

Chapter 2

Disentangling Selection and Causality in Assessing the Effects of Health Inputs on Child Survival: Evidence from East Africa

Gebrenegus Ghilagaber

2.1 Introduction

Many demographic data have a hierarchical or clustered structure. For example, the analysis of childhood mortality involves a natural hierarchy where children are grouped within mothers or families, and the latter, in turn, are grouped into communities. Children from the same parents tend to be more alike in their characteristics than children chosen at random from the population at large. To ignore this grouping risks overlooking the importance of group effects, and may render invalid many of the traditional statistical analysis techniques used for studying data relationships.

The present chapter addresses the relationship between childhood mortality on the one hand, and use of health care and other socioeconomic variables on the other, in three African countries – Egypt, Eritrea, and Uganda. In contrast to most previous works where the collection of children is assumed to be an independent random sample, we treat children with the same mother as correlated cases (level 1) within the same mother (level 2). This is consistent with the data collection where a nationally representative random sample of women is selected (National Statistics Office [Eritrea] and Macro International Inc 1995). Our formulation also enables us to allow for unobserved mother-specific heterogeneity in the models.

A second and important issue that is addressed in this chapter is that of selection bias. The public policy response to the problem of high childhood mortality in developing countries has primarily focused on encouraging prenatal care and institutional delivery. Since there are no randomized trials of standard prenatal care and hospital delivery, it is difficult to assess the impact of such health inputs on

G. Ghilagaber (✉)

Department of Statistics, Stockholm University, SE-106 91 Stockholm, Sweden

e-mail: Gebre@stat.su.se

survival chances without accounting for selection processes in the utilization of health facilities.

Generally, selection bias can arise when there is a systematic difference in characteristics between those who are selected for the study and those who are not. If these unobserved factors also influence the impacts of the studied subject, selection bias will occur. However, selection bias will only arise if these unobserved factors, which influence the selection, also influence the phenomena under investigation.

Two common types of selection are adverse selection and favorable selection. Suppose women who receive prenatal health care have a higher risk of losing their child than women who do not receive prenatal health care. It is possible that women who seek prenatal childcare in fact have characteristics that separate them from others. If such characteristics are such that they lead to poorer outcome of prenatal care than it really is, then the contribution of prenatal care may be underestimated due to adverse selection.

In contrast, favourable selection arises when the studied individuals have characteristics that lead to overestimation of the effects of a covariate on the phenomenon under investigation.

In the present chapter, we examine the effects of selection on estimates of the efficacy of prenatal care and hospital delivery (health inputs) by using multiprocess models developed and earlier used by Lillard and Panis (see, Lillard 1993; Panis and Lillard 1994; Lillard and Panis 2000).

2.2 Statistical Methods: Multilevel and Multiprocess Modeling

2.2.1 A Piece-Wise Log-Linear Hazard Model with Heterogeneity

A piece-wise log-linear hazards model of mortality is given by¹:

$$\ln \lambda_{ij}(t) = \gamma T_{ij}(t) + \beta' X_{ij} + \varepsilon_i \quad (2.1)$$

where $\ln \lambda_{ij}(t)$ is the log-hazard of death at age t associated to child j of mother i . The baseline log-hazard $\gamma T_{ij}(t)$ is assumed to be piecewise linear in the child's age; X_{ij} represents regressors, and ε captures unobserved heterogeneity, at the mother level, that is associated with mortality, $\varepsilon \sim N(0, \sigma_\varepsilon^2)$. The regressors may be time-varying but all covariates used in this chapter are fixed. Regressors add to the log-hazard

¹The presentation in this and the next sections follows largely that of Lillard and Panis (2000).

and thus shift the hazard proportionally. Time is written as an argument in $T_{ij}(t)$ to indicate that it varies continuously over the duration of an interval but the slope may vary between intervals.

This is a two-level piece-wise linear survival model with mothers as the level-2 units and children as repeated outcomes (level-1) within observations.

Conditional on the heterogeneity component ε , the likelihood of the hazard for child j from the i th mother is:

$$L_{ij}^{(M)}(\varepsilon) = \begin{cases} S_{ij}(t^*, \varepsilon), & \text{if the child is alive (censored) at time } t^* \\ S_{ij}(t^l, \varepsilon) - S_{ij}(t^u, \varepsilon), & \text{if the child died between at } t^l \text{ and } t^u \end{cases} \quad (2.2)$$

where $S_{ij}(t, \varepsilon)$ is the survivor function at time t . In the absence of time-varying covariates,

$$S_j(t, \varepsilon) = [S_{0j}(t, \varepsilon)]^{\exp(\beta' X_j + \varepsilon)}, \quad (2.3)$$

where

$S_{0j}(t, \varepsilon)$ representing the baseline survivor function at time t , i.e., the survivor function based on the baseline duration dependency (or dependencies) only:

$$S_{0j}(t) = \exp \left\{ - \int_{\tau=t_b}^t \lambda_{j0}(\tau) d\tau \right\},$$

where $\lambda_{j0}(t) = \gamma T_j(t)$ and t_b denotes the beginning of the hazard spell (interval).

Conditional on the heterogeneity, the likelihood contributions in (2.2) are independent. The joint likelihood of multiple hazard intervals in the presence of heterogeneity is thus found by the product of conditional likelihoods of individual hazard modules:

$$L^{(M)} = \prod_j L_j^{(M)} \quad (2.4)$$

The baseline duration pattern is the model's dependency on time without any covariates or heterogeneity. In the model above, it is represented by $\gamma T_{ij}(t)$. A constant baseline hazard (exponential model) may be achieved by defining a spline with intercept and without nodes, and fixing the slope coefficient to zero. A Gompertz (linear) log-hazard may be specified by defining a spline without nodes, so that the slope is the Gompertz slope. A piecewise-constant hazard may be achieved by estimating regression coefficient on time-varying indicator variables. Piecewise-linear duration patterns are very attractive because they adjust to any pattern in the data (with sufficiently many nodes), and because linear combinations of piecewise-linear patterns are again piecewise-linear (Lillard and Panis 2000).

2.2.2 Multilevel Probit Models with Unobserved Heterogeneity

A Probit Model of Prenatal Care

We model the i th mother's decision to visit a prenatal care center (as opposed to no such visit at all) during pregnancy of the j th child as a binary probit model:

$$P_{ij}^* = \alpha' X_{ij} + \delta_i \quad (2.5)$$

where X are mother-specific explanatory variables ($X_{ij} = X_i$ for all children j of the same mother); and δ_i represents unobserved heterogeneity at the mother level that is associated with utilization of prenatal care. We assume that the heterogeneity component is distributed normally, $\delta \sim N(0, \sigma_\delta^2)$. Thus, the likelihood for a binary probit model is given by

$$L_j^{(P)} = \begin{cases} \Phi(-\alpha' X), & \text{if } P_j = 0 \\ 1 - \Phi(-\alpha' X), & \text{if } P_j = 1 \end{cases}, \quad (2.6)$$

where $\Phi(\cdot)$ is the (cumulative) distribution function of the standard normal density:

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z \exp\left(-\frac{u^2}{2}\right) du \quad (2.7)$$

The Prenatal Care decision for child j is:

$$P_j = \begin{cases} 0, & \text{if } P_{ij}^* < 0 \text{ (no prenatal care)} \\ 1, & \text{if } P_{ij}^* \geq 0 \text{ (prenatal care)} \end{cases} \quad (2.8)$$

Conditional on the heterogeneity, the likelihood contributions in (2.6) are independent. The joint likelihood of multiple probit modules in the presence of heterogeneity is thus given by the product of conditional likelihoods of individual probit contributions:

$$L^{(P)} = \prod_j L_j^{(P)} \quad (2.9)$$

A Probit Model of Hospital Delivery

As with the prenatal care we model the decision to deliver in hospital (as opposed to home delivery) as a binary probit model:

$$H_{ij}^* = \phi' X_{ij} + \omega_i \quad (2.10)$$

where X are mother-specific explanatory variables, and ω represents unobserved heterogeneity at the mother level. We assume that the heterogeneity component is distributed normally, $\omega \sim N(0, \sigma_\omega^2)$.

The delivery decision for child j is:

$$H_j = \begin{cases} 0, & \text{if } H_j^* < 0 \text{ (delivery at home)} \\ 1, & \text{if } H_j^* \geq 0 \text{ (institutional delivery)} \end{cases} \quad (2.11)$$

The likelihood for a binary probit model (module) is then

$$L_j^{(H)} = \begin{cases} \Phi(-\phi'X), & \text{if } H_j = 0 \\ 1 - \Phi(-\phi'X), & \text{if } H_j = 1 \end{cases} \quad (2.12)$$

where $\Phi(\cdot)$ is the distribution function of the standard normal density as given in (2.8).

Conditional on the heterogeneity, the likelihood contributions in (2.12) are independent. The joint likelihood of multiple probit modules in the presence of heterogeneity is thus given by the product of conditional likelihoods of individual probit modules:

$$L^{(H)} = \prod_j L_j^{(H)} \quad (2.13)$$

2.2.3 Multiprocess Models: Disentangling Selection and Causality

A Joint Model of Child Mortality and Prenatal Care

Suppose we estimated a hazard model of child mortality and found a significant evidence of unobserved mother-specific characteristics that affect children's survival. If the mothers themselves are aware of at least some of those characteristics, they may respond to this private knowledge. Suppose that those women who are at above-average risk of losing their baby decide to reduce the risks by visiting prenatal care centers. The result will, then, be that prenatal care centers get a disproportionately high-risk mix of babies. If ignored, this adverse selection will underestimate the beneficial effect of prenatal care on childhood mortality. Conversely, prenatal care centers may get disproportionately low-risk mix of babies. This happens when selection is favorable – that women with below-average risk of losing their babies have a higher propensity of visiting prenatal care centers. These may include more educated women who are more aware of the benefits of prenatal care and/or urban residents for whom access is relatively easier. In this later type of selection, ignoring

the favorable selection will overestimate the effect of prenatal care. These problems prompt us to address the potential endogeneity of prenatal care and estimate a joint model of child mortality and prenatal care decisions.

The joint model consists of two sets of equations:

- A hazard of child mortality:

$$\ln \lambda_{ij}(t) = \gamma T_{ij}(t) + \beta' X_{ij} + \varepsilon_i \quad (2.14)$$

- A probit of prenatal care:

$$P_{ij}^* = \alpha' X_{ij} + \delta_i \quad (2.15)$$

The main issue addressed here is that we wish to allow for the possibility that unobserved mother-specific characteristics affect both child survival and prenatal care decisions, i.e., we wish to allow for correlation between ε and δ :

$$\begin{pmatrix} \varepsilon \\ \delta \end{pmatrix} \sim N \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_\varepsilon^2 & \sigma_{\varepsilon\delta} \\ \sigma_{\varepsilon\delta} & \sigma_\delta^2 \end{pmatrix} \right] \quad (2.16)$$

The bias due to selection effects is eliminated by making the source of the bias (the correlation) part of the model. In our present case, the effect of prenatal care on mortality may be biased because of non-random prenatal care decisions. We therefore estimate a joint or multiprocess (to borrow a word from Lillard and Panis 2000) model of child survival and the decision to visit a prenatal care center.

The joint likelihood of the continuous and probit outcomes may be separated into a continuous and a probit part:

$$L^{(MP)} = L_1^{(M)} L_2^{(P)} \quad (2.17)$$

where

$$L_1^{(M)} = \frac{1}{\sigma_\varepsilon \sqrt{2\pi}} \exp \left\{ -\frac{(y - \beta' X)^2}{2\sigma_\varepsilon^2} \right\} \quad (2.18)$$

and

$$L_2^{(P)} = \begin{cases} \Phi \left(\frac{\mu_{\delta|\varepsilon} - \alpha' X}{\sigma_{\delta|\varepsilon}} \right), & \text{if } P = 0 \\ 1 - \Phi \left(\frac{\mu_{\delta|\varepsilon} - \alpha' X}{\sigma_{\delta|\varepsilon}} \right), & \text{if } P = 1 \end{cases} \quad (2.19)$$

where the distribution of $\delta|\varepsilon$ is such that:

$$\begin{pmatrix} \varepsilon \\ \delta \end{pmatrix} \sim N \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_\varepsilon^2 & \\ & \sigma_{\varepsilon\delta} \sigma_\delta^2 \end{pmatrix} \right] \quad (2.20)$$

so that

$$\delta|\varepsilon \sim N \left[\frac{\sigma_{\varepsilon\delta}}{\sigma_\varepsilon^2} (\ln \lambda - \beta' X), \sigma_\delta^2 - \frac{\sigma_{\varepsilon\delta}^2}{\sigma_\varepsilon^2} \right] \quad (2.21)$$

From (2.17), (2.18), (2.19), (2.20), and (2.21) we note that the probit residual δ is conditional on the realized value of ε and, hence, $L_2^{(P)}$ is conditional on $L_1^{(M)}$.

A Joint Model of Child Mortality and Hospital Delivery

By analogous argument to the above subsection we address the potential endogeneity of institutional delivery by estimating a joint model of child mortality and hospital delivery decisions.

- A hazard of child mortality:

$$\ln \lambda_{ij}(t) = \gamma T_{ij}(t) + \beta' X_{ij} + \varepsilon_i \quad (2.22)$$

- A probit of hospital delivery:

$$H_{ij}^* = \phi' X_{ij} + \omega_i \quad (2.23)$$

Again, the joint likelihood of the continuous and probit outcomes may be separated into a continuous and a probit part,

$$L^{(MH)} = L_1^{(M)} L_2^{(H)} \quad (2.24)$$

where $L_1^{(M)}$ is as defined in (2.18) and

$$L_2^{(H)} = \begin{cases} \Phi \left(\frac{\mu_{\omega|\varepsilon} - \alpha' X}{\sigma_{\omega|\varepsilon}} \right), & \text{if } H = 0 \\ 1 - \Phi \left(\frac{\mu_{\omega|\varepsilon} - \alpha' X}{\sigma_{\omega|\varepsilon}} \right), & \text{if } H = 1 \end{cases} \quad (2.26)$$

where the distribution of $\omega|\varepsilon$ is such that:

$$\begin{pmatrix} \varepsilon \\ \omega \end{pmatrix} \sim N \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_\varepsilon^2 & \\ & \sigma_{\varepsilon\omega} \sigma_\omega^2 \end{pmatrix} \right] \quad (2.27)$$

so that

$$\omega|\varepsilon \sim N \left[\frac{\sigma_{\varepsilon\omega}}{\sigma_{\varepsilon}^2} (\ln \lambda - \phi'X), \sigma_{\omega}^2 - \frac{\sigma_{\varepsilon\omega}^2}{\sigma_{\varepsilon}^2} \right] \quad (2.28)$$

We now wish to investigate whether unobserved characteristics at the mother level that affect the prenatal care decision are correlated with those that affect the decision to deliver in hospital. If these characteristics are correlated and the correlation is not accounted for, the effects of prenatal care and hospital deliveries on child mortality may be incorrect because these two effects may compete with each other or reinforce each other depending on the direction of the correlation.

The next step is, therefore, to estimate the hazard of child mortality jointly with both prenatal care and hospital delivery in order to control for the correlation between unobserved characteristics that affect these two health care decisions.

A Joint Model of Child Mortality, Prenatal Care, and Hospital Delivery

The effect of prenatal care and hospital delivery on mortality may be biased because of non-random prenatal care and hospital delivery decisions. More importantly, these effects may be biased because of a disproportionately high number of hospital deliveries with mothers who have visited a prenatal care center. We, therefore, model prenatal care and hospital delivery decisions jointly with the hazard of mortality.

The three-process joint model consists of three sets of equations:

- A hazard of child mortality:

$$\ln \lambda_{ij}(t) = \gamma T_{ij}(t) + \beta' X_{ij} + \varepsilon_i \quad (2.29)$$

- A probit of prenatal care:

$$P_{ij}^* = \alpha' X_{ij} + \delta_i \quad (2.30)$$

- A probit of hospital delivery:

$$H_{ij}^* = \phi' X_{ij} + \omega_i \quad (2.31)$$

The key issue here is that we wish to allow for the possibility that unobserved mother-specific characteristics affect all three dimensions: child survival, prenatal care, and hospital delivery decisions. In other words, the mother-specific heterogeneities in the three models (ε , δ and ω) are allowed to be pairwise correlated:

$$\begin{pmatrix} \varepsilon \\ \delta \\ \omega \end{pmatrix} \sim N \left[\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{\varepsilon}^2 & & \\ \sigma_{\varepsilon\delta} & \sigma_{\delta}^2 & \\ \sigma_{\varepsilon\omega} & \sigma_{\delta\omega} & \sigma_{\omega}^2 \end{pmatrix} \right]$$

The joint likelihood in the three process model is given as a product of the three likelihoods in (2.18), (2.19), and (2.26).

2.3 Data and Correlates of Childhood Mortality in Egypt, Eritrea, and Uganda

2.3.1 Data Sources

Egypt

The 1995 Egypt Demographic and Health Survey is the third survey in a series of Demographic and Health surveys that have been carried out in Egypt. The survey is a nationally-representative survey of 14,779 ever married women aged 15–49. These women gave birth to a total of 56,681 children but information on antenatal care visits and assistance, etc. is available only for children born within 5 years before the survey. For the purpose of this chapter, therefore, we concentrate on these 12,051 children from 8,008 mothers. For comparison purposes we also use a subset of 7,483 children from 6,140 mothers who were born within 3 years before the survey. Details concerning the 1995 Egyptian Demographic and Health Survey is documented in El-Zanaty et al. (1996).

Eritrea

The data used for illustration in the present section come the 1995 Demographic and Health Surveys (DHS) in the three countries.

The Eritrean Demographic and Health Survey (EDHS) is a nationally-representative survey of 5,054 women age 15–49 and 1,114 men age 15–59. It is the first survey ever undertaken by the National Statistics Office (NSO) of the Department of Macro Policy and International Economic Cooperation, Office of the President. It was implemented through the worldwide Demographic and Health Surveys (DHS) program of Macro International Inc.

One of the main objectives of the EDHS was to collect reliable data on maternal and child health indicators among children in early ages of their life. These include, among others, antenatal care visits and assistance at delivery. While the 5,054 women had a total of 14,268 children, information on antenatal care visits and assistance, etc. is available only for children born within 3 years before the survey. After deleting children with incomplete information on important factors we were left with 2,284 children belonging to 1,969 mothers. The maximum number of children per women was 3.

More details concerning the EDHS sample design, estimations of sampling errors for selected variables as well as summary tabulations are provided in National

Statistics Office (Eritrea), and Macro International Inc. (1995). In the present illustration we use the 1,969 mothers as experimental units and treat the 2,284 children as levels nested within these 1,969 mothers.

Uganda

The 1995 Uganda Demographic and Health Survey is a second survey in a series of Demographic and Health surveys that have been carried out in that country. It is a nationally-representative survey of 7,070 women age 15–49. By the survey time these women had a total of 22,752 children but the usable records for the purpose of this chapter are the 5,677 children from 3,988 mothers who were born within 4 years before the survey and a subset of it – the 4,533 children from 3,670 women who were born within 3 years before the survey. Tables of preliminary results and other details on the 1,995 Uganda Demographic and Health Survey may be found in Statistics Department [Uganda] and Macro International Inc. (1996).

2.3.2 *Correlates of Child Mortality*

The dependent variable is the log-hazard (logarithm of the rate at which the event of death occurs). The time variable (duration) measures the number of months from birth to death or the survey date, whichever comes first. Time varies between 0 and 35 months. The period between 0 and 35 months was partitioned into four: (0, 1), (1, 6), (6, 12), and (12, 35). The slope of the log-hazard was assumed to be constant within each interval but may vary between intervals. Additional models with time ranging between 0 and 47 months (in Uganda) and between 0 and 59 months (in Egypt) were also fit. In these models additional intervals for the time variable were used. These were (12, 24) in the case of Uganda and (12, 24), and (24, 36) in the case of Egypt.

Three mother-specific and five child-specific variables were used as explanatory variables.

The mother-specific variables are:

- X_1 – Mother's Age-group at survey time (15–19, 20–24, 25–29, . . . , 45–49).
- X_2 – Mother's Level of Education (None, Primary, Secondary or higher).
- X_3 – Residence (Urban, Rural).

The child-specific variables are:

- X_4 – Preceding birth interval (First born, <18, 18–29, 30–47, 48+ months).
- X_5 – Prenatal Care during pregnancy (None, Some prenatal care).
- X_6 – Place of delivery of index child (Home, Hospital or clinic).
- X_7 – Sex of index child (Girl, Boy).
- X_8 – Multiplicity of index child (Single birth, One of multiple births).

The first level of each covariate was used as baseline (reference) level and, thus, no estimates are reported for these levels.

These variables are among those considered to be correlated with childhood mortality in previous analyses of the same data set (for the Eritrean data) or other data sets. However, because our main aim is of methodological nature, we have not strived to include all relevant covariates of mortality suggested in the literature or discuss the theoretical expectations of the effects of the covariates included in the analysis.

2.4 Results

2.4.1 *Covariates Effects*

The results of fitting the various models described in Sect. 2.2 to data on Egypt, Eritrea, and Uganda, are shown in Tables 2.1, 2.2, and 2.3, respectively. The five columns of results refer to the following situations:

- Model 1 refers to the multilevel piecewise log-linear hazards model in (2.1) but without heterogeneity term.
- Model 2 refers to the multilevel piecewise log-linear hazards model in (2.1) with mother-specific heterogeneity.
- Model 3 refers to the multilevel multiprocess model where (2.14) and (2.15) – that is a hazard model of child mortality and a probit model of prenatal care – are estimated simultaneously; we allow for mother-specific unobserved heterogeneity in both models; and allow these two heterogeneity terms to correlate.
- Model 4 refers to the multilevel multiprocess model where (2.22) and (2.23) – that is a hazard model of child mortality and a probit model of hospital delivery –

Table 2.1 Estimates of effects of prenatal care and hospital delivery on log-hazards of mortality (M) in various models: **Egypt (1995)**

| Parameters | Model 1 (No Hetro) | Model 2 (Hetro) | Model 3 (M and P) | Model 4 (M and H) | Model 5 (M, P & H) |
|----------------------------|-----------------------|--------------------|----------------------|----------------------|-----------------------|
| β_P | −0.3493* | −0.3959* | −1.4141* | −0.4037* | −1.2112* |
| β_H | 0.0748 | 0.0945 | 0.0800 | −0.6066* | −0.5283* |
| σ_ε | – | 0.7393* | 0.9082* | 0.7408* | 0.8469* |
| σ_δ | – | – | 2.2458* | – | 2.6563* |
| $\rho_{\varepsilon\delta}$ | – | – | 0.7531* | – | 0.6314* |
| σ_ω | – | – | – | 2.6997* | 2.2649* |
| $\rho_{\varepsilon\omega}$ | – | – | – | 0.6182* | 0.7741* |
| $\rho_{\delta\omega}$ | – | – | – | – | 0.4786* |

*Estimate significant at 10 % significance level

Table 2.2 Estimates of effects of prenatal care and hospital delivery on log-hazards of mortality (M) in various models: **Eritrea (1995)**

| Parameters | Model 1 (No Hetro) | Model 2 (Hetro) | Model 3 (M and P) | Model 4 (M and H) | Model 5 (M, P & H) |
|----------------------------|-----------------------|--------------------|----------------------|----------------------|-----------------------|
| β_P | -0.4624* | -0.4928** | 0.4632 | -0.5038* | 0.4432 |
| β_H | -0.1142 | -0.1107 | -0.0920 | -0.7188 | 0.1190 |
| σ_ε | — | 1.2488*** | 1.3324*** | 1.2367*** | 1.3165*** |
| σ_δ | — | — | 1.3193*** | — | 1.2913*** |
| $\rho_{\varepsilon\delta}$ | — | — | -0.5265* | — | -0.5412* |
| σ_ω | — | — | — | 1.9301*** | 1.7889*** |
| $\rho_{\varepsilon\omega}$ | — | — | — | 0.3138** | -0.2285 |
| $\rho_{\delta\omega}$ | — | — | — | — | 0.6233*** |

*Estimate significant at 10 % level; **Estimate significant at 5 % level; ***Estimate significant at 1 % level

Table 2.3 Estimates of effects of prenatal care and hospital delivery on log-hazards of mortality (M) in various models: **Uganda (1995)**

| Parameters | Model 1 (No Hetro) | Model 2 (Hetro) | Model 3 (M and P) | Model 4 (M and H) | Model 5 (M, P & H) |
|----------------------------|-----------------------|--------------------|----------------------|----------------------|-----------------------|
| β_P | -0.4163* | -0.4280* | -0.1284 | -0.4258* | -0.1713 |
| β_H | -0.2945* | -0.2996* | -0.2942* | 0.0271 | -0.0211 |
| σ_ε | — | 0.5118* | 0.4417** | 0.5180* | 0.4899* |
| σ_δ | — | — | 1.9376* | — | 2.0197* |
| $\rho_{\varepsilon\delta}$ | — | — | -0.3654*** | — | -0.4036*** |
| σ_ω | — | — | — | 2.0799* | 1.8867* |
| $\rho_{\varepsilon\omega}$ | — | — | — | -0.4099* | -0.3525** |
| $\rho_{\delta\omega}$ | — | — | — | — | 0.4625* |

*Estimate significant at 10 % level; **Estimate significant at 5 % level; ***Estimate significant at 1 % level

are estimated simultaneously; we allow for mother-specific unobserved heterogeneity in both models; and allow these two heterogeneity terms to correlate.

- Model 5 refers to the multilevel multiprocess model where the three Eqs. (2.29), (2.30), and (2.31) – that is a hazard model of child mortality and two probit models for hospital delivery and prenatal care, respectively, – are estimated simultaneously; we allow for mother-specific unobserved heterogeneity in all three models, and allow for pairwise correlation between these three heterogeneity terms.

We have reported results related to hazard models alone and left out those from probit models. Further, only estimates of Prenatal care and Hospital delivery are presented in the Tables while estimates of the other background variables are suppressed.

We can, however, mention that, in the Egyptian case for instance, children from older cohort of mothers (aged 35 years or above at the time of the survey) had higher

mortality risks than children from the very youngest cohort (15–19 years at survey time). Further, children of mothers with higher education (secondary or above level) had lower mortality; that 2nd and higher order births with short preceding birth intervals (< 18 months) had higher risks than first born children, while those born after long interval (at least 30 months) had significantly lower risks. The results for Eritrea and Uganda, in terms of the unreported covariate effects, were not much different.

2.4.2 Selection Bias in Prenatal Care Utilization

Again, beginning with Egypt (Table 2.1), a comparison of Models 2 and 3 shows that while both models show a significant beneficial effect of prenatal care on child mortality hazard, the magnitude is underestimated in the separate specification (from -1.4141 to -0.3959). This, again, is due to the positive correlation (0.7531) between the unobserved mother-specific characteristics that affect childhood mortality risks and the decision to visit a prenatal care during pregnancy. Thus, we can say that there is also adverse selection into prenatal care, and failure to account for this selectivity severely underestimates the magnitude of the beneficial effect of prenatal care.

The effect of selection in prenatal care is in the opposite direction in Eritrea. While separate specification (Model 2) shows a marginally significant beneficial effect of prenatal care (-0.4928), joint modeling (Model 3.) shows a positive but insignificant effect (0.4632). The correlation between the unobserved mother-specific characteristics that affect childhood mortality risks and the prenatal care is negative (-0.5265) and it is this negative correlation that pushed the effect of prenatal care far to the left of zero. In any case, we note that there is a mild favorable selection to prenatal care in Eritrea.

In the case of Uganda the selection bias is in the same direction as in Eritrea but it is stronger. The relatively weak and negative correlation (-0.3654) inflates the effect of prenatal care from an insignificant value (-0.1284) to a strongly significant effect (-0.4280) if this favorable selection is not accounted for.

2.4.3 Selection Bias in Hospital Delivery

Beginning with Egypt (Table 2.1), while the separate specification (Model 2) shows an insignificant and positive effect of hospital delivery on child mortality hazard (0.0748), joint estimation (Model 4) reveals a highly significant and strong negative effect (-0.6066). As stated in Lillard and Panis (2000), the mechanical reason lies in the positive correlation (0.6182) between the unobserved mother-specific characteristics that affect childhood mortality risks and the decision to delivery a child in hospital. An ignored positive correlation biases parameter estimates in positive direction, i.e., toward zero in the present case. Substantively, women

with above-average risks of losing a baby ($\varepsilon > 0$) also tend to have above-average propensities to deliver in a hospital ($\omega > 0$); and vice versa. In other words, there is adverse selection into hospital delivery, and failure to account for this selectivity severely underestimates the beneficial effect of hospital delivery.

The effect of selection bias is in the same direction in Eritrea as well (Table 2.2) but the effect is milder in the case of Eritrea than in Egypt. While the estimate changes from -0.1107 to -0.7188 , it is insignificant in both Models 2 and 4. This is due to the relatively weaker correlation (0.3138) between the unobserved mother-specific characteristics that affect childhood mortality risks and the decision to deliver a child in hospital, in the case of Eritrea.

A different picture is depicted in the case of Uganda (Table 2.3). To begin with, the correlation between the unobserved mother-specific characteristics that affect childhood mortality risks and the decision to deliver a child in hospital is negative (-0.4099) in the case of Uganda.

This implies that, women with above-average risks of losing a baby ($\varepsilon > 0$) tend to have below-average propensities to deliver in a hospital ($\omega < 0$); and vice versa. Thus, the effect of hospital delivery on child mortality shifts from a highly significant beneficial effect (-0.2996) to an insignificant effect (0.0271). In other words, there is favorable selection into hospital delivery in Uganda, and failure to account for this selectivity severely overestimates the effect of hospital care.

2.4.4 Correlation Between Prenatal Care Utilization and Hospital Delivery

The results in the above subsections indicated that there is significant correlation between the mother-specific unobserved heterogeneities in the hazard and probit models and that failure to account for such correlation would bias the parameter estimates of the effects of prenatal care and hospital delivery.

An important question that still remains to be answered is as to whether the two decisions (prenatal care and hospital delivery) are also correlated. It is quite likely that mothers who visited prenatal care centers during pregnancy would have a higher propensity to deliver their child in hospital than mothers who never did so. The result will, then, be that delivery centers get a disproportionately high proportion of babies whose mothers have visited prenatal care centers during pregnancy. If ignored, it would be difficult to distinguish between the relative strengths of the effects of prenatal care and hospital delivery on child mortality when both effects are considered together.

We therefore address the potential endogeneity of both decisions and estimate a joint (three-process) model of child mortality, prenatal care, and hospital delivery decisions.

The results from such a three-process models are shown in the last column (Model 5) of Tables 2.1, 2.2, and 2.3. In all three tables we see, as expected, that there is a highly significant positive correlation between the decisions of visiting

prenatal care during pregnancy and delivering the child in hospital. How does this affect the relative magnitudes of the effects of these two endogenous factors on the risk of childhood mortality?

In the case of Egypt Table 2.1 we note that failure to account for this positive correlation raises the magnitude of effect of both factors – from -1.2112 (Model 5) to -1.4141 (Model 3.) for prenatal care, and from -0.5283 (Model 5) to -0.6066 (Model 4) for hospital delivery.

The same is true for Eritrea Table 2.2 – from 0.4432 (Model 5) to 0.4632 (Model 3.) for prenatal care, and from 0.1190 (Model 5) to -0.7188 (Model 4) for hospital delivery.

In Uganda Table 2.3 the changes are from -0.1713 (Model 5) to -0.1284 (Model 3.) for prenatal care, and from -0.0211 (Model 5) to 0.0271 (Model 4) for hospital delivery.

2.4.5 Comparison of the Standard Model and the Multiprocess Model with Unobserved Heterogeneity and Correlated Health Input Variables

As a final remark in this section, it may be worth examining what happens to the effects of prenatal care and hospital delivery on the log-hazard of childhood mortality as we move from the standard model (Model 1) to the final model (Model 5). The changes in estimates of such effects may be examined by comparing the estimates in columns 1 and 5 of Tables 2.1, 2.2, and 2.3.

The results for Egypt Table 2.1 show that while the standard model reports a significant beneficial effect of prenatal care (-0.3493) but no effect of hospital delivery (0.0748), the final model, where selection and correlation are accounted for, shows that both factors have significant beneficial effects (-1.2112 and -0.5283 , respectively). Thus, it seems that at least part of the effect of hospital delivery was transferred to that of prenatal care in the standard model, which does not account for the correlation between these two factors.

The Eritrean case Table 2.2 shows to the contrary. While the true picture is that the two health inputs have no beneficial effects (0.4432 and 0.1190 , respectively for prenatal care and hospital delivery), failure to account for selection bias into these two processes and the correlation between them would lead to concluding that one of them (prenatal care) has strong beneficial effect (-0.4624) while the other (hospital care) has no effect at all (-0.1142).

Ugandan results Table 2.3 show another interesting case. The right picture (Model 5) is that none of these two health inputs has any beneficial effect on childhood mortality (-0.1713 and -0.0211 , respectively for prenatal care and hospital delivery). If one ignores selection biases and the correlation between the two health inputs, however, one would be led to the erroneous conclusion that the two health inputs have highly significant beneficial effects in reducing childhood mortality (-0.4163 and -0.2945 , respectively).

Are these changes statistically significant? This question may be answered by comparing the differences in log-likelihoods in the models under consideration because the models are nested within the next higher model.

It may also be of interest to examine the effects of unobserved heterogeneity, selection bias, and correlation between health input variables, affects the effects of the exogenous variables Education and Residence. A priori, one would suspect that these two variables are correlated with the health input variables (prenatal care and hospital delivery) because we expect the more educated women and those in urban areas would have a higher propensity to use health care facilities.

The results (details not shown here) show that as long as Prenatal Care is treated as exogenous variable, Education (at higher level) continues to have beneficial effects in reducing childhood mortality in the Egyptian data set (with estimates -0.4188 , -0.4429 , and -0.2279 , respectively in Models 1, 2, and 4). Once we treat Prenatal Care as endogenous variable (Model 3) and/or account for its correlation with Hospital Delivery (Model 5), however, the beneficial effect of Education fades away (the estimate reduces to -0.0916 and 0.0180 , respectively, in Models 3 and 5). One would, thus, be tempted to suspect that the effects of Education, at least in Egypt work, via higher propensity of educated women to make use of prenatal care centers. But we also need to reconcile this suspicion with our earlier results of adverse selection into prenatal care. The effect of Residence is more blurred though there is marginal evidence that the estimate shifts from insignificant difference towards rural advantages in childhood mortality when proper care is taken of selection effects.

The opposite is true in Eritrea. Results from model 1, 2 and 4 show that there is no effect of education on the hazard of childhood mortality. Once Prenatal Care is treated at endogenous variable and/or its correlation with Hospital Delivery is accounted for, it turns out that Education (now at primary level) has a strong beneficial effect in reducing the risk of childhood mortality. We already know that selection into prenatal care is favorable in Eritrea prompting that it is the more educated women who benefit from such services. Thus, accounting for such favorable selection brings to the surface the true and beneficial effect of education on childhood mortality risks. The effect of Residence is also interesting in the case of Eritrea. The standard model (Model 1) shows that Rural areas have significantly lower risks of childhood mortality than urban areas. In Models 3 and 5, however, it is shown that there are no differential mortality risks by mother's place of residence. We also know that we have favorable selection into prenatal care and that we suspect this would be so due to the fact that urban residents benefit more from prenatal care centers than their rural counterparts. Thus, failure to account for this selection would have underestimated the urban advantage.

Uganda provides another interesting result. Here, there is relatively weaker impact of our procedure on the effects of Education and Residence. If any, it is when we account for Hospital Delivery that the effects of Education are strengthened. It may be noted that there is a stronger correlation between the heterogeneity terms of Hospital Delivery and Mortality in Uganda than in Egypt and Eritrea. On the other hand, there is a weaker correlation between the heterogeneity terms of Prenatal Care and Mortality in Uganda than in Egypt and Eritrea.

The effects of accounting for selection biases and correlation between health input variables are relatively minor on the other exogenous variables (Interval, Sex, and Multiplicity) that we don't give much space to discuss them.

2.5 Summary and Concluding Remarks

For the last two decades Demographic and Health Surveys have been collected to provide information on family planning, maternal and child health, child survival, and reproductive health in Africa, Asia, the Near East, Latin America, and the Caribbean. The availability of such surveys has helped to shift the focus of investigations from indirect methods of estimation of summary measures to the use modern analytic methods in order to examine correlates of demographic behavior and their policy implications.

The surveys have been collected hierarchically at the family, household, and community levels. However, not many analysts seem to be aware of this nature of the data. The data in the surveys are collected by interviewing a nationally representative sample of women (and men in some cases). These women are independent observations once we account for their communities. Thus, in the analysis of marriage behavior, using these women as experimental units is a correct procedure.

In the analysis of childhood mortality, however, the situation is different. To analyze childhood mortality the original women data is converted to child data. In so doing a number of children are nested within the same woman (mother) and the data on children no longer consists of independent (random) observations unless we select just one child, say of a given birth order, from each mother. Children of the same mother are more alike than children selected at random from the population and analytical methods must pay due attention to this nature of the data.

Other issues of concern in the analysis of Demographic and Health Surveys Data include accounting for correlation structure among various determinants (such as that between death of previous child and preceding birth interval) and, more importantly, selection biases in the utilization of health facilities.

The present chapter attempts to address some of above issues through analyses of childhood mortality in three African countries – Egypt, Eritrea, and Uganda based on their 1995 DHS data.

In contrast to previous approaches where children are used as independent experimental units, we have treated children of the same mother as correlated cases (multilevels) within the same experimental unit (mother). We have also allowed for mother-specific unobserved heterogeneity at the mother level. Further, we have paid due account to selection into health care utilization by treating health care variables like prenatal care and hospital delivery as endogenous variables and modeling them simultaneously with the hazard of mortality.

Our results show that there are significant mother-level heterogeneities in the three countries. More interestingly, we have demonstrated that while there are

selection biases of health care utilization in all three countries, their effects and, hence, policy implications are different. In one of the countries (Egypt) we have shown that there is adverse selection bias and failure to account for this selection underestimates the beneficial effects of health care inputs. In the other two countries (Eritrea and Uganda) the selection is that of favorable selection and failure to account for it overstates the effect of health inputs.

We have also accounted for the possible correlation between the various health input variables and demonstrated that failure to account for such correlation would benefit one of the variables at the expense of the other. Further, we have demonstrated how the effects of exogenous variables like education and residence may be under/over-estimated if proper care is not taken to address selection into health care utilization.

References

- El-Zanaty, F., Hussein, E. M., Shawky, G. A., Way, A. A., & Kishor, S. (1996). *Egypt demographic and health survey 1995*. Calverton: National Population Council [Egypt] & Macro International Inc.
- Lillard, L. A. (1993). Simultaneous equations for hazards: Marital duration & fertility timing. *Journal of Econometrics*, 56, 189–217.
- Lillard, L. A., & Panis, C. W. A. (2000). aML multilevel multiprocess statistical software, Release 1.0. EconWare, Los Angeles.
- National Statistics Office [Eritrea], & Macro International Inc. (1995). *Eritrea demographic and health survey 1995*. Calverton: National Statistics Office [Eritrea] & Macro International Inc.
- Panis, C. W. A., & Lillard, L. (1994). Health inputs and child mortality: Malaysia. *Journal of Health Economics*, 13, 455–489.
- Statistics Department [Uganda], & Macro International Inc. (1996). *Uganda demographic and health survey 1995*. Calverton: Statistics Department [Uganda] & Macro International Inc.

Advanced Techniques for Modelling Maternal and Child
Health in Africa

Kandala, N.-B.; Ghilagaber, G. (Eds.)

2014, XIV, 330 p. 40 illus., 21 illus. in color.,

ISBN: 978-94-007-6778-2