UNOBSERVED HETEROGENEITY AND DEPENDENT COVARIATES: A STATE-SPACE MODEL OF INFANT GROWTH AND MORTALITY

Andrew Foster

December 1985
WP-85-87

This research was conducted in conjunction with a summer research seminar on heterogeneity dynamics, under the direction of James W. Vaupel and Anatoli I. Yashin, in the Population Program at IIASA led by Nathan Keyfitz. Funding provided by the American Academy of Arts and Sciences.

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
2361 Laxenburg, Austria

Foreword

A group of eleven Ph.D. candidates from seven countries--Robin Cowan, Andrew Foster, Nedka Gateva, William Hodges, Arno Kitts, Eva Lelievre, Fernando Rajulton, Lucky Tedrow, Marc Tremblay, John Wilmoth, and Zeng Yi--worked together at IIASA from June 17 through September 6, 1985, in a seminar on population heterogeneity. The seminar was led by the two of us with the help of Nathan Keyfitz, leader of the Population Program, and Bradley Gambill, Dianne Goodwin, and Alan Bernstein, researchers in the Population Program, as well as the occasional participation of guest scholars at IIASA, including Michael Stoto, Sergei Scherbov, Joel Cohen, Frans Willekens, Vladimir Crechuha, and Geert Ridder. Susanne Stock, our secretary, and Margaret Traber managed the seminar superbly.

Each of the eleven students in the seminar succeeded in writing a report on the research they had done. With only one exception, the students evaluated the seminar as "very productive"; the exception thought it was "productive". The two of us agree: the quality of the research produced exceeded our expectations and made the summer a thoroughly enjoyable experience. We were particularly pleased by the interest and sparkle displayed in our daily, hour-long colloquium, and by the spirit of cooperation all the participants, both students and more senior researchers, displayed in generously sharing ideas and otherwise helping each other.

A prize, the Peccei Fellowship, is awarded to the summer scientists who have excelled both in their own research and in helping other summer scientists with their research. This fellowship enables a summer scientist to return to IIASA for three months the following year. Three Peccei Fellowships were awarded this summer: Andrew Foster in the population seminar was one of the winners. This working paper summarizes the innovative research Foster carried out at IIASA. The research is not only mathematically sophisticated and demographically significant, but also policy relevant--quite an achievement.

James W. Vaupel
Anatoli I. Yashin

Acknowledgments

I would like to acknowledge my debt to fellow participants in the Population Program of the YSSP. Special thanks go to William Hodges and to the leaders of the summer program, James W. Vaupel and Anatoli I. Yashin. Also thanks to Susanne Stock who did the final formating of the manuscript.

Abstract

A policy oriented model of child growth and mortality is developed in the context of a stochastic state space model. The model incorporates unobserved heterogeneity as an unmeasured covariate which affects both mortality and an observed time varying covariate. It is demonstrated using Monte Carlo simulations that a model ignoring this unobserved heterogeneity will give biased parameter estimates; parameters are found to be unbiased if a model which allows for an unobserved variable is estimated. Monte Carlo simulations are then used to test the robustness of the model to misspecification of the distribution of the unobserved covariate. Estimates of the change in child survival are obtained using dynamic equations derived from a Kolmogorov-Fokker-Planck (KFP) equation. It is shown that the model which ignores unobserved heterogeneity produces incorrect estimates of the change in mortality that would result if certain types of mortality intervention programs were implemented.

UNOBSERVED HETEROGENEITY AND DEPENDENT COVARIATES: A STATE-SPACE MODEL OF INFANT GROWTH AND MORTALITY

Andrew Foster

Graduate Group in Demography
University of California at Berkeley
2234 Piedmont Avenue
Berkeley, California 94720
USA

1. Introduction

The work to date considering the effect of unobserved heterogeneity on parameter estimates of hazard models is quite diverse in terms of the applications, methodologies, and specifications considered. Nonetheless, one assumption has been retained in virtually all cases: unmeasured covariates are assumed to be independent of the measured covariates. From a theoretical perspective this limitation is certainly justified--it is not at all surprising that results should be biased when an unobserved variable is correlated with the observed variables. From the perspective of someone who wishes to use parameter estimates to inform policy, however, the independence assumption may be less desirable. In the first place, if the variables really are independent then the biased estimate may be more informative to the policy maker even when an unbiased estimate can be obtained. Conversely, if the assumption of independence is invalid then a model of heterogeneity which assumes independence may obscure important relationships and lead to inappropriate policy conclusions.

In this paper we develop a model of child mortality which incorporates unobserved heterogeneity which is related to an observed covariate. The model is presented in the context of the stochastic space model introduced by Woodbury and Manton (1977) and recently extended to include unobserved heterogeneity. (Yashin, Manton, and Vaupel, 1985; Yashin, 1984). Before presenting the details of the model, however, we provide a justification for its basic features by referring to recent work on infant and child mortality in the demographic and epidemiologi-

cal literature. Monte Carlo simulations are then used to compare parameter estimates from a model which ignores the unmeasured covariate of mortality with estimates from a model which incorporates the dependence of the measured and unmeasured covariates. The robustness of results to misspecification of the distribution of unobserved covariates is also explored. Finally, equations describing the dynamic properties of the moments of the distribution of the covariates of mortality are used to generate predictions of the impact on mortality of various policy alternatives using the parameters generated in the two models.

2. A General Model of Infant and Child Mortality

A general model of child mortality should consider three types of factors. First, mortality will be influenced by observable variables which tend to be fixed over the first few years of a child's life. The most typical example of such a covariate in a demographic study would be mother's education, but other covariates describing the general features of a child's household or community, such as whether there is a latrine or running water, will also be of this type. This kind of covariate is ignored in the model considered here, but when analyzing data, fixed covariates may be included in a relatively straightforward fashion.

Second, mortality will be affected by time varying covariates that are generated by a random process which is influenced by the values of other covariates as well as the age of the child. A typical example of such a covariate would be an anthropometric measure such as weight or some transformation of variables such as weight for age.

In an estimated model one might wish to include height, which is thought to be a measure of long term nutritional deprivation (stunting), and weight for height, which is thought to be a measure of short term deprivation (wasting). (See, for example, Mosely 1985.) In order to properly consider the effect of such covariates on mortality it is necessary to have longitudinal studies with anthropometric data taken at relatively short intervals, perhaps monthly.

Finally, from the point of view of this paper, it is desirable to have measures of time varying covariates which are, for any individual, stationary (in the sense of ARMA models) throughout the period of observation. For example, the measure $w^* = ln(w / w_s)$, where w_s is a standard weight for age taken from an appropriate schedule, may be approximately stationary. In practice one might want to try a number of transformations to see which transformation most closely approximates

stationarity.

The third set of factors affecting mortality are underlying unobserved covariates. These factors may, in principle, be either fixed or changing, but estimation is much more difficult if one does not believe that the unobserved covariates are fixed. Using the language of Vaupel et al. (1979) we may refer to a fixed unobserved covariate as frailty; however, the term frailty means something quite different for children growing up in a rural village of a poor country than it does, for example, for adults in developed countries. In the present context it is desirable to use the term frailty to consider not only unmeasured physical and mental attributes of an individual, but also the unmeasured components of the "disease environment" faced by a child. The motivation for considering frailty in this light can best be explained by a short review of a comparative study of child growth and survival in the rural areas of Costa Rica and Guatemala (Mata, 1973, 1985).

The basic subject of Mata's study is the interaction between disease and nutrition. While it it is generally agreed that poorly nourished children are more likely to succumb to infections than their well nourished counterparts, it is also thought that the level of nourishment of a child is likely to depend on the patterns of infection he faces. Mata's results demonstrate convincingly the importance of this second factor. The infant mortality rate in Cauqué, Guatemala was observed to be more that 6 times as high as that in Puriscal, Costa Rica, a circumstance which is attributable, for the most part, to the quite favorable public health environment in Puriscal, particularly the almost universal vaccination coverage, piped water, and use of of latrines and toilets. It was also observed that children of Cauqué had lower birth weights and less favorable growth then their counterparts in Puriscal.

Despite these very significant differences in the two towns, Mata demonstrated that diets of pregnant women and children in Cauqué were quite similar in quality and quantity to the diets in Puriscal. As such it seems that the low weights of the Cauqué children, as well as the high level of mortality, resulted not from inadequate feeding but from high levels of infection. Using the terminology discussed above, we would say that the children of Cauqué are especially frail because they face a less favorable disease environment. Unfavorable growth and higher mortality are both symptoms of this frailty.

In Mata's comparative study the disease environment in each town is reasonably well described by the series of variables considered. It is likely that at least at the level of aggregation considered, the towns themselves, the unobserved component of the the disease environment is relatively small. In general, however,

there will not be sufficient information for the model considered here where we wish to distinguish between the disease environment faced by different individuals rather than different towns.

At this point, the motivation for a model incorporating unobserved heterogeneity which is correlated with the observed variable becomes evident. Suppose we consider a policy which has the effect of raising a child's birth weight but no effect on the disease environment in which he spends his early years. If the birth weight is directly related to the mortality rate, and if both mortality and birth weight are affected by the disease environment, then we will overestimate the effect of the policy.

Before developing the model itself, it is helpful to consider the results of Trussell and Richards (1985) who used Heckman and Singer semi-parametric distributions to fit infant mortality data from Korea using several different time dependent hazard functions and sub-samples of the population. They found that parameter estimates for the mortality data were quite sensitive to the specification of the model as well as the sample considered. Noting that parameter estimates for fertility data from Korea were robust to changing specifications, they suggested that the volatility of the mortality model estimates may have resulted from the fact that relatively few closed intervals were observed in the mortality data.

A partial explanation is possible. When one considers a population in which few intervals are closed, there is relatively little scope for selection to affect results and therefore estimates of heterogeneity are likely to be unstable. Of course this fact taken by itself should not be a problem--one would expect that a maximum likelihood procedure would simply suggest that the addition of support points does not increase explanatory power. The authors found that they could add up to three support points and retain significance.

One explanation would attribute the observed results to the age pattern of mortality, which typically exhibits steep declines in the first few months or even weeks of life. While this drop in part may be attributed to heterogeneity, there are reasons to believe that the underlying curve is also reasonably steep. Since the parametric curves used by the authors are not able to capture this rapid decline well and still give a good fit to the remaining age specific mortality rates, an additional support point is required. Since the number of individuals dying in the first few months is small, the position of the support point is likely to be quite sensitive to changes in the sample; covariate parameter estimates may therefore also be sensitive. Since different curves have different abilities to capture the rapid

drop in mortality curves, we would also expect results to be sensitive to the form used for observed hazard function.

With this perspective it seems likely that the model developed here will not be subject to the same degree of instability encountered by Trussell and Richards. First, any population for which this model will be applicable is likely to have higher mortality than the Korean population Trussell and Richards considered. Since more events are observed there is more opportunity to observe selection. If the the estimation of unobserved heterogeneity depends on observing a selection effect, as it normally does, a population with higher mortality is likely to provide more stable results. A second fact is that the higher mortality populations that might be studied with the model discussed here are also likely to exhibit less rapid declines of mortality then the Korean population. The tendency for higher mortality populations to exhibit mortality curves which are relatively flat can be attributed to the prevalence of exogenous or environmentally determined mortality in high mortality populations in addition to the endogenous mortality which is experienced in all populations. The relative smoothness of infant mortality curves in high mortality population arises because exogenous infant mortality is concentrated at later ages than is endogenous mortality.

A far more important reason why the model developed here is not likely to exhibit the same instability is that unlike the more usual models of unobserved heterogeneity, the model being developed here provides estimates using factors other than the selection effect. If a substantial portion of the population survives to the end of the survey, as will be the case even in the highest mortality populations, it is essential to be able to rely on other factors than selection to estimate the underlying distributions of heterogeneity.

3. Model Specification

With this background we can construct the model that is the focus of the paper. As mentioned above, it is helpful to present the model in the context of the stochastic space framework first discussed by Woodbury and Manton (1977), especially when one wishes to consider the impact of various alternative policies on child survival, because the dynamics of the state variables over time may be represented by a Kolmogorov-Fokker-Planck equation (KFP).

The basic model considers an n dimensional state space through which individuals are postulated to move over time. Movement is governed in part by deterministic and in part by stochastic processes. Moreover, associated with each point in this space is a certain hazard of mortality.[1] Unfortunately, as analyzers of longitudinal survey data we are able to observe only $n-m$ dimensions; thus we must be content with analyzing the projections of the true movement of an individual onto this lower dimensional space.

The movement of an individual in this state space can be described using two equations, the equation of motion and the mortality equation. If x is an n dimensional vector and $\mu(t,x)$ is the mortality of an individual at point x at time t then the two equations can be written as follows:

$$dx = A(t,x)dt + B(t,x)d\varepsilon \tag{1}$$

$$\mu(t,x) = C(t,x) \tag{2}$$

where the term $d\varepsilon$ arises from a Wiener process. If $f_t(x)$ is the joint density function of x and the probability of the event that an individual survives to time t, then the following KFP equation governs the changing distribution of x over time (Woodbury and Manton, 1977):

$$\frac{df(x,t)}{dt} = f\ tr\left(\frac{dA}{dx}\right) + \frac{df}{dx}'A + tr\left(\frac{df}{dxdx'}BB'\right) - \mu f \quad .$$

In practice it is quite helpful to put several restrictions on the equations of motion. If we assume that $A(t,x)$ is linear in x, B is independent of x, and $C(t,x)$ is a proportional hazard which is quadratic in x then it will be the case that if the initial distribution of x is Gaussian then the distribution conditional on survival will be Gaussian at any future time. The value of a model which retains the distributional shape of x has been discussed by a number of authors (see, for example, Hougaard, 1982). In this paper the result is particularly helpful when we wish to make predictions about the effects on mortality of changing parameter values.

If x consists of one observed and one unobserved variable, and A is linear and C is quadratic only in the *unobserved* variable, then it will no longer be necessarily the case that the observed variable will be Gaussian; however, the unobserved variable conditional on the values of the observed variable up to time

[1] Actually Woodbury and Manton's model considered three levels of space, in the first of which mortality was entirely deterministic. For simplicity we focus on the second and third levels of space.

t will be Gaussian (Yashin et al., 1985). This is the basic model used in the estimation portion of the paper. It was selected because it is less restrictive than the completely Gaussian model discussed above and yet still has a likelihood function that can be evaluated without resorting to numerical integration.

Using this general framework we can describe the process of growth and mortality as follows: Assume that the transformed weight variable (w) follows a stationary stochastic process conditional on frailty (y) and that the frailty of an individual is fixed. Since mortality and weight are not likely to change very much over the course of a month, and since the chances of dying in the one month period are quite small, the continuous equations of motion are quite representative of the discrete equations that will be used in the analysis. Equations (1) and (2) become

$$\begin{aligned} \Delta w_t(w,y) &= (a_{11}w_t + a_{12}y) + b_{11}e_t \\ \Delta y_t(w,y) &= 0 \\ \mu_t(y) &= c_1 exp(c_2 t + c_3 w)y^2 \end{aligned} \tag{3}$$

where $e_t \sim N(0,1)$. In order to make w stationary conditional on y, we also need to specify the initial conditions:

$$w_0 | y \sim N(m_1, v_1)$$

and

$$y(0) \sim N(m_2, v_2)$$

where

$$m_1 = -a_{12}\frac{y}{a_{11}} \text{ and } v_1 = \frac{b_{11}^2}{1-(1+a_{11})^2} \quad .$$

The distributions have been specified as normal to take advantage of the distributional properties of the general model.

There are several features of the above relations which are worthy of note. First, a_{11} is assumed to be negative. This assumption is based on the phenomenon of "catch up growth" which is well recognized in the literature on human growth and nutrition (for example, Martorell et al., 1979). The idea is that children who fall below their appropriate weight for age due to some random environmental shock tend to experience more rapid growth in order to "catch up" with the appropriate time path of growth. If a_{11} is greater than -1, which it ought to be in practice,

then the equation for Δw conditional on y can then be written as a stationary AR(1) process with a mean m_1.

The dependence of m_1 on the frailty of individuals is incorporated in order to account for the fact that the biological tendency for catch up growth may be counteracted by continued adverse environmental conditions. Children exposed to high levels of infection may continue to lose ground with respect to a standard growth path.

The relation between the distribution of w_0 and frailty is a direct result of the assumption of stationarity of the process for each individual with a fixed frailty y. This assumption is not only helpful in practice, it is also plausible. As mentioned above in the discussion of Mata, children from an adverse environment have lower birth weights as well as less favorable growth.

The assumption that frailty is fixed for an individual also deserves some comment. While perhaps the most important reason for fixing frailty is that a model of varying frailty faces serious identification problems, the assumption does not seem unreasonable in the context being considered. It is unlikely that there will be substantial changes in the disease environment faced by a child over the course of a few years. Even when there are changes (as a result, perhaps of a development project on the one hand, or an epidemic on the other), it is likely that the changes will be shared by all the children in a village. If the ranking of frailty remains unaffected then the model is likely to still fit reasonably well.

One final note concerns the specification of the mortality function, which has been assumed fixed in each one month period. The mortality curve is thus not assumed to be a true Gompertz but a discrete approximation of the Gompertz. As long as the parameter c_2 is not large in absolute value this assumption will have little effect on the estimated parameters.

4. Monte Carlo Simulations

Before considering the results of the Monte Carlo simulations a short note on the approach that was used is in order. First, values were assigned to the various parameters of the model. Although the assumed values are not necessarily close to what would be observed if the model were fit to data, an attempt was made to chose coefficients that produced an observed pattern of weight dynamics and mortality that was at least representative of what is observed in practice. Second, each

simulated birth was assigned a random frailty drawn from a normal distribution. Using this value of frailty, an appropriate initial weight could be drawn for that individual conditional on his frailty. An iterative procedure was then used to simulate the series of weights and the death time of that child. At each time a probability of death was calculated conditional on t, y, and w_t. A random number was drawn from a uniform distribution to see if the individual survived the month. If he did, then a value of e_t was drawn from a normal distribution and weight in the $t+1$ period was calculated. Surviving individuals were censored at 30 months. The series of weights and the death time of each individual were then saved. The value of y was not saved since y represents information which would not be available in collected data that one wished to analyze.

4.1. Estimation

A maximum likelihood procedure was then used to estimate the parameters of the model. Since we wished to compare the results of a model incorporating unobserved heterogeneity and a model which ignored such heterogeneity, two different likelihood functions were maximized using the same simulated data.

The first likelihood function (Model A) can be described as the "true" model because with one small exception[2] it is based on the model used to generate the data. Consider an individual i who is observed for T_i months and then dies. The likelihood of observing the particular values that were simulated for each individual conditional on y is the following:

$$l_i \mid y = \varphi(w_0 \mid y) \prod_{s=1}^{T_i-1} [\varphi(w_{s+1} \mid w_s, y, s < T_i) \cdot exp(-\mu(s, y, w_s)] \times \mu(T_i, y, w_{T_i}) \quad . \tag{4}$$

[2]The mean and variance of the initial distribution of w were estimated separately, although the assumption of stationarity used in generating the data allows one to write these parameters in terms of other parameters in the model. The only real cost of this approach is that 2 degrees of freedom are lost. With real data, where one is unsure that the observed covariates are in fact stationary, one can compare the actual estimated values of these two parameter values with the values they should have based on the estimates of the other parameters. If the two estimates are inconsistent, then it may be advisable to try an alternative transformation of the time varying covariate.

If the individual was censored at time T_i, then the final multiplicand is omitted. Integrating over the possible values of y, we obtain:[3]

$$l_i = \int \varphi(y)\varphi(w_0|y) \prod_{s=1}^{T_i-1} [\varphi(w_{s+1}|w_s, y, s < T_i)\, exp(-\mu(s, y, w_s))]$$

$$\times \mu(T_i, y, w_{T_i})dy \quad .$$

Because each of the density functions is normal and because the mortality function is quadratic in y, the above equation can be written in the form:

$$l_i = \int y^2 \, exp(Ay^2 + By + C)dy$$

for which we can obtain an analytical expression. Taking the log of the unconditional likelihood and allowing for the possibility of censoring we obtain:

$$\ln(l_i) = C - \frac{B^2}{4A}\frac{1}{2} ln(-A)$$

$$+ g_i[ln(B^2 - 2A) - 2\, ln(-A) - ln(4)] \qquad (5)$$

where

$$A = \frac{1}{2}\left(\frac{m_1^2}{v_1} + \frac{1}{v_2} + T_i\frac{a_{12}}{b_{11}^2}\right) - \sum_{s=1}^{T_i-1} c_1\, exp(c_2 s + c_3 w_s)$$

$$B = w_0\frac{m_1}{v_1} + \frac{m_2}{v_2} + a_{12}\sum_{s=1}^{T_i-1}\frac{(w_{s+1} - (1 + a_{11})w_s)}{b_{11}^2}$$

$$C = \frac{1}{2}\left[\frac{m_2^2}{v_2} + \frac{w_0^2}{v_1} + \sum_{s=1}^{T_i-1}\frac{(w_{s+1} - (1 + a_{11})w_s)^2}{b_{11}^2}\right.$$

$$\left. + (T_i + 2)ln(2\pi) + T_i \ln(b_{11}^2) + \ln(v_2) + \ln(v_1)\right]$$

$$+ g_i[ln(c_1) + c_2 T_i + c_3 w_{T_i}]$$

and g_i takes the value zero if an individual is censored, one otherwise.

---

[3]Note that the probability of a death between time T_i and time T_{i+1}, is actually $1 - exp(-\mu(T_i, y, w_{T_i}))$. The expression used here is a good approximation, however, for monthly mortality rates because the monthly probability of death is quite small. If one uses the exact expression then the likelihood becomes more complicated but remains tractable.

The second model (Model B) to be considered ignores unobserved heterogeneity but is otherwise the same as Model A.[4] Ignoring heterogeneity amounts to assuming that the distribution of frailty has mean 1 and variance 0. The resulting likelihood function is a bit simpler:

$$ln(l_i) = -\frac{1}{2}\frac{(w_0 - m_1)^2}{v_1} - \sum_{s=1}^{T_i-1}\left\{c_1 \exp(c_2 s + c_3 w_s)\right.$$

$$-\frac{1}{2}[ln(v_1) - T_i \ln(b_{11}^2) - (T_i+1)ln(2\pi)$$

$$\left. + \frac{(w_{s+1} - (1+a_{11})w_s)^2}{b_{11}^2}]\right\}$$

$$+ g_i[ln(c_1) + c_2 T_i + c_3 w_{T_i}]$$

where again g_i is an indicator of whether or not the individual was censored. As Manton and Woodbury (1985) have pointed out, this type of model is separable into two parts. As a result, it is possible to estimate the weight dynamics equation independently of the mortality equation. Since we wished to compare the results of the two estimation processes, however, it seemed desirable to ignore this separability by estimating the two parts of the model simultaneously.

4.2. Basic Results

Tables 1a and 1b present the results of the Monte Carlo simulations for the two models along with the starting values that were used to generate the data. Each model was fit to 45 simulated data sets with 100 observations in each data set. Several aspects of the tables deserve comment. First, Model A, gives very reasonable parameter estimates. Not only are the parameter estimates within a 95% confidence interval of their true values, but the variances estimated from the inverse information matrix seem to be reasonably good estimates of the variances of the parameter estimates observed in the sample of 45. This result is quite expected. Since Model A is based on the true model, we expect the estimates to be both con-

[4]We are in the process of experimenting with an intermediate model in which unobserved heterogeneity is incorporated but assumed to be independent of the observed covariates. Some speculation is provided in the final section about what we expect to learn from these experiments.

sistent and efficient.

It is equally clear that Model B gives biased estimates to the parameters of interest. The parameter relating weight to mortality (c_3) is about twice its true value. In addition, the parameter estimates in the dynamic portion of the model are biased. The estimated value of a_{11}, for example, suggests that an individual who falls below the standard weight due to a random event will be slower to return than is the case in the true model (e.g. "catch up growth" has been underestimated).

It appears, then, that if one really can believe that the specified model captures the essence of the process operating in nature, then Model A is superior to Model B. Since in practice it is quite unlikely that the estimated model will accurately capture the process which generated the data, it is important to test the robustness of the conclusion that Model A is superior.

5. Robustness to Misspecification of Unobserved Distribution

There are reasons one might question any one of the assumptions incorporated into the model thus far. Nonetheless, one of the assumption seems particularly heroic: that the underlying frailty is distributed normally at the time of birth. Moreover, this assumption is perhaps the most difficult to test. If one suspects that the process of weight growth is not Markovian, then one can rewrite the likelihood function to incorporate an AR(2) process. Since the AR(1) and AR(2) processes are nested, one may use a Chi-Square test to test the hypothesis that the AR(1) model is appropriate. If one wishes to test alternative distributional assumptions, it is rarely possible to construct nested models.[5] It may be difficult to even fit models incorporating other distributional function since numerical integration of the conditional likelihood function over the possible values of frailty may be necessary.

In order to test the robustness of these estimation procedures to misspecifications of the underlying distribution of frailty, we generated data using a two point distribution and then fitted the data assuming a normal distribution. By fixing the mean and variance of the two point distribution and changing the probability associated with the first of the two points we could at least get a sense for the robustness of parameter estimates to changing distributional assumptions.

[5] With discrete point distributions, such as those of Heckman and Singer (1982), nested models are possible.

Table 1. Results from repeated simulations of model (100 individuals, 45 simulations).

Table 1a: Model A

Parameter	Input value	Average estimate	Sample variance	Average asymptotic variance
a_{12}	-0.5	-0.519	5.60E-03	7.80E-03
a_{11}	-0.5	-0.504	9.24E-04	6.30E-04
b_{11}^2	0.25	0.249	4.99E-05	8.42E-05
c_1	0.03	0.032	1.41E-04	2.01E-04
c_2	-0.01	-5.7E-03	2.11E-04	2.56E-04
c_3	-0.5	-0.555	5.94E-02	4.03E-02
m_1	-1.0	-1.030	2.16E-02	2.49E-02
v_1	0.33	0.325	1.96E-03	2.68E-03
m_2	1.0	0.975	1.55E-02	2.11E-02
v_2	0.1	0.093	8.32E-04	4.75E-04

Table 1b: Model B

Parameter	Input value	Average estimate	Sample variance	Average asymptotic variance
a_{12}	-0.5	-0.351	6.36E-04	4.81E-04
a_{11}	-0.5	-0.400	3.96E-04	4.18E-04
b_{11}^2	0.25	0.266	5.42E-05	9.10E-05
c_1	0.03	0.021	5.88E-05	4.60E-05
c_2	-0.01	-0.020	2.20E-04	2.48E-04
c_3	-0.5	-0.945	5.23E-02	3.64E-02
m_1	-1.0	-0.988	2.99E-03	4.24E-03
v_1	0.33	0.424	2.09E-03	3.52E-03

It is evident from Table 2 that the conclusion that Model A gives more accurate parameter estimates is retained even when the underlying distribution is misspecified. It also seems that the parameter estimates are biased, especially for a two point distribution with a high probability associated with the first point (and thus a strong negative skewness). Nonetheless, given the estimated variances that are obtained for sample sizes of 250, it is generally not possible to reject the hypothesis that the actual value of the parameter is -0.5 even with the most extreme distribution tested. In any case, the Model B parameter estimates of c_3 are still severely biased in a negative direction.

Of course the two point distribution is not representative of all possible distributions of underlying frailty. One cannot conclude on the basis of these results that the estimates will be equally robust, for example, when confronted with data generated by extreme value distributions of frailty. Nonetheless, the results obtained thus far are encouraging.

6. Prediction Equations

Up until this point it has been assumed that a model that produces biased parameter estimates is always inferior to a model that produces unbiased parameter values. In practice this does not always follow. First, although it may be the case that certain of the parameters are biased, it is also possible that some combination of those parameters will be well estimated by a simpler model. For example, the ratio $\frac{\alpha_{12}}{\alpha_{11}}$ is accurately estimated by Model B despite the fact that the estimates of α_{12} and α_{11} are both biased towards zero. Second, as was mentioned earlier, it is possible that the biased parameter estimates will provide a policy maker with more accurate information than the unbiased ones.

A simple example of this second idea can be constructed. Suppose there are two groups of people with mortality curves that are related as follows. For the first group $\mu(t) = \mu\, exp(y)$ and for the second group $\mu(t) = \mu\, exp(A + y)$. It is further assumed that at time 0 the distributions of y are the same in the two groups. However, due to the process of selection individuals in group A will have a different distribution of y at some later time than will those not in group A. As a result a model which ignores frailty will give an estimate of A which is biased toward zero as suggested by the Monte Carlo work of Ridder and Verbakel (1983). Now suppose we wish to know the mortality curve that will be experienced by an in-

Table 2. Effect of misspecified distribution of unobserved variables on estimate of c_3 (250 individuals, 12 simulations per probability).

Table 2a: Model A

Probability of point	1	0.2		0.4		0.6		0.8
Simulation	c_3	Variance	c_3	Variance	c_3	Variance	c_3	Variance
1	-0.45	0.02	-0.49	0.02	-0.75	0.02	-0.54	0.02
2	-0.79	0.02	-0.58	0.02	-0.38	0.02	-0.61	0.01
3	-0.20	0.01	-0.50	0.02	-0.68	0.01	-0.90	0.01
4	-0.46	0.02	-0.87	0.02	-0.60	0.01	-0.59	0.01
5	-0.55	0.02	-0.54	0.02	-0.46	0.01	-0.87	0.02
6	-0.56	0.02	-0.32	0.02	-0.90	0.02	-0.68	0.02
7	-0.62	0.02	-0.53	0.01	-0.59	0.02	-0.79	0.02
8	-0.20	0.02	-0.41	0.02	-0.57	0.02	-0.61	0.02
9	-0.35	0.01	-0.60	0.02	-0.48	0.02	-0.64	0.02
10	-0.61	0.01	-0.51	0.01	-0.60	0.01	-0.60	0.01
11	-0.52	0.02	-0.48	0.02	-0.60	0.02	-0.65	0.02
12	-0.40	0.02	-0.35	0.02	-0.83	0.02	-0.54	0.02
Mean:		-0.48		-0.52		-0.62		-0.67
Estimated Variance:		0.029		0.020		0.023		0.015

Table 2b: Model B

Probability of point	1	0.2		0.4		0.6		0.8
Simulation	c_3	Variance	c_3	Variance	c_3	Variance	c_3	Variance
1	-0.94	0.01	-0.93	0.01	-1.11	0.02	-0.79	0.02
2	-1.17	0.01	-1.08	0.01	-0.72	0.01	-0.69	0.01
3	-0.64	0.01	-1.05	0.01	-1.00	0.01	-1.59	0.01
4	-1.01	0.01	-1.25	0.01	-0.90	0.01	-0.92	0.01
5	-0.96	0.01	-1.08	0.01	-0.73	0.01	-1.15	0.01
6	-1.05	0.01	-0.74	0.01	-1.21	0.02	-0.97	0.02
7	-1.07	0.01	-0.94	0.02	-0.90	0.01	-1.00	0.02
8	-0.69	0.01	-0.82	0.01	-0.97	0.01	-0.88	0.02
9	-0.82	0.01	-0.99	0.01	-0.91	0.01	-0.97	0.02
10	-1.18	0.01	-0.95	0.01	-0.85	0.01	-0.89	0.01
11	-0.94	0.01	-0.90	0.01	-0.97	0.01	-1.05	0.01
12	-0.86	0.01	-0.81	0.01	-1.31	0.01	-0.81	0.01
Mean:		-0.94		-0.96		-0.97		-0.98
Estimated Variance:		0.029		0.020		0.020		0.053

dividual who moves into group A from the other group at time 0. Since the initial distribution of frailty in the two populations is the same, it is evident that the best estimate of his new mortality pattern will be the observed mortality pattern of the group he has joined. Since the biased estimate is based on the observed mortality it is likely to be a more accurate representation of the effect of moving into group A then the unbiased estimate. The unbiased parameter will give an unbiased estimate of the relative risk of members of the two populations only at time 0.

With this perspective in mind it seems essential to have a way of determining what the effect of a biased parameter estimate will be on the predicted result of some policy which serves to alter the value of one or more of the parameters in the model. One way that this may be done is to simply change a parameter and then generate a new set of data using the simulation procedure. By then comparing the observed mortality rates in the two cases we will be estimating the effect of some policy. In practice, however, this approach may be impractical since a large sample size is required to obtain estimates of the observed mortality which have small variance. (Figure 1, for example, illustrates the size of the variation which arises from simulated data with a sample size of 1000.) Moreover, in most cases it is difficult to determine the variance of the estimated mortality so that an appropriately large sample size may be selected in the first place.

An alternative approach is to use the equations describing the dynamics of the moments of the underlying distribution and the observed mortality in the context of the stochastic state space model. These equations for the multivariate case are derived in Woodbury and Manton (1977). Yashin, Manton, and Vaupel (1985) generalize the equations by including an observed and an unobserved variable in the model together. Yashin (1984) provides the mathematical underpinnings of the general model and also presents an elegant derivation of the multivariate prediction equations.

The basic results can be expressed as follows: if the initial distribution of unobserved variables is multivariate normal, if the equations of motion are linear, and if the mortality function is quadratic in the state variables, e.g.

$$d\boldsymbol{x} = (A_0(t) + A_1(t)\boldsymbol{x})dt + B(t)d\boldsymbol{\varepsilon}$$

$$\mu(\boldsymbol{x},t) = C_0(t) + C_1(t)\boldsymbol{x} + \boldsymbol{x}'C_2(t)\boldsymbol{x}$$

then the following equations are true:

$$\frac{d\,\mathbf{m}}{dt} = A_0(t) + A_1(t)\mathbf{m}(t) - \mathbf{V}(t)B_1 - 2\mathbf{V}(t)B_2(t)\mathbf{m}(t)) \qquad (7)$$

$$\frac{d\mathbf{V}}{dt} = A_1(t)\mathbf{V}(t) + \mathbf{V}(t)A_1(t)' + B(t)B(t)' - 2\mathbf{V}(t)C_2(t)\mathbf{V}(t) \tag{8}$$

$$\bar{\mu}(t) = C_0(t) + C_1(t)\mathbf{m}(t) + \mathbf{m}(t)'C_2(t)\mathbf{m}(t) + tr(C_2\mathbf{V}(t)) \tag{9}$$

where $\mathbf{m}(t)$ and $\mathbf{V(t)}$ are the vector of means and matrix of covariances, respectively, of the mortality determinants conditional on survival to time t.

If the true model met the conditions of this theorem completely then there would be no hesitancy in applying these equations in order to predict the effect on observed mortality of certain policy experiments. As may be remembered, however, the specification of the mortality curve was exponential rather than quadratic in w. As a result it is necessary to use an approximation of the actual mortality curve so that the above equations may be applied. The effect of making this simplifying assumption on results can be estimated through the use of simulations.

A Taylor expansion of $\mu(t,w,y)$ about the means m_1 and m_2 of w and y at time zero can be written as follows:

$$\begin{aligned}\mu(t,w,y) = c_1\, exp(c_2 t + c_3 m_1)(m_2^2 + c_3 m_2^2(w - m_1) \\ + 2m_2 y + \frac{1}{2}[(c_3 m_2(w - m_1))^2 \\ + 4m_2 c_3(w - m_1)(y - m_2) + 2(y - m_2)^2] \quad .\end{aligned} \tag{10}$$

If this equation is then used to generate a series of $\bar{\mu}$, then we will have at least an approximation of the expected observed mortality curve. In order to test the extent to which the approximation deviates from the true results we simulated 1000 data points using the "true" exponential model and the quadratic approximation of that model. In Figure 1, these two curves are plotted along with the values of $\bar{\mu}(t)$ obtained using the prediction equations for the quadratic model, (7), (8), and (9). The data generated from the quadratic mortality function should have as its expectation $\bar{\mu}(t)$. If the approximation is reasonable, then the curve generated from the exponential model should be similar to the curve generated from the quadratic model. It seems from Figure 1 that the approximation is acceptable.

Even if large differences were observed it would not be necessary to abandon the approach altogether, since even if the mortality curves tend to be quite different, the relative mortality arising from a change in one of the coefficients may be well estimated using the approximation. This issue needs to be investigated in greater detail by comparing simulated results with predicted results arising from the dynamic equations.

Figure 1

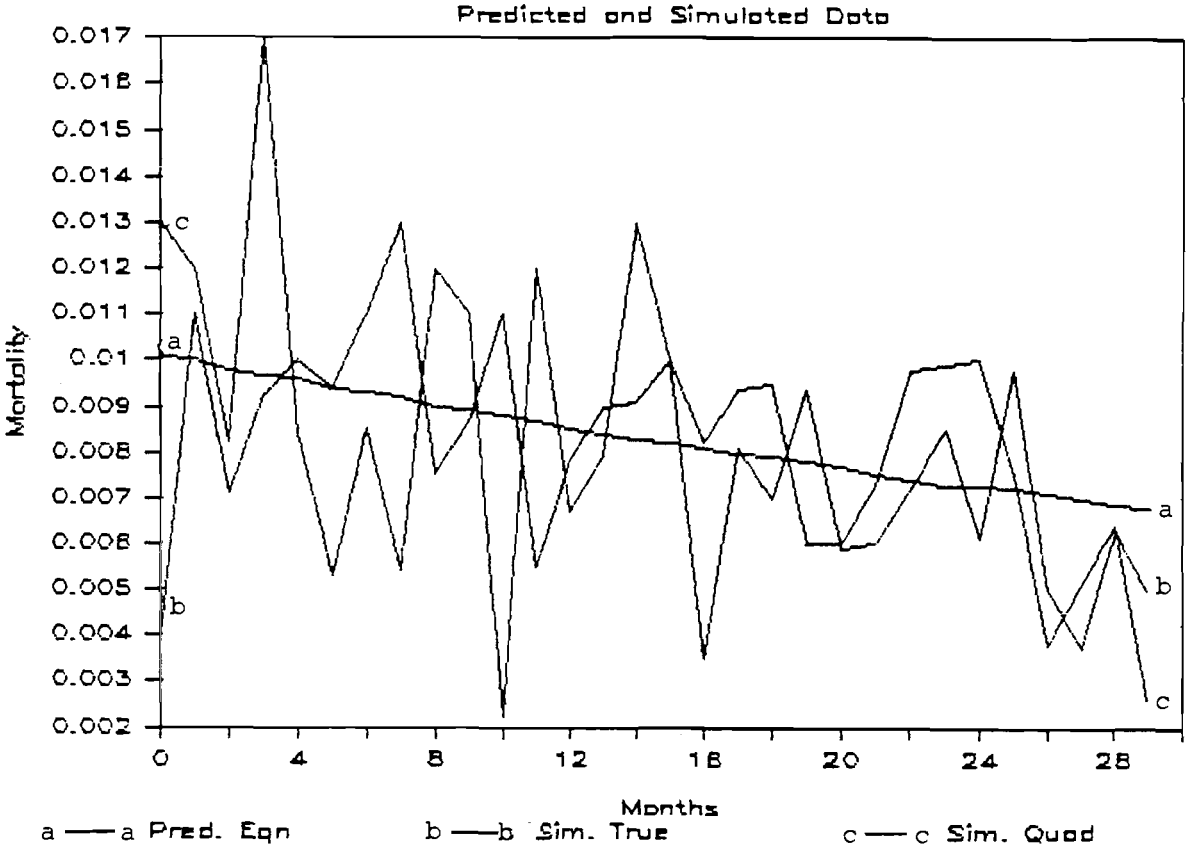

Table 3 provides a printout generated by the prediction equations when the true (e.g. input) parameters are known. The observed rising mean weight and falling mean frailty are certainly expected due to selection. The relatively small changes observed in the moments of the covariates suggests that the process of selection is having relatively little effect on parameter estimation. This result is not surprising since 70% of the individuals survive until their 30th month.

Nonetheless, the covariance term, which is negative and decreases in magnitude over time, provides insight into the likely result of a model which incorporates unobserved heterogeneity but assumes that the unobserved variables are uncorrelated with the observed variables (see footnote 4). Since the effect of selection is to make the two covariates less dependent on each other over time, and since the bias observed in Model B is a direct result of the covariance between them, we may postulate that a model of unobserved heterogeneity which controls

Table 3. Prediction estimates based on Model A.

Month	$\bar{\mu}_x$	l_x	Mean (w)	Variance (w)	Mean (y)	Variance (y)	Covariance (w,y)
0	0.0147	1.000	-1.000	0.333	1.000	0.100	-0.100
1	0.0144	0.985	-0.995	0.349	0.997	0.100	-0.099
2	0.0142	0.971	-0.991	0.349	0.994	0.099	-0.099
3	0.0140	0.958	-0.988	0.348	0.992	0.099	-0.099
4	0.0138	0.944	-0.985	0.348	0.989	0.099	-0.098
5	0.0135	0.931	-0.982	0.347	0.986	0.099	-0.098
6	0.0133	0.919	-0.980	0.347	0.984	0.098	-0.098
7	0.0131	0.907	-0.977	0.347	0.981	0.098	-0.097
8	0.0129	0.895	-0.975	0.347	0.979	0.098	-0.097
9	0.0127	0.883	-0.972	0.346	0.976	0.097	-0.097
10	0.0126	0.872	-0.970	0.346	0.974	0.097	-0.097
11	0.0124	0.861	-0.967	0.346	0.971	0.097	-0.096
12	0.0122	0.851	-0.965	0.346	0.969	0.097	-0.096
13	0.0120	0.840	-0.963	0.345	0.966	0.096	-0.096
14	0.0118	0.830	-0.960	0.345	0.964	0.096	-0.096
15	0.0116	0.821	-0.958	0.345	0.962	0.096	-0.095
16	0.0115	0.811	-0.956	0.345	0.960	0.096	-0.095
17	0.0113	0.802	-0.954	0.344	0.957	0.095	-0.095
18	0.0111	0.793	-0.952	0.344	0.955	0.095	-0.095
19	0.0110	0.784	-0.950	0.344	0.953	0.095	-0.094
20	0.0108	0.775	-0.948	0.344	0.951	0.095	-0.094
21	0.0107	0.767	-0.945	0.344	0.949	0.094	-0.094
22	0.0105	0.759	-0.943	0.343	0.947	0.094	-0.094
23	0.0104	0.751	-0.941	0.343	0.945	0.094	-0.094
24	0.0102	0.743	-0.940	0.343	0.943	0.094	-0.093
25	0.0101	0.736	-0.938	0.343	0.941	0.094	-0.093
26	0.0099	0.728	-0.936	0.342	0.939	0.093	-0.093
27	0.0098	0.721	-0.934	0.342	0.937	0.093	-0.093
28	0.0097	0.714	-0.932	0.342	0.935	0.093	-0.093
29	0.0095	0.707	-0.930	0.342	0.933	0.093	-0.092

only for the effect of selection would give parameter estimates with *more* bias than a model of infant growth and mortality which ignores unobserved variables altogether. If this hypothesis is supported by future work then there will be additional reason for giving careful consideration to possible relationships between unobserved variables and observed ones in the process of model estimation.

6.1. Prediction Equation Results

Table 4 represents, in some sense, the motivation for the entire discussion up to this point. In it we compare the effect of four policy alternatives on the estimated probability of surviving to the 30th month for the two different models. Estimates were generated by simulating a large data set (1000 children) and then fitting the two models (Table 5). Parameter estimates were then used in the dynamic

equations described above in order to produce estimated survival curves. The dynamic equations were similiar in the two cases, with the mean of y set to one and the variance to zero in Model B. For each policy we estimated the elasticity of the probability of dying by the 30th month by increasing the appropriate parameter by 10% and noting the percentage change in the estimated probability of dying that the prediction equations produced. Since A is the fitted version of the true model the values obtained from Model A are good estimates of the true impact of a given change in a parameter. Model B, then, gives poor estimates of the effect of specific polices on survival chances when results of Model B differ by a large amount from those of Model A.

The first policy considered involves raising the mean birth weight and assuming that all other parameters remain fixed. A program which provides nutritional suppiementation, prenatal care, or other benefits to pregnant women would produce this result and is in some ways relatively easy to administer (compared, for example, to making sure that infants are adequately fed). In any case, given a program of this sort there are two reasons that we might expect the estimates from Model B to be overly optimistic. First, since the effect of weight on mortality was overestimated in Model B, we will overestimate the effect of raising the weight at birth on mortality during the first month of life. Second, since Model B produces an estimate of α_{11} which is biased towards zero, suggesting a first order autoregressive process which adjusts to shocks relatively slowly, we will overestimate the mean weights that will be observed during the months immediately following birth. This effect increases the extent to which mortality declines will be overestimated in the model which ignores heterogeneity. Table 4 confirms the bias. While Model A suggests that the policy will have little effect, Model B predicts a small fall in mortality if birth weights rise. Specifically, Model B predicts that a 10% rise in mean weight at birth will lead to a 0.9% fall in the probability of dying by age 30 months. At the predicted level of mortality, this increases the number of surviving children per 1000 births by 2.

A second policy involves a decrease in the effect of weight on mortality. A practical example of a policy that might generate this result would be a program that provides medical care to low weight children. Again we expect Model B to overestimate the mortality declines that will be observed if such a policy is implemented because it assumes that the only factor affecting mortality is weight. Since in the true model mortality is also affected by an individual's frailty, a program targeting low weight children will not necessarily be reaching the ones with the

Table 4. Predicted changes in mortality.

	Model A		Model B	
	$1 - 1_{30}$	Estimated elasticity[1]	$1 - 1_{30}$	Estimated elasticity[1]
Base values	0.224	--	0.217	--
$m_1 := m_1 * 1.1$	0.224	0.00	0.219	0.09
$c_3 := c_3 * 1.1$	0.235	0.49	0.240	1.06
$a_{12} := a_{12} * 1.1$	0.234	0.46	0.233	0.74
$m_2 := m_2 * 1.1$	0.248	1.07	--	--

[1]Percent change in the proportion dying for a 1% change in the specified parameter. In each case mortality will rise if the parameter value rises.

Table 5. Large sample size estimates (1000 individuals).

		Model A		Model B	
Parameter	Input value	Parameter estimate	Variance estimate	Parameter estimate	Variance estimate
a_{12}	-0.5	-0.497	9.0E-04	-0.379	3.3E-05
a_{11}	-0.5	-0.507	3.7E-05	-0.396	2.5E-05
b_{11}^2	0.25	0.248	5.0E-06	0.267	5.5E-05
c_1	0.03	3.8E-03	3.2E-07	2.7E-03	2.4E-07
c_2	-0.01	0.010	6.4E-05	7.3E-03	6.4E-05
c_3	-0.5	-0.486	2.1E-03	-0.903	1.0E-03
m_1	-1.0	-0.998	3.7E-03	-1.008	4.0E-04
v_1	0.33	0.309	2.3E-04	0.398	3.1E-04
m_2	1.0	1.005	3.8E-02	--	--
v_2	0.1	0.099	1.0E-03	--	--

highest mortality. (They will, of course, be reaching some of the ones with high mortality because of the correlation between weight and frailty). Again, as expected, Table 4 indicates that Model B overestimates the effect of the specified policy.

A third possible policy involves a rise in the mean weight at all ages. An example might be a feeding program that reaches children of the appropriate ages. From the perspective of the model considered here this policy amounts to changing the value of α_{12}. Alternatively, one might incorporate a constant term into the equations of motion. As before we expect an overestimate of the effect on mortality from model B. In this case Model A predicts that a 10% increase in mean weight at all ages will increase the number of surviving children per 1000 by 10. Model B overestimates this increase by 6.

A final policy involves changing the disease environment or frailty, perhaps through building latrines, educating women, providing running water, or undertaking effective vaccination and oral rehydration campaigns. Since Model B ignores frailty, one cannot give a prediction based on Model B. It is clear from the results in Table 4, that this sort of policy is likely to have a large effect, at least if one accepts the parameter estimates used to generate the data as realistic. Estimation of parameters from real data will prove helpful in determining the relative mortality reductions that will be obtained under various policies.

One final note is necessary. In practice, any proposed policy cannot be described in terms of a *ceteris paribus* change in a single parameter. For example a policy designed to reduce the level of frailty in the environment may also reduce the variation of frailty observed in the population. While careful thought can be helpful in predicting what the results of a given program will be, it is certainly no substitute for estimation. If the model is repeatedly applied in different communities experiencing different sorts of intervention programs then some reasonable generalizations are likely to emerge about the effect on parameters of the model of certain kinds of programs.

7. Conclusion

In conclusions, it seems most appropriate to point to the additional work that needs to be done before the model discussed here can be adequately understood. First, the model begs to be applied; although the simulation results are certainly encouraging the real worth of the model can only be tested by fitting the model to available data. Second, more complicated specifications may be considered. The variance of the error term in the equations of motion in weight might, for example, be allowed to depend on the level of frailty. Third, the robustness of the results to other types of distributions than the two point distribution need to be tested. Robustness to other assumptions such as the linear and Markovian assumptions in the equations of motion must also be explored. Fourth, the robustness of the prediction equations to misspecification needs to be considered more carefully. If these equations prove to be reasonably robust then they can be powerful tools for exploring policy alternatives.

Despite the substantial amount of additional work that remains to be done, the significance of the results obtained so far deserve to be underscored. It seems clear from the above results, that it is possible to construct a model of child mortality which explicitly incorporates unobserved environmental frailty. Moreover, at least under certain assumptions, a model of this sort leads to parameter estimates which provide more accurate information to policy makers than do models which ignore unobserved heterogeneity.

REFERENCES

Bairagi, R. (1984) *Why mortality discriminating power of anthropometric indicators differs among populations.* Carolina Population Centers, University of North Carolina, Chapel Hill.

Bairagi, R., M.K. Chowdhury, Y. Kim, and G. Curlin (1984) *Alternative antrhopometric indicators of mortality: Is growth velocity a better indicator than attained weight*? Carolina Population Centers, University of North Carolina, Chapel Hill.

Chen, L.C., A.K.M.A. Chowdhury, and S.L. Huffman (1980) Anthropometric assessment of energy--protein malnutrition and subsequent risk of mortality among preschool age children. *American Journal of Clinical Nutrition* 33:1836-45.

Elbers, C. and G. Ridder (1982) True and spurious duration dependence: The identifiability of the proportional hazard model. *Review of Economic Studies* 49:403-9.

Heckman, J.J. and B. Singer (1982) Population heterogeneity in demographic models. *Multidimensional Mathematical Demography*. Academic Press.

Heckman, J.J. and B. Singer (1984) A method for minimizing the impact of distributional assumptions in econometric models for duration data. *Econometrica* (March) 52(2):271-320.

Hobcraft, J., J.W. McDonald, and S. Rutstein (1983) Socioeconomic factors in infant and child mortality: a cross-national comparison. *Population Studies* 38:193-233.

Hougaard, P. (1984) *Frailty models derived from the stable distribution.* Institute of Mathematical Statistics, University of Copenhagen. September 1984. Preprint #7.

Hougaard, P. (1982) *Life table methods for heterogeneous populations: distributions describing the heterogeneity.* Statistical Research Unit, Danish Medical Research Council. Research Report 82/5.

Manton, M.A. and M.A. Woodbury (1985) A multivariate Gaussian stochastic process model for describing both discrete and continuous state changes in longitudinal studies. *Sociological Methodology 1985*, pp. 277-315. Jossey-Bass: San Francisco.

Manton, K.G., E. Stallard, and J.W. Vaupel (1984) Alternative estimates of the heterogeneity of mortality risks among the aged. Submitted to *Journal of the American Statistical Association.*

Martorell, R. (1984) Nutritional status, morbidity and mortality. *Population and Development Review* Vol. 10: Supplement.

Martorell, R., C. Yarbrough, R.E. Klein, and A. Lechtig (1979) Malnutrition, body size and skeletal matruation: Interrelationships and implications for catch up growth. *Human Biology* 51:371-389.

Mata, L. (1985) Control of infection as a means to improve nutrition and reduce childhood mortality. *IUSSP Conference Proceedings* 4:39-50. Florence.

Mosely, W.H. (1985) Biological and Socio-Economic Determinants of Child Survival. A Proximate Determinants Framework Integrating Fertility and Mortality Variables. *IUSSP Conference Proceedings* 4:189-208. Florence.

Mosely, W.H. and L.C. Chen, eds. (1984a) Child survival: Strategies for research. *Population and Development Review* Vol. 10: Supplement.

Mosely, W.H. and L.C. Chen (1984b) An analytical framework for the study of child survival in developing countries. *Population and Development Review* Vol. 10: Supplement.

Ridder, G. and Wim. Verbakel (1983) On the estimation of the proportional hazards model in the presence of unobserved heterogeneity. University of Amsterdam, Faculty of Actuarial Science and Econometrics. Unpublished manuscript. December, 1983.

Trussell, J. and T. Richards (1984) Correcting for unmeasured heterogeneity in hazard models using the Heckman-Singer procedure. *Sociological Methodology 1985.*, pp. 242-276. Jossey-Bass: San Francisco.

Woodbury, M.A. and K.G. Manton (1977) A random-walk model of human mortality and aging. *Theoretical Population Biology* (February) 11(1):37-48.

Yashin, A.I. (1984) *Dynamics in Survival Analysis.* WP-84-107. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Yashin, A.I. and K.G. Manton (1984) *Evaluating the Effects of Observed and Unobserved Diffusion Processes in Survival Analysis of Longitudinal Data.* CP-84-58. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Yashin, A.I., K.G. Manton and J.W. Vaupel (1985) Mortality and aging in a heterogeneous population: A stochastic process model with observed and unobserved variables. *Theoretical Population Biology* (April) 27(2):154-175.