the semiparametric estimator (3.5). These are also derived in Appendix B.

4. Data

The preceding framework suggests that to appropriately estimate the effectiveness of health interventions based on nonexperimental data in environments in which child survival is not assured, it may be necessary to have or collect information on i) the household resources allocated to children that may affect their health and ii) a complete fertility roster indicating the date of birth of all children and their current survival status. The latter is important because if mortality censoring is significant, it is necessary to construct the 'population' of children inclusive of those who have died. We have identified two data sets that have the requisite information – the 1981–82 Nutrition Survey of Rural Bangladesh

---

[2] We have tried different values of c in a grid search and used the value which provides the largest log-likelihood value (for the survival equation) or the smallest sum of squared residuals (for the outcome equations). Monte Carlo studies in Lee (1990, 1995) provide numerical evidence that this cross-validation procedure is reasonably good. For the estimation of variances of semiparametric estimators, it is desirable to use wider bandwidth. Motivated by Monte Carlo results in Lee (1995), as a rule of thumb, a five times larger bandwidth is used for variance estimation.

[3] Misspecification of the parametric distribution may, in general, cause maximum likelihood parameter estimators of ratios of coefficients in α to be inconsistent (see the Appendix B for the definition of α). Exceptions can occur in special cases where the joint multivariate distribution of the explanatory variables has the property that the expectation of each explanatory variable in x_1 conditional on $\bar{x}_1\alpha$ is a linear function of $\bar{x}_1\alpha$ (Ruud, 1983). These cases include the multivariate normal distribution and distributions in the family of spherical symmetricals. For those special cases, this test statistic may have no power. These special cases are unlikely to be relevant for our data as many explanatory variables are discrete.

(Ahmad and Hassan, 1986) and the 1984–85 IFPRI Bukidnon, Philippines Survey.

The 1981–82 Nutrition Survey of Rural Bangladesh is a national probability sample of 25 households in each of 12 randomly-selected villages plus two villages chosen for their seasonal characteristics and one selected for its proximity to an industrial area (in which 35 households were surveyed). There is thus information on all individuals in 385 households located in 15 villages. Information was obtained by standard survey methods on socioeconomic characteristics, including demographic information, a roster of births and deaths, and activities, educational levels and literacy, landholding, wealth, and wage rates. Information on a time-series of village food prices is also provided. Anthropometric measurements were obtained for all household members, including height, weight, skinfold thickness, and mid-arm circumference. There is also detailed information on the individual distribution of foods obtained over a 24-hour period, based on weight, under the supervision of (female) survey investigators. The Bukidnon survey is a stratified random panel of 448 households in Bukidnon in northern Midanao, Philippines, who were interviewed in four rounds at four-month intervals in 1984–85 as part of a study by Bouis and Haddad (1990). This data set also provides information on births and deaths, anthropometric measures, and individual-specific food intakes, based on 24-hour recall of recipes prepared and proportions eaten, along with detailed socioeconomic and farm production data, prices, and wages.

An important feature of both the Bangladesh and Bukidnon surveys is that there is detailed information on each household's source of drinking water and on waste disposal. The Bangladesh survey provides information on five categories of water sources – tubewell, well, pond, river or canal, and other – and indicates whether the household has a specific place for wastes disposal. Among the sample households with children aged 1 through 14, approximately 63 percent obtain their drinking water from a tubewell, while 20 percent use a pond, river, or canal. Less than 20 percent also have a specific place for the disposal of waste. The Bukidnon data provides six categories of drinking water sources – piped, artesian well (covered), open dry well, improved spring, unimproved spring, and rainwater. Just over 10 percent of households have access to piped water, while approximately 40 percent obtain their drinking water from a spring. There are also four classifications for methods of waste disposal – none (16 percent of households), open pit (37 percent), antipolo (28 percent), and water sealed (21 percent).

Because of the different age structures of death by cause and of activities by children (Pitt et al., 1990), for each data set, we constructed subsamples of all births in each household occurring in the last one to six years and in the last seven to fourteen years, including in each the current anthropometric measures and the individual calorie consumption for each surviving child aged 1 through 6 and 7 through 14, respectively. Age at death information, available

Table 1
Sample characteristics (standard deviations in parentheses)

Variable	Nutrition survey of rural Bangladesh, 1975–76	IFPRI Bukidnon, Philippines survey, 1984–85
Children aged 1–6		
Proportion surviving to survey	0.594	0.924
Mean age of survivors (yrs)	3.67 (1.58)	3.97 (1.61)
Mean caloric intake, survivors	888 (469)	1229 (650)
Mean weight, survivors (kg)	11.2 (2.78)	12.8 (3.00)
Mean head schooling level (yrs)	2.25 (1.51)	6.26 (2.74)
Number of children born	611	837
Number of households	291	423
Children aged 7–14		
Proportion surviving to survey	0.550	0.907
Mean age of survivors (yrs)	10.3 (2.29)	10.5 (2.28)
Mean caloric intake, survivors	1666 (552)	1714 (737)
Mean weight, survivors (kg)	22.5 (68.3)	24.5 (6.78)
Mean head schooling level (yrs)	2.22 (1.41)	5.92 (2.59)
Number of children	693	789
Number of households	331	305

only in the Bangladesh data set, indicates that 55 percent of all deaths among children less than age 15 (312 deaths) occurred in the first year of life, with 25 percent of all deaths occurring between the ages of 1 and 6.[4] Table 1 provides descriptive statistics for all four samples from the two data sets. The potential for selectivity bias is quite strong in the Bangladesh survey area, as less than 60 percent of children born one through six years prior to the survey survived to the survey date. Mortality rates are considerably lower in the Bukidnon survey area. Average calorie intake within each age group of children in Bangladesh is considerably lower than in the other sample. Table 1 also indicates that the number of households represented in each sample is considerably less than the number of births or survivors, so that the effective sample sizes are less than the totals. The estimation procedures that we use take into account this source of nonindependence.[5]

[4]Bangladesh experienced severe flooding and famine in the 1974–75 crop year, which led to elevated levels of mortality in that year and the subsequent year. In the Bangladesh data set, those children aged 6 at the time of the survey in the 1–6 year old group were born in the famine year, while those children in the 7–14 group who were aged 7 at the time of the survey were conceived during the famine year. Razzaque et al. (1990) report that both of these groups had lower survival rates prior to age 1 than other children born or conceived prior to or after the famine. Rates of mortality for older children in the famine year were not significantly higher compared to other years.

[5]Only one observation is taken (the first) for each surviving child, although children were surveyed as many as four times in the Bukidnon survey.

5. Results

Table 2 provides reduced-form least squares estimates of the determinants of weight for the (surviving) children. These specifications correspond to what would typically be used in evaluating the effects of the health infrastructure or health interventions on health. We have also estimated the same specifications and the complete model for other measures of health, including weight-for-height and skinfold thickness with similar results. The estimates in Table 2 indicate that in Bangladesh, although there are significant positive effects of the schooling of the household head and, for the older children, landholding size on children's weight, there are no significant effects of the water sources or of having a specific area for waste disposal. For the Philippines sample, only household wealth has a significant positive effect on the younger children's weight.[6] Among the older children in the Bukidnon sample, however, while neither wealth nor parental schooling affects children's weight significantly, there is a positive effect of access to a modern, water-sealed toilet.

The lack of significant effects of water sources or, for the most part, for the waste disposal mechanisms on children's weight is not inconsistent with prior evidence in the literature. Because, as discussed, household allocative behavior and mortality selectivity may tend to bias downward the health effects of better water and sanitation facilities, we cannot yet conclude whether or not such interventions are effective in improving children's body weight. To assess whether better water and sanitation facilities affect child survival, and as the first step in implementing the full simultaneous equations selectivity model, we estimated probit functions and also obtained semi-parametric estimates of the determinants of survival based on the sample of all births in the relevant years. Tables 3 and 4 present both sets of estimates for the Bangladesh and Bukidnon samples, respectively. As noted, the semi-parametric estimates require a normalization using a continuous variable and we have used household landholdings and wealth; i.e., the coefficients on these variables in their respective samples were set to unity. To compare the probit and semi-parametric estimates we also present the probit estimates similarly normalized. All coefficient standard errors are corrected for the presence of multiple children within households as described in the appendices. Standard errors are not corrected for village clustering effects. The probit results for all samples indicate that neither water sources nor sanitation facilities significantly affect child survival. However, greater landholdings and increased parental schooling levels in the Bangladesh sample significantly improve survival

[6] We use the schooling of the household head in the specification estimated from the Bangladesh samples because there is almost no variation in the schooling of adult women. The schooling of the mother is used in the specification estimated from the Bukidnon samples as that is the schooling variable commonly used. Replacement of the mother's schooling by the head's schooling does not affect the results.

Table 2
Reduced-form determinants of the weight of surviving children in Bangladesh and Bukidnon, Philippines, by age group (asymptotic standard error in parentheses corrected for nonindependence of errors within households)

Variable	Age group 1–6		Age group 7–14	
	Mean	Coefficient	Mean	Coefficient
Bangladesh[a]				
Land size ($\times 10^{-4}$)	0.181	0.0945 (0.392)	0.180	37.5 (7.94)**
Head's schooling	2.25	0.219 (0.0047)**	2.22	0.402 (0.179)**
Male	0.524	0.752 (0.167)**	0.534	0.299 (0.366)
Drinking water source (omitted=tubewell)				
Well	0.173	−0.287 (0.329)	0.144	1.38 (0.924)
Pond	0.102	−0.238 (0.579)	0.101	−0.195 (1.23)
River/canal	0.037	0.429 (0.783)	0.031	−1.62 (2.31)
Other	0.062	0.614 (0.492)	0.062	1.77 (0.956)*
Place for refuse disposal	0.196	−0.115 (0.270)	0.178	0.892 (0.655)
Bukidnon, Philippines[b]				
Wealth ($\times 10^{-4}$)	2.18	0.0277 (0.0129)**	2.69	0.0438 (0.0303)
Mother's schooling	6.32	0.0201 (0.0294)	5.97	0.0445 (0.0744)
Male	0.532	0.425 (0.113)**	0.506	1.10 (0.203)**
Drinking water source (omitted=piped)				
Artesian well	0.136	−0.169 (0.277)	0.130	0.0540 (0.870)
Dry well, open	0.340	−0.0210 (0.214)	0.325	−0.115 (0.748)
Spring, improved	0.229	−0.277 (0.251)	0.253	−0.214 (0.785)
Spring, unimproved	0.189	0.377 (2.44)	0.187	−0.0702 (0.746)
Rainwater	—	—	0.021	2.08 (1.43)
Disposal (omitted=none)				
Open pit	0.364	−0.199 (0.219)	0.378	0.0726 (0.461)
Antipolo	0.269	−0.0342 (0.249)	0.297	0.316 (0.488)
Water sealed	0.208	0.216 (0.262)	0.211	1.13 (0.534)**

[a] Specification also includes head's age, child's age and its square and village dummies (15).
[b] Specification also includes mother's age, child's age and its square.

* Significant at 0.10 level. **Significant at 0.05 level.

Table 3
Determinants of survival in Bangladesh, by age group and estimation procedure; all specifications also include the head's age and its square and village dummy variables (15) (asymptotic standard errors in parentheses corrected for nonindependence of errors within households)

	Children ages 1–6			Children ages 7–14		
Variable	ML probit	Normalized probit ML	Semi-p ML[a]	ML probit	Normalized probit ML	Semi-p ML[b]
Owned land ($\times 10^{-3}$)	0.588	—	—	0.365	—	—
	(0.294)**			(0.215)*		
Schooling of head	0.0887	0.151	0.165	0.124	0.340	1.17
	(0.0493)*	(0.0878)*	(0.0887)*	(0.0443)**	(0.144)**	(0.483)**
Male child	0.0125	0.0213	0.0318	−0.0612	−0.168	−0.0472
	(0.117)	(0.198)	(0.142)	(0.0986)	(0.290)	(0.308)
Drinking water source (left out=tubewell)						
Well	−0.0585	−0.0995	−0.0407	−0.355	−0.973	−1.19
	(0.203)	(0.593)	(0.494)	(0.189)*	(0.547)	(1.06)
Pond	−0.119	−0.0203	0.317	−0.289	−0.792	−1.30
	(0.349)	(1.16)	(0.780)	(0.225)	(0.624)	(1.74)
River/canal	0.283	0.481	0.983	−0.323	−0.885	−0.909
	(0.683)	(0.741)	(0.882)	(0.584)	(1.60)	(2.14)
Other	0.734	1.25	0.753	0.886	2.43	2.54
	(0.544)	(0.926)	(1.77)	(0.548)	(1.51)	(1.91)
Place for waste disposal	0.233	0.396	0.367	0.144	0.395	0.539
	(0.168)	(0.284)	(0.250)	(0.154)	(0.426)	(0.687)
H_0: Accept normal, χ^2_{23}	—	—	17.9	—	—	98.4
n, sample size	611	611	611	693	693	693

[a] Bandwidth factor=0.5.
[b] Bandwidth factor=3.0.
*Significant at 0.10 level. **Significant at 0.05 level.

probabilities. In the Philippines samples, increases in maternal schooling, but not wealth, improve survival. Tests of the normal parameterization versus the unknown alternative fail to reject normality in all but one of the four subsamples, that for the children aged 7–14 in Bangladesh. In that subsample, comparison of the rejected normalized probit coefficients with the comparable semi-parametric estimates suggests that the survival-enhancing effects of maternal schooling and

Table 4
Determinants of Survival in Bukidnon, Philippines, by age group and estimation procedure

Variable	Children ages 1–6			Children ages 7–14		
	ML probit	Normalized probit ML	Semi-p ML[a]	ML probit	Normalized probit ML	Semi-p ML[b]
Wealth ($\times 10^{-4}$)	0.145	—	—	0.0185	—	—
	(0.262)			(0.0142)		
Mother's schooling	0.0571	3.92	3.84	0.0463	2.50	2.92
	(0.0309)*	(24.7)	(2.57)	(0.0304)*	(10.7)	(3.36)
Drinking water source (left out=piped water)						
Artesian well (deep)	0.130	8.97	9.05	−0.558	−30.2	−28.9
	(0.259)	(29.9)	(16.2)	(0.292)*	(20.8)	(29.3)
Dug well, open	0.180	12.4	11.1	−0.179	−9.69	−10.5
	(0.231)	(28.8)	(10.8)	(0.283)	(20.0)	(19.4)
Spring, improved	0.136	9.35	7.79	0.0225	1.22	1.25
	(0.242)	(28.3)	(13.1)	(0.285)	(20.5)	(21.2)
Spring, unimproved	0.119	8.22	7.01	0.112	6.06	5.45
	(0.262)	(29.5)	(19.1)	(0.318)	(21.3)	(24.1)
Rainwater	—	—	—	−0.831	−44.9	−45.3
				(0.542)	(34.3)	(41.6)
Waste disposal (left out=none)						
Open pit	0.128	8.15	5.43	0.0361	1.95	2.32
	(0.191)	(28.6)	(13.2)	(0.235)	(17.3)	(12.5)
Antipolo	0.288	19.8	18.2	0.257	13.9	14.0
	(0.203)	(30.0)	(11.4)	(0.239)	(16.8)	(18.2)
Water sealed flush	0.239	16.4	15.6	0.349	18.9	19.1
	(0.271)	(34.2)	(15.0)	(0.267)	(17.6)	(25.9)
H_0: Accept normal, χ^2_{11}	—	—	0.364	—	—	0.122
n, sample size	837	837	837	789	789	789

[a] Bandwidth factor=7.50.

[b] Bandwidth factor=9.00.

*Significant at 0.10 level. **Significant at 0.05 level.

Table 5
TSLS weight production function estimates for Bangladesh, by age group and type of survival selection correction; all specifications also include the age of the child and its square and for children aged above 6 two indicators of activities stratified by their energy intensity, treated as endogenous (asymptotic standard errors in parentheses corrected for nonindependence of errors within households)

	Children ages 1–6			Children ages 7–14		
Variable	No correction	Bivariate normal	Semi-parametric	No correction	Bivariate normal	Semi-parametric
Calories[a]	0.474	0.482	0.549	0.365	0.326	0.281
	(0.176)**	(0.184)**	(0.184)**	(0.161)**	(0.178)*	(0.169)*
Male child	1.23	1.23	0.833	2.17	2.27	2.99
	(0.461)**	(0.462)**	(0.496)*	(2.20)	(2.78)	(2.60)
Drinking water source (left out=tubewell)						
Well	−0.492	−0.519	−0.797	−0.229	0.471	0.675
	(0.459)	(0.470)	(0.607)	(0.772)	(1.02)	(0.882)
Pond	−0.433	−0.442	−0.320	1.03	1.35	1.06
	(0.342)	(0.349)	(0.415)	(1.14)	(1.14)	(1.12)
River/canal	0.655	0.631	0.593	0.295	0.852	0.444
	(0.885)	(0.900)	(0.927)	(3.25)	(3.28)	(3.21)
Other	0.700	0.796	1.10	0.772	−0.119	−1.36
	(0.604)	(0.634)	(0.782)	(0.943)	(1.23)	(1.53)
Place for waste disposal	0.371	0.382	0.293	1.35	1.39	1.61
	(0.313)	(0.321)	(0.391)	(0.815)*	(0.883)*	(0.852)*
H_0: No selectivity	—	t=0.42	χ^2_{11}=8.38	—	t=1.17	χ^2_{13}=8.70
n, sample size	363	363	363	381	381	381

[a] Endogenous variable; instruments include food prices, landholdings, head's age and schooling.
* Significant at 0.10 level. ** Significant at 0.05 level.

of tubewells relative to ponds may be understated based on probit survival estimates. The normalized semi-parametric coefficient estimate for schooling is almost four times its normalized probit counterpart; the semi-parametric coefficient for pond water is 64 percent greater than the probit estimate in absolute value. Other coefficients are similar across the two procedures, suggesting that these specific differences in the normalized coefficients do not arise merely because the normalizing landholding variable coefficient is upward biased.

Tables 5 and 6 report the estimates of the health (weight) production functions for the Bangladesh and Bukidnon samples, respectively. For each subsample, there are three sets of estimates reported. The first is from the standard simul-

Table 6

TSLS weight production function estimates for Bukidnon, Philippines, by age group and type of survival selection correction; all specifications also include the mother's age (asymptotic standard errors in parentheses corrected for nonindependence of errors within households)

Variable	Children ages 1–6			Children ages 7–14		
	No correction	Bivariate normal	Semi-parametric	No correction	Bivariate normal	Semi-parametric
Calories[d]	0.107	0.102	0.108	0.190	0.136	0.143
	(0.0469)[c]**	(0.0574)*	(0.0786)	(0.124)	(0.124)	(0.161)
Male	0.322	0.326	0.280	1.08	1.10	1.15
	(0.123)**	(0.124)**	(0.134)**	(0.272)**	(0.266)**	(0.305)**
Drinking water source (left out=piped water)						
Artesian well (deep)	0.0957	0.0347	0.310	−0.313	0.506	0.980
	(0.285)	(0.326)	(0.417)	(0.948)	(1.04)	(1.04)
Dug well, open	0.0714	0.0277	0.225	−0.486	−0.190	0.453
	(0.234)	(0.265)	(0.339)	(0.771)	(0.855)	(0.760)
Spring, improved	−0.189	−0.244	−0.0810	−0.706	−0.702	0.077
	(0.262)	(0.297)	(0.359)	(0.851)	(0.958)	(0.820)
Spring, unimproved	0.443	0.433	0.668	−0.478	−0.600	0.310
	(0.257)**	(0.287)*	(0.361)*	(0.793)	(0.901)	(0.764)
Rainwater	—	—	—	0.846	2.35	4.59
				(1.96)	(2.09)	(2.33)**
Waste disposal (left out=none)						
Open pit	−0.0412	−0.147	−0.255	0.128	0.0104	0.00835
	(0.225)	(0.264)	(0.267)	(0.477)	(0.591)	(0.446)
Antipolo	0.185	0.0831	−0.157	0.554	0.138	0.176
	(0.278)	(0.347)	(0.336)	(0.509)	(0.632)	(0.517)
Water sealed flush	0.395	0.181	−0.010	1.48	0.787	0.892
	(0.258)	(0.342)	(0.325)	(0.590)**	(0.750)	(0.719)
H_0: No selectivity	—	t=1.07	χ^2_{11}=12.8	—	t=1.69	χ^2_{12} = 8.10
n, sample size	773	773	773	716	716	716

[a] Bandwidth factor=7.50.

[b] Bandwidth factor=9.00.

*Significant at 0.10 level. **Significant at 0.05 level.

taneous equations model and thus ignores survival selectivity. The second set of estimates is obtained from the simultaneous equations selection model in which errors are assumed to be jointly normally distributed (Lee et al., 1980). The third set of estimates is from the semi-parametric selectivity-corrected simultaneous equations model outlined in the previous sections. In all models the coefficient standard errors are corrected for the nonindependence of the errors because of the household-based sample clustering of children. We report the selectivity model estimates based on the parametric assumption of bivariate normality for comparison with the estimates that correct for selectivity without imposing any parametric distributional assumption because the bivariate normal selectivity model is often used when sample selection is taken into account (although it has never been applied to mortality). It is important to note, however, that the acceptance of the normality assumption for the latent survival function, as was the case for some of the subsamples, does not necessarily imply that the assumption of joint normality, inclusive of the health equation, that is the basis of the bivariate normal selection model, is correct. The semi-parametric estimates of the full model impose no parametric distributional assumptions at all, as is true for the standard (two-stage least squares) simultaneous equations model without selectivity correction.

The sets of estimates do not appear sensitive to selectivity correction or to the parametric assumption of normality. The test statistics for selectivity based on the semi-parametric model indeed indicate that there is no death selection, that the probability of survival is random with respect to health, at least as measured by the standard anthropometric measures. The bivariate normal model does however yield a test statistic rejecting the nonselectivity of mortality for the older-children subsample of the Bukidnon data. We have seen that the assumption of the normality of the survival function was not rejected for this subsample. If the joint normality assumption is correct, then the bivariate-normal model yields more efficient estimates than the semi-parametric model; however, its estimates are not very different from those of the uncorrected model. The test results based on the semi-parametric model which is robust to assumptions about error distributions thus indicate that the standard simultaneous equations estimates of the effects of calorie intakes, the water sources and the sanitation facilities are preferred.[7] Tests of whether the calorie intake variable is correlated with the error in the health production function, containing the health endowment, also indicated rejection of the hypothesis of orthogonality for the two Bangladesh subsamples but nonrejection for both of the Bukidnon subsamples.[8] Thus, calorie allocation

[7]The test statistics yield identical conclusions for weight-for-height and skinfold thickness.

[8]The relevant test statistics for the Bangladesh subsamples of younger and older children are $F(2,349) = 10.0$ and $F(4,363) = 3.85$, respectively. Both specifications include in addition to calories a variable indicating how many family members consumed meals outside the household that is also treated as an endogenous variable. In addition, for the older children, there are two endogenous variables measuring type of work activity. The relevant t-statistics for the test of endogeneity of the calorie variable in the Philippines sample are 1.29 and 1.13 for the older and younger children, respectively.

appears to be sensitive to endowments in the former environment but not the latter. This may be due to the substantially higher average level of consumption (and income) in the Bukidnon area compared to Bangladesh (Table 1). Indeed, the estimated effects of variation in calorie intake on weight are substantially higher for the Bangladesh subsamples than for the Bukidnon subsamples, suggesting that there may not only be nonlinearities in the sensitivity of the allocation of calories among children differentiated by health endowments but also nonlinearities in the health technology with respect to calorie effects.

The preferred estimates that do not correct for mortality selectivity also indicate that in all but one of the subsamples there are no effects of different water sources for the weight of children. The exception is that water from unimproved springs appears to significantly positively affect the weight of the younger children in the Bukidnon subsample compared to all other water sources. Comparison with the reduced-form estimate of Table 2 from that subsample suggests that that specification understates the 'effectiveness' of this water source by 18 percent presumably due to household calorie allocations differing by water source. The uncorrected simultaneous equations estimates also indicate that the method of waste disposal matters among the older children, although not among the younger children. In Bangladesh older-children residing in households having a specific area for waste disposal appear to be over a kilogram heavier than similarly-aged children in households having no specified area, controlling for calorie intake. The estimates from the Bukidnon subsample indicate that children aged 7–14 are almost 1.5 kilograms heavier in households having a modern toilet facility compared to children in households with other waste disposal methods and receiving the same calorie allocation. Again, the reduced-form estimates, which do not take into account that household calorie allocations are not orthogonal to waste disposal facilities, underestimate the effectiveness of altering waste disposal methods, by 56 percent for the Bangladesh subsample and by 31 percent in the Bukidnon subsample.

6. Conclusion

In this paper, we have set out a framework for estimating the effects of health interventions that takes into account the three principal mechanisms by which improvements in the health infrastructure augment children's health: by affecting the magnitude of parental resources allocated to the health of children, by influencing who among children born survive, and by directly affecting the health of survivors given parental resource allocations. We devoted particular attention to the nonrandom allocation of household resources to children and to the selectivity effects of alterations in health interventions via their effects on child survival. Estimates were obtained using data sets describing child health and survival, drinking water sources, sanitation facilities, and child-specific nutritional intake from two countries – Bangladesh and the Philippines. Semi-parametric estimators were used in the context of a structural equations system to estimate the effects of the health infrastructural variables on child survival and the effects of increased calorie consumption and of improvements in the water and sanitation on measures of children's nutritional status net and gross of the effects of these interventions on child survival. The results indicated that the reduced-form estimates obtained in most studies evaluating health interventions understate the effectiveness of improving sanitation facilities in augmenting health as measured by anthropometric indicators, particularly among older children. However, this appeared to be due to the reduced allocation of household resources to children's health in households with better facilities and not to survival selectivity. The test statistics did not reject the hypothesis of random mortality in all but one of the subsamples (older children in Bangladesh); deaths among children in these low-income countries thus do not appear for the most part to be more likely among children with lower health endowments. Moreover, neither variation in water sources nor improvements in sanitation facilities appeared to significantly affect child survival, although wealth and parental schooling levels were significantly and positively associated with higher survival.

Appendix A: Asymptotic variances

The asymptotic distribution of the SML estimate of (3.2′) can be derived as in Klein and Spady (1993) (see also Lee, 1995). It can be shown that, by a Taylor expansion,

$$\sqrt{n}(\hat{\eta}-\eta) = D_n^{-1}\frac{1}{\sqrt{n}}\sum_{i=1}^{n} I_T(x_{1i}\hat{\eta})\left\{\sum_{l=1}^{2} d_{li}\,\frac{\partial \ln \mathrm{E}_n(d_l|x_{1i}\bar{\eta})}{\partial \eta}\right\} + o_P(1),$$

where $o_P(1)$ refers to terms which converge to zero in probability. The estimator is $\sqrt{n}$-consistent and is asymptotically normally distributed. If sample observations are i.i.d., $\sqrt{n}(\hat{\eta}-\eta) \xrightarrow{D} \mathrm{N}(0,\Omega_c)$ where the variance matrix Ω_c can be consistently estimated by D_n^{-1} evaluated at $\hat{\eta}$. For our empirical study, sample observations of children within each family unit are likely correlated. The $\hat{\eta}$ is consistent but its asymptotic covariance matrix needs to take into account the interfamily correlation. Let m denote the number of families, and let i_f denote the number of children in the family f. If the number of children of each family is exogenous, the members of a family can be grouped together such that

$$\begin{aligned}&\frac{1}{\sqrt{n}}\sum_{i=1}^{n} I_T(x_{1i}\bar{\eta})\sum_{l=1}^{2} d_{li}\,\frac{\partial \ln \mathrm{E}_n(d_l|x_{1i}\bar{\eta})}{\partial \eta}\\ &= \frac{1}{\sqrt{n}}\sum_{f=1}^{m}\left\{\sum_{j=1}^{i_f} I_T(x_{1j}^f\bar{\eta})\sum_{l=1}^{2} d_{lj}^f\,\frac{\partial \ln \mathrm{E}_n(d_l|x_{1j}^f\bar{\eta})}{\partial \eta}\right\}.\end{aligned}$$

In this case,

$$\sqrt{n}(\hat{\eta} - \eta) = D_n^{-1} \frac{1}{\sqrt{n}} \sum_{f=1}^{m} \left\{ \sum_{j=1}^{i_f} I_T(x_{1j}^f, \bar{\eta}) \sum_{l=1}^{2} d_{lj}^f \frac{\partial \ln \mathrm{E}_n(d_l | x_{1j}^f, \bar{\eta})}{\partial \eta} \right\} + \mathrm{o}_P(1). \tag{A.1}$$

Under the assumption that all the families in the sample are unrelated, the terms within the bracket in (A.1) are independent for different f and the summation is effectively summing over m independent samples. The limiting variance matrix of $\sqrt{n}(\hat{\eta} - \eta)$ can be consistently estimated by $\Omega_{c,n} = D_n^{-1} H_n D_n^{-1}$, where

$$H_n = \frac{1}{n} \sum_{f=1}^{m} \left\{ \sum_{j=1}^{i_f} I_T(x_{1j}^f, \bar{\eta}) \sum_{l=1}^{2} d_{lj}^f \frac{\partial \ln \mathrm{E}_n(d_l | x_{1j}^f, \bar{\eta})}{\partial \eta} \right\}$$

$$\times \left\{ \sum_{j=1}^{i_f} I_T(x_{1j}^f, \bar{\eta}) \sum_{l=1}^{2} d_{lj}^f \frac{\partial \ln \mathrm{E}_n(d_l | x_{1j}^f, \bar{\eta})}{\partial \eta'} \right\},$$

evaluated at $\hat{\eta}$. For the estimates in our empirical study, the corrected asymptotic variances are reported.

The $\hat{\beta}$ in (3.5) is a two-stage estimator. We first discuss the asymptotic distribution of $\hat{\beta}$ under the assumption that sample observations are i.i.d., and then discuss necessary modifications to take into account the family correlation structure. Eq. (3.5) implies that

$$\hat{\beta} - \beta = [\hat{W}'\hat{X}_3(\hat{X}_3'\hat{X}_3)^{-1}\hat{X}_3'\hat{W}]^{-1}\hat{W}'\hat{X}_3'(\hat{X}_3'\hat{X}_3)^{-1}\hat{X}_3'\hat{V}_n,$$

where $\hat{V}_n = (\hat{v}_{n1}, \cdots, \hat{v}_{nn_1})'$ with $\hat{v}_{ni} = v_i - \mathrm{E}_n(v | x_{1i}\hat{\bar{\eta}})$. As shown in Lee (1994),

$$\frac{1}{\sqrt{n}} \hat{X}_3' \hat{V}_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} d_{1i} I_T(x_{1i}\bar{\eta})(x_{3i} - \mathrm{E}_n(x_3 | x_{1i}\bar{\eta}))'(v_i - \mathrm{E}_n(v | x_{1i}\bar{\eta}))$$

$$- \frac{1}{n} \sum_{i=1}^{n} d_{1i} I_T(x_{1i}\bar{\eta})(x_{3i} - \mathrm{E}_n(x_3 | x_{1i}\bar{\eta}))' \frac{\partial \mathrm{E}_n(v | x_{1i}\bar{\eta})}{\partial \eta'} \cdot \sqrt{n}(\hat{\eta} - \eta)$$

$$+ \mathrm{o}_P(1) \to \mathrm{N}(0, \Delta),$$

where $\Delta = \mathrm{E}[I_T(x_1\bar{\eta})(x_3 - \mathrm{E}(x_3 | x_1\bar{\eta}))'\omega(x_1\bar{\eta})(x_3 - \mathrm{E}(x_3 | x_1\bar{\eta}))] + C\Omega_c C'$, $\omega(x_1\bar{\eta})$ is the variance matrix of $d_1 v$, $C = \mathrm{E}(d_1 I_T(x_1\bar{\eta})(x_3 - \mathrm{E}(x_3 | x_1\bar{\eta}))'(\partial \mathrm{E}(v | x_1\bar{\eta}) / \partial \eta'))$, and $\mathrm{E}(\cdot | x_1\bar{\eta})$ stands for $\mathrm{E}(\cdot | x_1\bar{\eta}, d_1 = 1)$. Let $A = \mathrm{plim}(\hat{X}_3'\hat{X}_3)^{-1}\hat{X}_3'\hat{W}$ and $B = \mathrm{plim}(1/n)(\hat{W}'\hat{X}_3)(\hat{X}_3'\hat{X}_3)^{-1}\hat{X}_3'\hat{W}$. The $\hat{\beta}$ is $\sqrt{n}$-consistent and $\sqrt{n}(\hat{\beta} - \beta) \xrightarrow{D} \mathrm{N}(0, \Omega_h)$, where $\Omega_h = B^{-1}A'\Delta A B^{-1}$.

With correlation across members of a family, in addition to the correction of the variance matrix of the first-stage estimator $\hat{\eta}$, a correction is needed for the first term in Δ. With m families in the sample,

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n} d_{1i}I_T(x_{1i}\bar{\eta})(x_{3i}-\mathrm{E}_n(x_3|x_{1i}\bar{\eta}))'(v_i-\mathrm{E}_n(v|x_{1i}\bar{\eta}))$$
$$=\frac{1}{\sqrt{n}}\sum_{f=1}^{m}\sum_{j=1}^{i_f} d_{1j}^f I_T(x_{1j}^f\bar{\eta})(x_{3j}^f-\mathrm{E}_n(x_3|x_{1j}^f\bar{\eta}))'(v_j^f-\mathrm{E}_n(v|x_{1j}^f\bar{\eta})). \tag{A.2}$$

As samples are independent across families, the limiting variance matrix of (A.2) can be estimated by $Q_{c,n}$, where

$$Q_{c,n}=\frac{1}{n}\sum_{f=1}^{m}\left\{\sum_{j=1}^{i_f} d_{1j}^f I_T(x_{1j}^f\bar{\eta})(x_{3j}^f-\mathrm{E}_n(x_3|x_{1j}^f\bar{\eta}))'(v_j^f-\mathrm{E}_n(v|x_{1j}^f\bar{\eta}))\right\}$$
$$\times\left\{\sum_{j=1}^{i_f} d_{1j}^f I_T(x_{1j}^f\bar{\eta})(v_j^f-\mathrm{E}_n(v|x_{1j}^f\bar{\eta}))(x_{3j}^f-\mathrm{E}_n(x_3|x_{1j}^f\bar{\eta}))\right\}.$$

Denote

$$A_n=(\hat{X}_3'\hat{X}_3)^{-1}\hat{X}_3'\hat{W},$$
$$B_n=\frac{1}{n}(\hat{W}'\hat{X}_3)(\hat{X}_3'\hat{X}_3)^{-1}\hat{X}_3'\hat{W},$$
$$C_n=\frac{1}{n}\sum_{i=1}^{n} d_{1i}I_T(x_{1i}\bar{\eta})(x_{3i}-\mathrm{E}_n(x_3|x_{1i}\bar{\eta}))'\,\frac{\partial\mathrm{E}_n(v|x_{1i}\bar{\eta})}{\partial\eta'}.$$

It follows that

$$\sqrt{n}(\hat{\beta}-\beta)=B_n^{-1}A_n'\left\{\frac{1}{\sqrt{n}}\sum_{f=1}^{m}\sum_{j=1}^{i_f} d_{1j}^f I_T(x_{1j}^f\bar{\eta})(x_{3j}^f-\mathrm{E}_n(x_3|x_{1j}^f\bar{\eta}))'\right.$$
$$\left.\times\,(v_j^f-\mathrm{E}_n(v|x_{1j}^f\bar{\eta}))-C_n\sqrt{n}(\hat{\eta}-\eta)\right\}+o_P(1). \tag{A.3}$$

The limiting variance matrix of $\sqrt{n}(\hat{\beta}-\beta)$ can be estimated by $B_n^{-1}A_n'(Q_{c,n}+C_n\Omega_{c,n}C_n')A_nB_n^{-1}$.

Appendix B: Test statistics

Let $\bar{x}_1=(1,x_1)$ be the extended vector of x_1 which incorporates a constant term. With the normalization that ε has zero mean and a unit variance, let α be the coefficient vector of $\bar{x}_1$ in (2.2). Let Φ denote the standard normal distribution. If ε were normally distributed, the probit log-likelihood function for (2.2) would be

$$\ln L_p=\sum_{i=1}^{n}\{d_{1i}\ln\Phi(\bar{x}_{1i}\alpha)+d_{2i}\ln[1-\Phi(\bar{x}_{1i}\alpha)]\}.$$

Let $\hat{\alpha}_p$ be the probit MLE. Taking into account the family correlation structure,

$$\sqrt{n}(\hat{\alpha}_p - \alpha) = D_{p,n}^{-1}\frac{1}{\sqrt{n}}\sum_{f=1}^{m}\sum_{j=1}^{i_f}\frac{\phi^2(\bar{x}_{1j}^f\alpha)}{\Phi(\bar{x}_{1j}^f\alpha)(1-\Phi(\bar{x}_{1j}^f\alpha))}\bar{x}_{1j}^{\prime f} \times(d_{1j}^f - \Phi(\bar{x}_{1j}^f\alpha)) + o_P(1), \tag{B.1}$$

where $\phi(\cdot)$ is the standard normal density function, and

$$D_{p,n} = \frac{1}{n}\sum_{i=1}^{n}\frac{\phi^2(\bar{x}_{1i}\alpha)}{\Phi(\bar{x}_{1i}\alpha)(1-\Phi(\bar{x}_{1i}\alpha))}\bar{x}_{1i}'\bar{x}_{1i}.$$

Let $\alpha = (\alpha_0, \alpha_1, \alpha_2)$, where α_0 is the intercept coefficient and α_1 is the coefficient of the continuous regressor in x_1 which coefficient is normalized to be unity in the semi-parametric estimation. Thus $\eta = \alpha_2/\alpha_1$, and the probit MLE of η is $\hat{\eta}_p = \hat{\alpha}_{2,p}/\hat{\alpha}_{1,p}$. By the Delta method of C.R. Rao,

$$\sqrt{n}(\hat{\eta}_p - \eta) = PD_{p,n}^{-1}\sum_{f=1}^{m}\sum_{j=1}^{i_f}\frac{\phi^2(\bar{x}_{1j}^f\alpha)}{\Phi(\bar{x}_{1j}^f\alpha)(1-\Phi(\bar{x}_{1j}^f\alpha))}\bar{x}_{1j}^{\prime f}(d_{1j}^f - \Phi(\bar{x}_{1j}^f\alpha)) +o_P(1), \tag{B.2}$$

where $P = (\partial\eta/\partial\alpha_0, \partial\eta/\partial\alpha_1, \partial\eta/\partial\alpha_2') = (0, -\alpha_1^{-2}\alpha_2, \alpha_1^{-1}I)$. Under the probit hypothesis, (A.1) and (B.2) imply that

$$\sqrt{n}(\hat{\eta} - \hat{\eta}_p) = [D_n^{-1}, -PD_{p,n}^{-1}]\frac{1}{\sqrt{n}}\sum_{f=1}^{m}\begin{pmatrix}\displaystyle\sum_{j=1}^{i_f} I_T(x_{1j}^f\bar{\eta})\sum_{l=1}^{2} d_{l,j}^f\frac{\partial \ln \mathrm{E}_n(d_l|x_{1i}^f\bar{\eta})}{\partial\eta} \\ \displaystyle\sum_{j=1}^{i_f}\frac{\phi^2(\bar{x}_{1j}^f\alpha)}{\Phi(\bar{x}_{1j}^f\alpha)(1-\Phi(\bar{x}_{1j}^f\alpha))}\bar{x}_{1j}^{\prime f}(d_{1j}^f - \Phi(\bar{x}_{1j}^f\alpha))\end{pmatrix} +o_P(1),$$

which is asymptotically normal with zero mean and the variance matrix of its limiting distribution can be consistently estimated by

$$\Omega_{p,n} = [D_n^{-1} - PD_{p,n}^{-1}]\frac{1}{n}\sum_{f=1}^{m}\begin{pmatrix}\displaystyle\sum_{j=1}^{i_f} I_X(x_j^f)\sum_{l=1}^{2} d_{l,j}^f\frac{\partial \ln \mathrm{E}_n(d_l|x_{1i}^f\bar{\eta})}{\partial\eta} \\ \displaystyle\sum_{j=1}^{i_f}\frac{\phi^2(\bar{x}_{1j}^f\alpha)}{\Phi(\bar{x}_{1j}^f\alpha)(1-\Phi(\bar{x}_{1j}^f\alpha))}\bar{x}_{1j}^{\prime f}(d_{1j}^f - \Phi(\bar{x}_{1j}^f\alpha))\end{pmatrix}$$

$$\times\begin{pmatrix}\displaystyle\sum_{j=1}^{i_f} I_T(x_{1j}^f\bar{\eta})\sum_{l=1}^{2} d_{l,j}^f\frac{\partial \ln \mathrm{E}_n(d_l|x_{1i}^f\bar{\eta})}{\partial\eta} \\ \displaystyle\sum_{j=1}^{i_f}\frac{\phi^2(\bar{x}_{1j}^f\alpha)}{\Phi(\bar{x}_{1j}^f\alpha)(1-\Phi(\bar{x}_{1j}^f\alpha))}\bar{x}_{1j}^{\prime f}(d_{1j}^f - \Phi(\bar{x}_{1j}^f\alpha))\end{pmatrix}'[D_n^{-1}, -PD_{p,n}^{-1}]',$$

evaluated at $\hat{\alpha}$ and $\hat{\eta}$ (or $\hat{\eta}_p$). The statistic $n(\hat{\eta}-\hat{\eta}_p)\Omega_{p,n}^{+}(\hat{\eta}-\hat{\eta}_p)$ is asymptotically chi-square with the degree of freedom equal to the rank of the limiting matrix of $\Omega_{p,n}$. This statistic is a generalized Hausman-type test statistic.

It is also possible to derive a test of simultaneity by a generalized Hausman-type statistic. If z were exogenous, the β in (2.3) could be semi-parametrically estimated by $\hat{\beta}_x = (\hat{W}'\hat{W})^{-1}\hat{W}'\hat{H}_2^*$. Taking into account the family correlation,

$$\sqrt{n}(\hat{\beta}_x - \beta) = \left(\frac{1}{n}\hat{W}'\hat{W}\right)^{-1}\frac{1}{\sqrt{n}}\hat{W}'\hat{V}_n,$$

where

$$\frac{1}{\sqrt{n}}\hat{W}'\hat{V}_n = \frac{1}{\sqrt{n}}\sum_{f=1}^{m}\sum_{j=1}^{i_f} d_{1j}^f I_T(x_{1j}^f,\bar{\eta})(w_j^f - \mathrm{E}_n(w|x_{1j}^f,\bar{\eta}))' \times (v_j^f - \mathrm{E}_n(v|x_{1j}^f,\bar{\eta})) - C_{n,w}\sqrt{n}(\hat{\eta} - \eta) + o_P(1),$$

with

$$C_{n,w} = \frac{1}{n}\sum_{i=1}^{n} d_{1i}I_T(x_{1i},\bar{\eta})(w_i - \mathrm{E}_n(w|x_{1i},\bar{\eta}))'\frac{\partial \mathrm{E}_n(v|x_{1i},\bar{\eta})}{\partial \eta'}.$$

Let

$$B_{n,w} = \frac{1}{n}\hat{W}'\hat{W}.$$

With (A.3), it follows that

$$\sqrt{n}(\hat{\beta} - \hat{\beta}_x) = [B_n^{-1}A_n', -B_{n,w}^{-1}]\frac{1}{\sqrt{n}}\sum_{f=1}^{m}\sum_{j=1}^{i_f} d_{1j}^f I_T(x_{1j}^f,\bar{\eta})\begin{pmatrix}(x_{3j}^f - \mathrm{E}_n(x_3|x_{1j}^f,\bar{\eta}))' \\ (w_j^f - \mathrm{E}_n(w|x_{1j}^f,\bar{\eta}))'\end{pmatrix} \times (v_j^f - \mathrm{E}_n(v|x_{1j}^f,\bar{\eta})) - (B_n^{-1}A_n'C_n - B_{n,w}^{-1}C_{n,w})\sqrt{n}(\hat{\eta} - \eta) + o_P(1),$$

which is asymptotically normal with zero mean under the exogeneity hypothesis for z. The variance matrix of the limiting distribution of $\sqrt{n}(\hat{\beta} - \hat{\beta}_x)$ can be estimated by

$$\begin{aligned}\Omega_{x,n} = {} & [B_n^{-1}A_n', -B_{n,w}^{-1}] \\ & \times \frac{1}{n}\sum_{f=1}^{m}\left[\sum_{j=1}^{i_f} d_{1j}^f I_T(x_{1j}^f,\bar{\eta})(v_j^f - \mathrm{E}_n(v|x_{1j}^f,\bar{\eta}))\begin{pmatrix}(x_{3j}^f - \mathrm{E}_n(x_3|x_{1j}^f,\bar{\eta}))' \\ (w_j^f - \mathrm{E}_n(w|x_{1j}^f,\bar{\eta}))'\end{pmatrix}\right] \\ & \times \left[\sum_{j=1}^{i_f} d_{1j}^f I_T(x_{1j}^f,\bar{\eta})(v_j^f - \mathrm{E}_n(v|x_{1j}^f,\bar{\eta}))\begin{pmatrix}(x_{3j}^f - \mathrm{E}_n(x_3|x_{1j}^f,\bar{\eta}))' \\ (w_j^f - \mathrm{E}_n(w|x_{1j}^f,\bar{\eta}))'\end{pmatrix}\right]' \\ & \times [B_n^{-1}A_n', -B_{n,w}^{-1}]' + (B_n^{-1}A_n'C_n - B_{n,w}^{-1}C_{n,w})\Omega_{c,n}(B_n^{-1}A_n'C_n - B_{n,w}^{-1}C_{n,w})',\end{aligned}$$

evaluated at $\hat{\eta}$ and $\hat{\beta}$. A test statistic for simultaneity is $n(\hat{\beta}-\hat{\beta}_x)'\Omega_{x,n}^{+}(\hat{\beta}-\hat{\beta}_x)$, where $\Omega_{x,n}^{+}$ is the generalized inverse of $\Omega_{x,n}$. Under the null hypothesis, this statistic is asymptotically chi-square distributed with the degree of freedom equal to the rank of the limiting matrix of $\Omega_{x,n}$.

Under the hypothesis that ε and v are uncorrelated, i.e., no selection bias in observed h_2^*, (2.3) can be estimated by the classical Theil–Basmann two-stage least squares (TSL) procedure. For the classical simultaneous equations model, v is assumed to have zero mean and the intercept term is identified. To incorporate an intercept term for estimation, let $\bar{w}=(1,z,x_2)$, $\bar{x}_3=(1,x_3)$ and $\bar{\beta}=(\beta_0,\beta_1',\beta_2')'$. The classical TSL estimator is

$$\hat{\bar{\beta}}_s=\left\{\sum_{i=1}^{n_1}\bar{w}_i'\bar{x}_{3i}\left(\sum_{i=1}^{n_1}\bar{x}_{3i}'\bar{x}_{3i}\right)^{-1}\sum_{i=1}^{n_1}\bar{x}_{3i}'\bar{w}_i\right\}^{-1}$$
$$\times\sum_{i=1}^{n_1}\bar{w}_i'\bar{x}_{3i}\left(\sum_{i=1}^{n_1}\bar{x}_{3i}'\bar{x}_{3i}\right)^{-1}\sum_{i=1}^{n_1}\bar{x}_{3i}'h_{2i}^*. \tag{B.3}$$

Let $J=[0,I]$ be the selection matrix such that $\beta=J\bar{\beta}$. The TSL estimate of β is $\hat{\beta}_s=J\hat{\bar{\beta}}_s$. Let

$$B_{n,s}=\frac{1}{n}\sum_{i=1}^{n_1}\bar{w}_i'\bar{x}_{3i}\left(\sum_{i=1}^{n_1}\bar{x}_{3i}'\bar{x}_{3i}\right)^{-1}\sum_{i=1}^{n_1}\bar{x}_{3i}'\bar{w}_i,$$

$$A_{n,s}=\left(\sum_{i=1}^{n_1}\bar{x}_{3i}'\bar{x}_{3i}\right)^{-1}\sum_{i=1}^{n_1}\bar{x}_{3i}'\bar{w}_i.$$

It follows from (A.3) and (B.3) that

$$\begin{aligned}&\sqrt{n}(\hat{\beta}-\hat{\beta}_s)\\&=[B_n^{-1}A_n',-JB_{n,s}^{-1}A_{n,s}']\\&\quad\times\frac{1}{\sqrt{n}}\sum_{f=1}^{m}\sum_{j=1}^{i_f}\begin{pmatrix}d_{1j}^fI_T(x_{1j}^f,\hat{\eta})(x_{3j}^f-\mathrm{E}_n(x_3|x_{1j}^f,\hat{\eta}))'(v_j^f-\mathrm{E}_n(v|x_{1j}^f,\hat{\eta}))\\ \bar{x}_{3j}^{f'}(h_{2j}^{*f}-\bar{w}_j^f\bar{\beta})\end{pmatrix}\\&\quad-B_n^{-1}A_n'C_n\sqrt{n}(\hat{\eta}-\eta),\end{aligned}$$

which is asymptotically normal with zero mean under the hypothesis of no selectivity. The variance of its limiting distribution can be consistently estimated by

$$\begin{aligned}\Omega_{s,n}&=[B_n^{-1}A_n',-JB_{n,s}^{-1}A_{n,s}']\\&\quad\times\frac{1}{n}\sum_{f=1}^{m}\left[\sum_{j=1}^{i_f}\begin{pmatrix}d_{1j}^fI_T(x_{1j}^f,\hat{\eta})(x_{3j}^f-\mathrm{E}_n(x_3|x_{1j}^f,\hat{\eta}))'(v_j^f-\mathrm{E}_n(v|x_{1j}^f,\hat{\eta}))\\ \bar{x}_{3j}^{f'}(h_{2j}^{*f}-\bar{w}_j^f\bar{\beta})\end{pmatrix}\right]\end{aligned}$$

$$\times \left[\sum_{j=1}^{i_f} \begin{pmatrix} d_{1j}^f I_T(x_{1j}^f \bar{\eta})(x_{3j}^f - \mathrm{E}_n(x_3|x_{1j}^f \bar{\eta}))'(v_j^f - \mathrm{E}_n(v|x_{1j}^f \bar{\eta})) \\ \bar{x}_{3j}^{f'}(h_{2j}^{*f} - \bar{w}_j^f \bar{\beta}) \end{pmatrix} \right]'$$
$$\times [B_n^{-1} A_n', -J B_{n,s}^{-1} A_{n,s}']' + B_n^{-1} A_n' C_n \Omega_{c,n} C_n' A_n B_n^{-1},$$

evaluated at $\hat{\eta}$ and $\hat{\beta}$. A test statistic of no selectivity is $n(\hat{\beta} - \hat{\beta}_s)' \Omega_{s,n}^{+} (\hat{\beta} - \hat{\beta}_s)$ which is asymptotically chi-square distributed with the degree of freedom equal to the rank of the limiting matrix of $\Omega_{s,n}$.

References

Ahmad, K. and N. Hassani, 1986, Nutrition survey of rural Bangladesh 1981–82 (Institute of Nutrition and Food Science, University of Dhaka, Bangladesh).

Behrman, Jere R. and Anil Deolalikar, 1988a, How do food prices and income affect individual nutritions and health status: A latent variable fixed effects analysis, Manuscript (University of Pennsylvania, Philadelphia, PA).

Behrman, Jere R. and Anil Deolalikar, 1988b, Health and nuitrition, in: H. Chenery and T.N. Srinivasan, eds., Handbook of development economics (Elsevier Science, Amsterdam).

Bierens, H.J., 1985, Kernel estimators of regression functions, in: T.W. Bewley, ed., Advances in econometrics, Fifth world congress, Vol. 1 (Cambridge University Press, New York, NY) 99–144.

Bouis, H.E. and L.J. Haddad, 1990, Agricultural commercialization, nutrition, and the rural poor: A study of Philippine farm households (Lynne Rienner Publishers, Boulder, CO).

DaVanzo, Julie, 1988, Infant mortality and socioeconomic development: Evidence from Malaysia household data, Demography 25, 581–596.

Heckman, J., 1979, Sample selection bias as a specification error, Econometrica 47, 153–162.

Ichimura, H., 1993, Semiparametric least squares (SLS) and weighted SLS estimation of single-index models, Journal of Econometrics 58, 71–120.

Ichimura, H. and L.F. Lee, 1991, Semiparametric estimation of multiple index models: Single equation estimation, in: W.A. Barnett, J. Powell, and G. Tauchen, eds., Nonparametric and semiparametric methods in econometrics and statistics (Cambridge University Press, New York, NY).

International Center for Diarrhoeal Disease Research, Bangladesh, 1984, Vital events and migration-tables 1982, ICDDR.B scientific report no. 62.

Klein, R.W. and R.H. Spady, 1993, An efficient semiparametric estimator for binary response models, Econometrica 61, 387–421.

Lee, L.F., G.S. Maddala, and R.P. Trost, 1980, Asymptotic covariance matrices of two-stage probit and two-stage tobit methods for simultaneous equations models with selectivity, Econometrica 48, 491–503.

Lee, L.F., 1990, Efficient semiparametric scoring estimation of sample selection models, Discussion paper no. 255 (Center for Economic Research, University of Minnesota, Minneapolis, MN).

Lee, L.F., 1994, Semiparametric instrumental variable estimation of simultaneous equation sample selection models, Journal of Econometrics 63, 341–388.

Lee, L.F., 1995, Semiparametric maximum likelihood estimation of polychotomous and sequential choice models, Journal of Econometrics 65, 381–428.

Mroz, T.A. 1987, The sensitivity of an empirical model of married women's hours of work to economic and statistical assumptions, Econometrica 55, 765–799.

Olsen, R.J. and K.I. Wolpin, 1983, The impact of exogenous infant mortality on fertility: A waiting-time regression with dynamic regressors, Econometrica 51, 731–749.

Pitt, M.M. and M.R. Rosenzweig, 1990, Estimating the intrafamily incidence of health: Child illness and gender-inequality in Indonesian households, International Economic Review.

Pitt, M.M., M.R. Rosenzweig, and M.N. Hassan, 1990, Productivity, health and inequality in the intrahousehold distribution of food in low-income countries, American Economic Review 80, 1139–1156.

Powell, J.L., 1987, Semiparametric estimation of bivariate latent variable models, Discussion paper no. 8704 (Social Systems Research Institute, University of Wisconsin, WI).

Razzaque, A., N. Alam, L. Wai, and A. Foster, 1990, Sustained effects of the 1974–5 famine on infant and child mortality in a rural area of Bangladesh, Population Studies 44, 145–154.

Robinson, P.M., 1988, Root-n-consistent semiparametric regression, Econometrica 56, 931–954.

Rosenzweig, M.R. and T.P. Schultz, 1982a, Market opportunities, genetic endowments, and intrafamily resource distribution: Child survival in rural India, American Economic Review 72, 803–815.

Rosenzweig, M.R. and T.P. Schultz, 1982b, Estimating a household production function: Heterogeneity, the demand for health inputs, and their effects on birthweight, Journal of Political Economy 91, 723–746.

Ruud, P.A., 1983, Sufficient conditions for the consistency of maximum likelihood estimation despite misspecification of distribution in multinomial discrete choice models, Econometrica 51, 225–228.

Strauss, J. and K. Mehra, 1987, Cote d'Ivoire: Estimates for child anthropometry from two surveys, 1985 and 1986, LSMS working paper no. 51 (World Bank, Washington, DC).

Thomas, D., J. Strauss, and M.H. Henriques 1988, Child survival, height for age and household characteristics in Brazil (Yale University, New Haven, CT).

Vaupel, J.W., 1988, Inherited frailty and longevity, Demography 25, 277–298.