5-1-1981

Fertility, Child Schooling, and Exogenous Mortality: A Waiting Time Regression with Dynamic Regressors

Randall Olsen

Kenneth Wolpin

ECONOMIC GROWTH CENTER

YALE UNIVERSITY

Box 1987, Yale Station
New Haven, Connecticut

CENTER DISCUSSION PAPER NO. 378

FERTILITY, CHILD SCHOOLING, AND EXOGENOUS MORTALITY:

A WAITING TIME REGRESSION WITH DYNAMIC REGRESSORS

Randall J. Olsen and Kenneth I. Wolpin

May 1981

Introduction

The declining fertility and increasing levels of schooling that accompany economic development have been ascribed, in part, to the introduction of new medical inputs which have led to reductions in child mortality (Schultz (1971), O'Hara (1975)). Loosely speaking, these changing fertility and schooling patterns are hypothesized to have been "caused" by the exogenous mortality decline induced by the use of new medical technology. Recognizing that the longevity of offspring may be subject to choice, as is fertility and child schooling, implies joint adaptation of all three to exogenous influences. In this theoretical choice framework, the pattern of declining mortality, declining fertility, and increased schooling can result from exogenous changes in the provision of medical services only with a priori restrictions placed upon the preference structure (O'Hara, Rosenzweig and Wolpin (1980b)). There is no convincing evidence that fertility and schooling decisions are conformable with the hypothesis in a regime in which child survival responds to parental decisions.

In this paper we develop and implement an econometric methodology for estimating a family-specific exogenous component of life-expectancy which may be thought of as having analogous effects to medical technology. While waiting time models of the sort we estimate have a long tradition in statistics, economists are more familiar with models which seek to discover which characteristics of individuals and alternatives explain why an individual has made a particular choice when there is only a finite set of unordered alternatives. These studies are most often carried out using cross-sectional data which provide, at a particular point in time, attributes of choosers and the alternative choices they face as well as information on which alternative was selected.

When longitudinal data is available it is still possible to analyze the data as if it were cross sectional data. One would be free to pick an arbitrary time to study why some people are engaged in some activity and others aren't, but this strategy ignores the information available concerning how long people were engaged in that activity. After all, if the probability that someone is working is high, it seems reasonable to conclude that the same person would be in the state "working" for a long spell before shifting to the state "not working," or perhaps in the state "not working" for a short time before entering the state "working." Longitudinal data provides much more information than is contained in a simple cross-sectional survey of current behavior.

Markov models are most often used to analyze transitions among states. Discrete time Markov chain models assign some probability to the transition from state i to state j during some discrete interval of time. This probability is assumed only to depend upon the current value of i and not the path taken to arrive in i. The length of the time interval chosen is arbitrary. Moreover, such models cannot be used to predict transition probabilities over time intervals which are not integer multiples of the time unit chosen for the analysis (Singer and Spilerman [1976]). When the discrete interval of time is allowed to become infinitesimally small, a transition is made to a continuous time Markov model. The continuous time Markov model is closely related to the Poisson process. Poisson processes also assume that for a given interval Δt the probability of an event is $\alpha \Delta t$ with independence for events in nonoverlapping time periods. Under these conditions the distribution of the number of events in an interval is Poisson, and the **distribution of waiting time between events is exponential.**

While the simple exponential model provides a good place to begin the

analysis of waiting times, most applications are likely to require
extensions of the basic model. Many of the extensions to the exponential
waiting time model **have** antecedents in earlier work on extensions to the
Poisson distribution. For example, an early extension was to allow the
values of α for each individual to be drawn from a gamma density function
(see Greenwood and Yule [1920]). While their discussion was in terms of
distributions of events which departed from the Poisson, it should be clear
their approach also generates the gamma-exponential waiting time distribution.
Newbold [1927] took the idea a step further by inferring the distribution of
α from the observed distribution of events **using the assumption**
that given α the distribution of events is Poisson. A similar exercise
has been carried out with waiting time models in the more recent literature
starting with Silcock [1954].

While allowing for heterogeneity in α improved the fit of the model
to the data, early writers recognized that the departure of the distribution
of events from the Poisson did not imply heterogeneity. The alternative to
"accident **proneness**" models in which α varied across individuals was a
"contagion" model in which successive events are not independent. Workers
either learned from accidents, and subsequently had fewer, or accidents
degraded response time generating yet more accidents. In either case the
fundamental Poisson assumption of independence in non-overlapping intervals
is dropped. This indeterminacy between heterogeneity and time dependence
was dramatized when Eggenberger and Polya [1923, 1924] unwittingly arrived
at the same distribution of events as Greenwood and Yule except by pos-
tulating a contagion or time dependent process (see **Feller** [1943]). The
indeterminancy introduced by time dependence and heterogeneity has apparently

been recognized by actuaries and demographers in the study of life tables.
In pointing this out Hoem [1972] makes a citation to Higham [1851]. The
problem of identifying whether heterogeneity or time dependence in waiting
times is the true model has recently appeared in the economics literature.
See, for example Salant [1977], Heckman and Borjas [1979], Flinn and Heckman
[1980] and Coppock [1980].

One source of heterogeneity which is reasonably easy to identify in
waiting time models is that produced by differences in observable character-
istics across individuals. Once we control for differences in our exogenous
variables the exponential model may appear adequate (see Coleman [1964, 1973]
and Tuma and Groeneveld [1979]). This heterogeneity can be easily introduced
in the exponential waiting time model since the mean of waiting time is $1/\alpha$
so one can set $\alpha = 1/\theta x$ and regress t on x .

An additional extension to the usual waiting time model introduces a
dynamic change to the system while "waiting" is going on. For example, the
heterogeneity due to differences in individual characteristics may be dynamic
in the sense that the characteristics of the individual change through time. The
work of Flinn and Heckman introduces such time varying regressors into a
general waiting time model. They utilize the hazard function to model the
time specific rates of transition and then derive the density for waiting
time **from** the hazard function. Tuma and Groeneveld discuss the
introduction of parameters which take on different values during certain
discrete periods of time.

When applied work is done using waiting time models an often encountered problem is open intervals. Retrospective data provides information about when an individual was in a particular state, and the state currently occupied. The respondent does not usually know when he will depart his present state. For this last observed state the eventual waiting time to a change must be longer than the time spent in the state. Maximum likelihood solutions to this problem have been suggested by Salant, Tuma and Hannan [1978] and Flinn and Heckman. Heckman and Borjas suggest a sample selection correction to be used when one ignores the final open interval in a series of spells.

In the next section, we discuss a static utility maximization model which highlights the fertility and child investment decisions. We show that the effect of an exogenous component of child longevity on fertility and child schooling can only be determined with knowledge of compensated price effects and other utility function relationships. The following section develops a generalized waiting time regression model applied to life expectancy. Our approach diverges from usual waiting time models in that we use a non-exponential waiting time distribution, incorporating nevertheless many of extensions found in the more recent literature. We allow for family-specific heterogeneity in life-expectancy and for time-varying explanatory variables. The structural equation for longevity is more akin to a "production function" for life in the sense that we use choice variables, which may be thought of as inputs, as regressors. Given this, we treat heterogeneity as a fixed-effect which is permitted to be correlated with the regressors. The strategy is to retrieve the heterogeneity component from the production function, assuming of course the inclusion of all inputs,

and to estimate its impact on fertility and child schooling. The third section discusses the Malaysian data used in the estimation, the fourth section reports results, and the last section summarizes and discusses further work.

I. Theoretical Considerations

We adopt a very simple static lifetime utility maximization model to illustrate the relationships between child endowments of health and parental decisions concerning fertility and investments in child health and schooling. Prospective parents are assumed to have preference orderings over family size, the characteristics or "quality" of these children, and non-child related consumption goods. We consider two major components of child quality, survival or life expectancy, and schooling. The lifetime utility function is given by

(1) $U = U(N, L, S, Z)$ $U_i > 0, U_{ii} < 0$ $i = N, L, S, Z$

where N is the number of live births, L is the average length of life of children, S is the average schooling of children and Z is a composite consumption item.

Average child longevity is assumed to be produced with purchased inputs, X,

conditional on an endowment of health, V, which is unalterable and given

at birth to every child in the family. The relationship given by

(2) $L = L(X, V)$ $L_X, L_V > 0$

represents a production function for length of life. It is assumed that

V is known with certainty and that there either is no within family child-

specific stochastic variation in longevity, i.e., child longevity can be chosen

with certainty, or if such variation exists it is unknown to the family

and ignored in the maximization process. In addition, contraception is

assumed to be costless and perfect, although there is assumed to be a fixed

cost of bearing a live child. The budget constraint, with exogenous life-

time wealth, Y, is

(3) $Y = p_N N + p_X X + p_S S + p_Z Z$

where p_N is the price of a live birth, and p_X, p_S, p_Z are the respective

per unit prices of X, S, and Z.

In the context of this model, it is easy to see the difference between

exogenously altering longevity and exogenously altering the health endowment,

V . An experiment which randomly manipulated longevity in a population around

each family's equilibrium would alter the level of fertility and child school-

ing according to compensated substitution-complementarity relationships only.

This follows in a straightforward manner from conditional demand theory; in

particular, fertility will be inversely related to and child schooling direct-

ly related to exogenous changes in longevity if fertility and longevity are

Hicks-Slutsky substitutes, and schooling and longevity are complements. On

the other hand, population variation in V , even though randomly distributed,

does not imply an exogenous allocation of L as long as there are other mechan-

isms for altering L, i.e., as long as a production function like (2) exists.

To the extent that medical technology dominates other inputs in producing longevity, the experience during the economic development process would be more akin to the experiment of exogenously altering life expectancy. Estimation of compensated price effects would provide sufficient information to verify that the predicted movements in N and S would ensue from such an experiment, assuming the availability of appropriate prices. Confirmation of the theory, however, would also require direct evidence on the efficacy of health inputs.

The impact of V on N and S when longevity can be varied is obtained in the usual way by maximizing the utility function (1) given the production function (2) and the budget constraint (3). Totally differentiating the resulting first-order conditions yields[1]

$$(4) \quad \frac{dN}{dV} = \frac{1}{L_X} \left[L_V P_X \frac{dN}{dY} + \frac{1}{\lambda} (U_L L_V L_{XX} - U_L L_X L_{XV}) \left(\frac{dN}{dP_X} \right)_{U = \bar{U}} \right]$$

$$(5) \quad \frac{dS}{dV} = \frac{1}{L_X} \left[L_V P_X \frac{dS}{dY} + \frac{1}{\lambda} (U_L L_V L_{XX} - U_L L_X L_{XV}) \left(\frac{dS}{dP_X} \right)_{U = \bar{U}} \right]$$

where λ is the marginal utility of wealth, and where subscripts refer to partial derivatives.

The impact of V on fertility and schooling is composed of two components, an income effect and a compensated substitution effect. If N and S are normal goods then an increase in V will increase consumption of N and S as some of the real income gain induced in L is transferred to other consumption goods. In addition, the impact of V on N and S depends upon the substitution-complementarity relationships between L and N, and L and S, and on second-derivatives of the longevity production function. If we assume that $L_{XV} \simeq 0$ as we do in the estimation, then with $L_{XX} < 0$, the compensated price effect term will lead to a reduction in N and an increase in S if N and L are substitutes, and S and L are complements. Therefore, $\frac{dN}{dV}$ and $\frac{dS}{dV}$ cannot be signed without income and price effect information. The point is, however, that if medical inputs have effects analogous to V, the consistency of the hypothesis can be ascertained from (4) and (5) without estimates of their separate parts. Of course, if information on the use of medical inputs was available and if input usage was not subject to choice, this roundabout procedure would be unnecessary.

In the next section, we show how the production function given by (2) can be estimated using a waiting time regression framework allowing for the assumed correlation between V and X. The productive inputs are permitted to change through time, possibly in response to price variation through time, and stochastic individual child variation is introduced. The assumption of a static optimization model is rigidly maintained, however, in the sense that responses to child deaths in terms of subsequent input choices are ruled out. A dynamic adjustment model would require a different and more complex estimation approach.

II. A Generalized Waiting Time Regression Model

In this section we develop a regression approach to waiting time
models which incorporates many of the extensions to the exponential model.
The introduction provided a brief survey of some of the many approaches
taken in applying waiting time models to data. The chief advantage of the
method we suggest here is that it combines a number of these features in
a very simple regression framework.

Maximum likelihood estimators possess very desirable large sample properties.
Unfortunately they are often awkward to compute, requiring numerical methods
to obtain not just a local maximum, but the global maximum for functions which
are exceedingly nonlinear. While our method produces less efficient estimators,
its ease of application is an important advantage.

Our approach uses a nonexponential waiting time distribution, although
it could be amended to roughly approximate the exponential. We also
incorporate many of the other extensions to the basic exponential
waiting time model. First, we allow for heterogeneity which arises due to
differences in both observed and unobserved attributes. The unobservable
heterogeneity component is allowed to be correlated with the observable
heterogeneity component. Our approach is closer to fixed effects models than
random effects models. Whenever unobservable heterogeneity has been introduced
in waiting time models a random effects structure has been assumed.

Second, we allow for explanatory variables whose value changes through
time. We show how the effects of such variables can be summarized using
certain integrals. This is a useful result since it is not feasible to
use all the values of such variables as regressors, nor is it immediately
apparent how the information contained in such a variable can be summarized

for use in a regression equation. We also extend our method to allow for a generalized distributed lag process on these variables.

Third, we allow the coefficients of the exogenous variables to vary continuously through time.

Finally, the regression equation developed here allows for the presence of open intervals in the data so that data on incomplete spells can be combined with data on complete spells.

The traditional approach to waiting time models has been to parameterize the instantaneous rate of failure, the hazard function. This function, $\phi(t)$ is related to the probability density function (p.d.f.) $f(t)$ by

$$\phi(t) = f(t) / [1 - \int_0^t f(\tau) \, d\tau]$$

The p.d.f. can be expressed in terms of the hazard function using

$$f(t) = \phi(t) \exp [-\int_0^t \phi(\tau) \, d\tau].$$

Whether the analysis goes forward in terms of the pdf or hazard function is largely a question of convenience since the two are functionally related. An analogous situation is the choice of the time or frequency domain in time series analysis.

To facilitate the following analysis we assume the form of the p.d.f. for the duration of life for the ith child of the jth mother (parents) as

$$(6) \qquad f_{ij}(t) = X_{ij} B + \alpha Z_{ij}(t) + v_j$$

$$0 \leq t_{ij} \leq L_{ij}$$

Where L_{ij} satisfies $1 = \int_0^{L_{ij}} f_{ij}(\tau) \, d\tau$

with $f_{ij}(t) \geq 0$ $0 \leq t \leq L_{ij}$. [2/]

This formulation allows for attributes of the child, X_{ij}, which do not change over time, attributes $Z_{ij}(\tau)$ which do change over time, and a fixed effect, v_j, which represents random mortality components which are particular to the mother.

Censoring is a common feature of duration data. If we examine recent data on lifetimes many people will still be alive at the time the data are gathered. Let us assume there is censoring from above at $T_{ij} \leq L_{ij}$.[3]

In this event we have

$$(7) \qquad E(t_{ij} \mid t_{ij} \leq T_{ij}) = E(t_{ij} \mid t_{ij} < T_{ij}) \cdot Pr(t_{ij} \leq T_{ij}) +$$

$$T_{ij} \cdot Pr(t_{ij} \geq T_{ij}).$$

We can evaluate the mean of t_{ij}, subject to the right censoring, by using the density function in (6). To ease the notation let $\Theta_{ij} = X_{ij}B$ and temporarily drop the subscripts. We must also define the following expressions:

$$\overline{Z}_{\tau_1}^{\tau_2} = \frac{1}{\tau_2 - \tau_1} \int_{\tau_1}^{\tau_2} Z(t) \, dt$$

$$\overline{tZ}_{\tau_1}^{\tau_2} = \frac{1}{\tau_2 - \tau_1} \int_{\tau_1}^{\tau_2} t \, Z(t) \, dt$$

$$\overline{t^2 Z}_{\tau_1}^{\tau_2} = \frac{1}{\tau_2 - \tau_1} \int_{\tau_1}^{\tau_2} t^2 Z(t) \, dt$$

The first term on the right hand side of (7) is

$$T^2(\Theta + v)/2 + \alpha T \, \overline{tZ}_0^T$$

and the second term is

$$T[1 - T(\Theta + v) - \alpha T \, \overline{Z}_0^T].$$

Combining terms we obtain

$$(8) \qquad E(t \mid T, \Theta, v, Z(\tau)) = -(\Theta + v) T^2/2 - \alpha T (T \overline{Z}_0^T - \overline{tZ}_0^T) + T$$

Similarly,

$$(9) \qquad E(t^2 \mid T, \Theta, V, Z(\tau)) = -2(\Theta + v) T^3/3 - \alpha T (T^2 \overline{Z}_0^T - \overline{t^2 Z}_0^T) + T^2$$

clearly, $t = E(t \mid T, \Theta, v, Z(\tau)) + u$ where u is a random term with mean zero and finite variance.

Substituting from (8), rearranging terms, and adding the proper subscripts

$$(10) \quad \frac{2(t_{ij}-T_{ij})}{-T^2_{ij}} = X_{ij}B + 2\alpha(T_{ij} \overline{Z}_{ij0}^{T_{ij}} - \overline{tZ}_{ij0}^{T_{ij}})/T_{ij}$$

$$+v_j + \frac{2u_{ij}}{-T^2_{ij}} \quad .$$

We cannot estimate the regression in (10) by ordinary least squares (OLS) since v_j is unknown and presumably correlated with the X_{ij} or the time series for $Z(\tau)$.[4] However, we can use the usual within group estimator for fixed effects models to estimate B and α, see Mundlak [1978]. It is clear by inspection that u_{ij} is heteroscedastic since

$$Var(u_{ij}) = E(t^2_{ij}|T_{ij}, X_{ij}, Z_{ij}(\tau), v_j, \alpha, B) -$$

$$- E(t_{ij}|T_{ij}, X_{ij}, Z_{ij}(\tau), v_j, \alpha, B)^2.$$

Once we estimate B and α the estimated residuals from (10) may be averaged within the family to obtain an estimate of v_j, and then $Var(u_{ij})$ can be estimated using (8) and (9).

The heteroscedasticity of the u_{ij} makes the first step within estimator inefficient, so we must weight using

$$\hat{W}_{ij} = \{4 \hat{Var} (u_{ij})/ T^4_{ij}\}^{-1/2}$$

These weights must be used to obtain the group means, and used again to weight the deviations from group means.[5] The covariance estimator

for α and B is consistent if we increase either the number of children per family or the number of families, provided there is more than one child per family. When the number of children per family becomes large the estimated $\hat{v}_j$ converges to v_j and the estimated value of $\widehat{Var}(u_{ij})$ approaches the true value and our weighted least squares estimator becomes the Aitken GLS estimator.[6]

Censoring and Exogeneity

Each observation must have an exogenously chosen upper limit T_{ij} which is feasible in the sense that $f_{ij}(t)$ is positive for $t \leq T_{ij}$ and the probability of living past T_{ij} is positive . If it appears that this condition does not hold for some observations a smaller value of T_{ij} can be chosen. In the empirical section we focus on the first 24 months of life, so T_{ij} is the lesser of 24 months or the age of the child had it survived to the date of the survey.

We assume that while the family may adjust its inputs to the production of child survival to take into account its value of v_j, it does not alter its time profile of $Z(t)$ in response to random child deaths unexplained by either the inputs or the v_j. A dynamic strategy of adjusting inputs to deaths will bias the results obtained with this method since the time varying explanatory variables will be functions of the dependent variable.

Time Varying Coefficients and Pure Time Effects

If it is suspected that the coefficients α and B themselves change through time this can be captured using the above method for time

varying regressors. For example, if it is believed that the coefficient for $Z(t)$ follows the pattern $\alpha_o + \alpha_1 t + \alpha_2 t^2$ then estimation is simply a matter of defining new time varying regressors $tZ(t)$ and $t^2 Z(t)$ and treating them as $Z(t)$ was treated above.

If there are pure time effects then one could define expressions such as t, t^2, t^3 to be used as time varying regressors. Alternatively, one could construct time varying indicator variables $i(t)$ which take on the value one when the child is in a particular time interval.

In order to identify the pure time effects on the p.d.f. it is necessary to have sample variation in T_{ij} since the ultimate regressors are integrals over t from zero to T_{ij}. If it is important to estimate pure time effects one may select values of T_{ij} to induce sample variation. The only restriction is that the values of T_{ij} chosen be independent of actual waiting times and that there be positive probability that $t > T_{ij}$. The coefficients of the nonvarying regressors X_{ij} could also be made to change over time simply by multiplying the expressions obtained for the pure time effects by X_{ij}.[7]

By allowing coefficients to change through time we have relaxed the implication of the basic model that the effects of the exogenous variables have the same additive effect at all times. If the ordinate of the probability density function is very different through time then one might suspect that a model which was more nearly multiplicative than additive would be more desirable. Such a change could be achieved by transforming our measure of time so that the ordinate of the empirical pdf is more nearly constant through time.[8]

Distributed Lags in Continuous Time

One of the advantages of the above method is that it permits time varying explanatory variables to be used in a waiting time regression. The use of certain integrals as summary statistics enables us to deal with the embarrassment of riches contained in panel data. If we suspect that the p.d.f. for waiting time at t depends not only on the values of $Z(\tau)$ at time t, but also upon values of $Z(\tau)$ prior to t, we again confront a plethora of data since we may wish to allow $Z(\tau)$ to be lagged by any continuous amount of time. To deal with this problem we will use a continuous time analog of the Almon polynomial method. That method is often used for distributed lags of variables measured over discrete intervals of time and employs terms of the form

$$\sum_{i=1}^{r} i^k \, Z(t-i) \ .$$

Since here our explanatory variables are measured continuously (or at least are subject to changes at irregularly spaced times) an alternative is to use

$$\int_0^t (t-\tau)^k \, Z(\tau) \, d\tau \ .$$

This transformation yields another time varying regressor which can be handled in the way described above. If we include a number of these terms in our regression for $k=1, \ldots s$, each with coefficient

γ_k , the effect of a change in Z at time τ on the p.d.f. at time t is

$$\sum_{k=1}^{s} \gamma_k (t-\tau)^k .$$

Since the variable $Z(\tau)$ usually changes only at irregular intervals, the calculation of the required integrals is not difficult.

The hazard function formulation of waiting time models usually specifies the hazard of time τ to depend upon $Z(\tau)$. When one examines the formula for the p.d.f. at time τ in terms of the hazard function, it is clear $f(\tau)$ depends upon an integral over $Z(t)$ for $t \leq \tau$. Intuitively, lagged values of $Z(t)$ influence the fraction of the population at risk at time $\tau \geq t$ but the hazard is usually postulated only to be a function of $Z(\tau)$. A more general formulation would allow the hazard at τ to depend upon a distributed lag in Z.

In the absence of any firm theoretical basis for choosing a particular hazard function it is largely a matter of convenience whether one parameterizes $f(\tau)$ or the hazard since one can always express one in terms of the other.

Truncation and Grouping Durations

It may be the investigator is reluctant to assume stability of the estimated coefficients across durations and does not wish to parameterize the shift in coefficients across time by making the coefficients explicit functions of time. In this case the data may be examined sequentially through time. One model may be fit for durations from zero to t_1, another model from t_1 to t_2, and so forth. Only those who survived to the end of the previous stage of the analysis would be included in the next stage. The estimated

p.d.f. at each stage would be conditional upon survival to that stage. The complete p.d.f. for waiting time (as well as the mean of waiting time) could be numerically computed using the results from the various stages.

Summary

The method developed here is a fairly general regression approach to waiting time models. Problems such as censoring, time varying regressors and their distributed lags, fixed effects (heterogeneity), time varying coefficients and time dependent hazards can be accommodated. There are some restrictions on the method, so the avoidance of maximum likelihood is not costless. First, the entire history of any time varying regressors must be exogenous in a statistical sense and known up to the point of right censoring. Second, the point of right censoring must be exogenously chosen and there must be positive probability the waiting time exceeds the right censoring point. Finally, information about the distribution of the residuals is not used so the method, while computationally convenient, is not fully efficient.

III. The Data

The density function for duration of life can be estimated given information about child-specific inputs within families containing at least two children. The permanent family component of child longevity, v_j, can be calculated if all inputs that produce survival and that are subject to parental choice are included among the time-invariant or time varying regressors. A complete family life history is, therefore, required.

The 1976 Malaysian Family Life Survey contains 1262 households consisting of at least one ever married woman under 50 years of age as of the survey data. Except for some slight oversampling of Indian families and of fishing communities, the households form a random sample of all of Peninsular Malaysia. The essential feature of the survey for our purpose is that it contains a retrospective life history of each woman to the earlier of age 15 or age at marriage. The survey contains detailed information about pregnancy outcomes, child mortality, employment and housing characteristics.

The variables used in our analysis are presented in Table 1 together with descriptive statistics. They are cross classified by two characteristics, whether the variables change through time and whether they pertain to the child or to the household. The sample chosen for the analysis consists only of Malay families with at least two live births and only of children for whom the male respondent in the survey corresponded to the father. The final sample consisted of 1938 children from 311 families.

Table 1

Descriptive Statistics

	Mean	Standard Deviation
Child Specific Time Invariant Variables		
Birth Interval: months to preceding birth if not first born, zero if first born	24.99	21.02
Sex: 1 = male, 2 = female	1.51	.50
AMB: age of mother at birth (months)	302.4	74.11
Mother over 40 at Birth: one if age of mother at birth 40 or more, zero otherwise	.01	.11
Birthweight	7.01	1.41
First Born: one if first born, zero otherwise	.21	.41
Second Born: one if second born, zero otherwise	.18	.39
Third Born: one if third born, zero otherwise	.15	.36
Months breastfed (max. 24 months)	13.55	8.56
Length of life " "	21.58	6.29
Potential life	22.86	4.17
Child Specific Time Varying Variables		
Breastfeeding: unity if breastfed at time t, zero otherwise	.58	.37
Undivided Attention of Mother: fraction of potential time spent by mother at home at time t, not working at home	.81	.20
Divided Attention of Mother: fraction of potential time spent by mother at home at time t, working at home	.02	.10
Father's Time at Home: fraction of potential time spent by father at home at time t	.63	.21
Number of live siblings, age under one year at time t	.08	.15
Number of live siblings at least one year old but less than five years old at time t	.82	.67
Number of live siblings at least five years old but less than 10 years old at time t	.88	.96

	Mean	Standard Deviation
Number of live siblings at least 10 years old but less than 15 years old at time t	.48	.81
Number of grandparents in household at time t	.44	.82
Number of relatives under 10 years old in household at time t	.22	.84
Number of relatives over 10 years old in household at time t	.36	.95
Child Care Help - Own Children: fraction of potential time at t	.01	.06
Child Care Help - Grandparents: fraction of potential time at t	.04	.11
Child Care Help - Other Relatives, Neighbors, Servants: fraction of potential time at t	.02	.09
Access to Piped Water: one if piped water is accessible, zero otherwise	.26	.19
Access to Toilet Facilities: one if toilet facilities are accessible, zero otherwise	.64	.23
Electricity in House: one if house has electricty at time t, zero otherwise	.25	.43
Number of sleeping rooms in house at time t	2.13	.98
Modern Walls: one if walls constructed of modern materials, zero otherwise	.16	.36

The child specific time invariant variables are sex, birth order, age of mother at birth, interval to the preceding birth, and birthweight. For approximately half of the children weight was reported only as a categorical variable; very small, small, average, large, and very large. Weights of 3,5,7,9, and 11 pounds were assigned in those cases and an approximate weight variable was created which takes on the value zero if real weight was reported and the approximate weight otherwise. These child endowment variables should reflect only their inherent biological relationship to longevity as long as all other parental inputs are observed. Weight at birth is, for instance, in part determined by pre-natal maternal health care although it is clearly an endowment of the child at birth.

The household or **family-specific** time-invariant variables consist of parental schooling levels and parental ages. Notice that together with age of mother at each birth, age of mother at the time of the survey fixes the cohort of each child. Of course, the effect of these variables cannot be disentangled from the fixed effect since they do not vary within families. However, they will enter in the second stage of our analysis.

The **child-specific** time-varying variables include a dichotomous breast-feeding variable, i.e., whether, at any time t during the 24 months of potential life we consider, that the child is being breastfed, and a set of variables which accounts for the number of other live siblings of different ages with whom household level inputs must be shared at each time t of the 24 month period.[9] This latter set is divided into four categories, the number of live siblings under a year old, the number between one and five, the number between five and ten, and the number between ten and fifteen. It should be

noted that the manner in which time varying variables enter the formulation

of expected duration, namely as integral expressions over the entire po-

tential life of the child, requires information about the dynamic

variables over potential life independent of actual duration of life. For

children who die (prior to 24 months), it is therefore necessary to know

what the value of the time-varying variable would have been had the child

not died. Clearly, if a child dies while being breastfed, the necessary

information is not available. Operationally, we have assigned the average

duration of breastfeeding of those children within the family who survived

for at least 24 months to those children who were not weaned prior to

death. This provides an unbiased estimate of "desired" breastfeeding since

we have assumed that random child deaths unexplained by the inputs or the

fixed effect are uncorrelated with the inputs, although desired breast-

feeding is not measured without error by this procedure.

The final category consists of family level variables that change

through time. Since children are born at different stages of the family

life cycle, each child will face a different profile of each household level

input even though at any time t each child faces the same level of the input.

As long as the only determinants of the within-family distribution of these

inputs at time t are fully captured by the age distribution of living

children at time t and possibly also by child endowments, or as long as

the inputs are "public" goods, the family level variable will capture the

child-specific input. These variables consist of family composition other than

siblings, non-parental child care time, parental home time, and housing and sanitation

characteristics as described in Table 1.

The date of birth and date of death were not always known precisely
for each child. Length of life was, nonetheless, computed
in fractional months for each child by assigning values for the missing
information based upon the assumption that the day and month of birth (and
death) followed uniform distributions. When birth and death information
were both missing, death dates were constructed conditional on the death date be-
ing subsequent to the birth data. In addition, length of life was
always constrained to be as least as large as the duration of breastfeeding.
Approximately 18% of all children who died had no missing information,
another 52% were missing either or both the day of birth and the day of
death, and 21% had no information other than year of birth or year of death.[10]

The population histogram for duration of life of Malay children
whose potential life is at least as great as 24 months is depicted in
Table 2. Almost 8% of all live born children die within the first 24
months. Over 1% die in the first 2 weeks and over 5% die within the first
six months. The density is basically flat after 12 months although the
number of deaths is so sparse that it may not be possible to accurately
ascertain the shape of the density after 12 months. We attempt to capture
the shape of the density by the inclusion of a first month dummy variable.
There are not enough deaths among children with potential lives under 24
months to estimate the shape of the density with more detail.

Table 2

Duration of Life Frequency Distribution in Months:
Children with Potential Life of 24 Months or More

Length of Life,	Frequency	Per Cent	Cum. Per Cent
0 - .49	25	1.136	1.136
.50 - .99	12	.546	1.682
1.00 - 1.49	10	.454	2.136
1.50 - 1.99	13	.591	2.727
2.00 - 2.49	10	.454	3.181
2.50 - 2.99	6	.274	3.455
3.00 - 3.49	7	.318	3.773
3.50 - 3.99	6	.274	4.047
4.00 - 4.49	8	.362	4.409
4.50 - 4.99	7	.318	4.727
5.00 - 5.49	10	.454	5.181
5.50 - 5.99	1	.045	5.226
6.00 - 6.49	1	.045	5.271
6.50 - 5.99	9	.365	5.636
7.00 - 7.49	4	.182	5.818
7.50 - 7.99	2	.091	5.909
8.00 - 8.49	4	.182	6.091
8.50 - 8.99	5	.227	6.318
9.00 - 9.49	2	.091	6.409
9.50 - 9.99	1	.045	6.454
10.00 -10.49	3	.137	6.591
10.50 -10.99	4	.182	6.773
11.00 -11.49	5	.227	7.000
11.50 -11.99	0	.000	7.000
12.00 -12.99	4	.182	7.182
13.00 -13.99	1	.045	7.227
14.00 -14.99	2	.091	7.318
15.00 -15.99	2	.091	7.409
16.00 -16.99	1	.045	7.454
17.00 -17.99	1	.045	7.500
18.00 -18.99	1	.045	7.545
19.00 -19.99	3	.137	7.682
20.00 -20.99	2	.091	7.773
21.00 -21.99	1	.045	7.818
22.00 -22.99	1	.045	7.864
23.00 -23.99	0	.000	7.864
24.00 -	2027	92.136	100.000

IV. Results

A. The Production Function Estimates

The estimated parameters of the density function are presented in Table 3 both with and without a fixed effect. The magnitude of the coefficients are interpreted as the effect of unit changes on the monthly probability of death; for the time-varying variables the unit change occurs at each instant of potential life. Table 4, column 1, converts the density parameters to unit effects on life expectancy, i.e., on mean duration.

Consider first the endowment variables. An extra pound of birthweight reduces the monthly probability of death by 1 1/2 tenths of one percent or increases expected duration of life (over the first 24 months) by almost 2 weeks. Using approximate weight yields a similar, though slightly smaller, effect. Females have a lower monthly probability of death and higher mean duration of over 3 weeks. First borns have greater likelihood of death in each month than do second borns, second borns a greater likelihood of death than third borns, and third borns a greater likelihood of death than higher order births. Notice that we control for sibling age composition and age of mother at birth; either higher parity children are inherently more resilient or parents gain experience in childrearing. Women who give birth at age 40 or more (given parity, etc.) have children with lower life expectancy. We experimented with four other age categories, under 25, 25-29, 30-34, and 35-39, but there was no descernible age gradient across these groups. Finally, the length of the interval from the previous birth has very little

Table 3

Probability Density Function for Waiting Time to Death--

First 24 Months of Life [1]
(t-statistics in parentheses)

	Fixed Effects	No Fixed Effects
Birth Weight	-.001517	-.000871
	(5.07)	(3.48)
Birth Weight (approx.)	.000577	.000250
	(3.18)	(2.47)
Sex	-.002762	-.001939
	(5.47)	(3.38)
Birth Interval	.000021	-.000038
	(.86)	(1.87)
First Born	.008960	-.003311
	(4.57)	(1.94)
Second Born	.003613	-.003465
	(3.19)	(2.63)
Third Born	.001971	-.000341
	(2.10)	(.32)
Mother over 40 at Birth	.009541	.003915
	(3.73)	(.99)
Number Live Siblings Under One	.033070	.039310
	(9.70)	(9.04)
Number Live Sibling One to Five	.004615	.000097
	(5.92)	(.13)
Number Live Siblings Five to Ten	.000086	-.001253
	(.19)	(2.27)
Number Live Siblings Ten to Fifteen	.000021	-.000438
	(.05)	(1.06)
Number Relatives Under Age Ten Present	.000177	-.001017
	(.31)	(2.67)
Number Relatives Over Age Ten Present	-.001106	.000543
	(2.09)	(1.54)
Number Grandparents Present	.000374	.000733
	(.50)	(1.48)
Child Care Help Own Children	.005713	-.001432
	(1.18)	(.34)
Child Care Help Grandparents	-.008363	-.002957
	(2.16)	(1.21)
Child Care Help Other	-.019120	-.007998
	(4.38)	(2.25)
Electricity	-.000772	-.005577
	(.60)	(6.61)
Access to Piped Water	.000824	.001180
	(.55)	(1.47)
Access to Toilet Facilities	.002043	.000388
	(1.12)	(1.36)
Number Sleeping Rooms	.000631	.000014
	(1.46)	(.05)
Modern Walls	.001087	.001524
	(.70)	(2.01)
Undivided Attention of Mother	-.006271	.001967
	(1.53)	(1.44)

Table 3 continued

Divided Attention of Mother	-.016060	-.002608
	(2.35)	(1.06)
Father's Time at Home	.001205	.001959
	(.35)	(1.48)
Breastfeeding	-.003328	.001408
	(2.45)	(1.43)
First Month (Dummy)	.015650	-.004241
	(1.01)	(.30)
Constant	–	.010990
		(3.69)

[1] The coefficients are the parameters of the probability density function given by equation (6). A positive (negative) sign implies an increase (decrease) in the instantaneous probability of death.

Table 4

Estimated Mean Duration Effects from Probability Density Function

and from Ordinary Least Squares - Fixed Effects

	Probability Density Function[1]	Ordinary Least Squares
Birthweight	.437	.401
Birthweight (approximate)	-.166	-.083*
Sex	.795	.776
Birth Interval	-.006*	.001*
First Born	-2.58	-4.60
Second Born	-1.04	-2.75
Third Born	-.568	-1.93
Mother over 40 at Birth	-2.75	-2.90
Number Live Siblings under One	-9.52	-6.30
Number Live Siblings One to Five	-1.33	-1.87
Number Live Siblings Five to Ten	-.025*	-.560
Number Live Siblings Ten to Fifteen	-.006*	-.263*
Number Relatives under Age Ten Present	-.051*	-.000*
Number Relatives over Age Ten Present	.319	.540
Number Grandparents Present	-.108*	-.280*
Child Care Help - Own Children	-1.65*	-3.36*
Child Care Help - Grandparents	2.41	3.92
Child Care Help - Others	5.51	3.13
Electricity	.222*	1.45
Access to Piped Water	-.237*	-.620*

Table 4 continued

Access to Toilet Facilities	-.588*	-1.68
Number of Sleeping Rooms	-.182*	-.210*
Modern Walls	-.313*	-.650*
Undivided Attention of Mother	1.81*	3.52*
Divided Attention of Mother	4.64	6.19
Father's Time at Home	-.347*	-1.05*
Breastfeeding	.033	.033*

$^1 \dfrac{\partial Et}{dx} = \dfrac{-T^2}{2} \times$ density function parameters; $T = 24$.

$\dfrac{\partial Et}{\partial LBF} = (LBF - PL) \times -.003328$; $LBF = 14$, $PL = 24$; $LBF =$ length of breast-feeding, $PL =$ potential life.

*indicates standard error that exceeds one-half coefficient value

effect on the monthly probability of death. But, lengthening that interval,

given parity and mother's age at birth, will generally alter the age composi-

tion of living siblings which is also included in the regression. Thus,

as in the case with the other endowment variables, interpretations must be

made with care. Indeed, since we do not have explicit information about

the distribution of inputs within the family, the endowment variables such

as sex and parity may reflect the allocation process unless that

process depends only upon child age.

The existence of siblings under one year old during the first 24 months

of life has a substantial negative impact on life expectancy. For example,

being separated by 18 months from the next birth as opposed to

a 24 month separation, i.e., number of live siblings under one equals .25

as opposed to zero, implies a reduction in life expectancy of 2.38 months.

Note, however, that the mean number of live siblings under one is only .08

and that the standard deviation is .15.[11] An increase in the number of

older live siblings also increases the monthly probability of death, parti-

cularly for siblings under five. An additional sibling between the ages

of one and five reduces life expectancy in the first 24 months by 1.33

months. Live siblings over five, however, have much smaller impact on survival.

Additional relatives in the household over age ten increase life-expectancy,

possibly because they free parental time for child care, while additional relatives
under ten and grandparents are detrimental to survival though the effects
are measured imprecisely. Child care help supplied by older siblings in-
creases mortality while child care help by grandparents or by others includ-
ing relatives and servants increase survival. Given that we do not have
information on exact parental child care time, it is not clear whether
these alternatives replace or add to parental child care time on net.
Moreover, within the range of the variables observed in the data, the
effects are not exceptionally large.

The housing and sanitation facilities are generally not important
determinants of child mortality. Indeed, only electricity has the anticipated
sign. Selectively eliminating subsets of these variables does not change results for
the others and eliminating the entire group has no influence on the other findings.

An increase in the fraction of each week spent at home by the mother,
regardless of whether some home time is spent in production of items for
market sale, reduces the monthly probability of death. Surprisingly,
divided attention has a larger and more precisely measured favorable impact.
To get some notion of magnitudes, a one standard deviation increase in un-
divided attention time (.20) increases mean duration of life by about 10 days while a
one standard deviation increase in divided attention time (.10) increases mean
duration by about 14 days. Father's home time, however, increases mor-
tality but the magnitude is not very great and the point estimate not precise.

Breastfeeding is estimated to reduce the monthly probability of death by three-tenths of one percent. Thus, a child who is not breastfed has a cumulative probability of death over the first 24 months of life that is 8% higher than a child breastfed for the entire 24 months.
Put differently, increasing the length of breastfeeding by one month at the average length of breastfeeding of about 14 months, increases the expected length of life by 1 day; breastfeeding the first month, however, increases mean duration by 2 1/2 days.

Finally, the dummy variable for the first month indicates a likelihood of death in the first month that is .0157 higher than in any of the next 23 months. The uniform probability of death is .0033 per month so that the greater risk in the first month is substantial. Unfortunately, a more detailed description of the density function is not possible with this data. Such an analysis requires a significant number of deaths of children whose potential life is under 24 months and we have only a handful of such observations. The same shortcoming prevents the estimation of duration interactions either with endowments or time-varying variables.

Table 3 also presents the results without fixed effects. Families with different inherent monthly probabilities of infant death are anticipated, according to the theory of the first section, to choose different life-cycle input paths. A comparison of the two columns in Table 3 reveal substantively altered point estimates for some of the variables. For example, birth order effects are essentially reversed, the effect of live siblings between one and five is annihilated and the effect of live

siblings five to ten reversed in sign. Reversals also occur for relatives
under ten, relatives over ten, undivided attention of mother and breast-
feeding. As a general rule, the non-fixed effect estimates are biased down
when the fixed effect estimate is positive and biased up when the
fixed effect estimate is negative. This suggests
that families with inherently less healthy children attempt to compensate
by purchasing more inputs.

The second column of Table 4 provides estimates obtained from the
naive regression of actual duration of life on the same determinants,
except that time-varying variables are set equal to the mean value over
potential life. In addition potential life and its square are added as
regressors. The essential qualitative results are preserved, but quanti-
tative results are greatly distorted, both in the case of static and dyna-
mic variables.

B. The Effect of Exogenous Mortality on Fertility and Child Schooling.

Since most of the women in the sample are young, fertility is for many
not complete and many children are still in school. We, therefore, adopt
a waiting time model for both of these decisions. With respect to fertility,
we consider the duration to first birth as a proxy for completed fertility.[12]
The upper truncation point (maximum potential duration) is set at 300 months,
i.e., we estimate the density function only up to that point. In the case
of schooling, we assume that all children begin at age six (we do not know
school entry ages) with the upper truncation point set at 6 years of schooling.

Children who have already left the household have no reported schooling informa-
tion.[13] Children who died prior to entering school, but who had a positive
potential level of schooling, are treated as follows. If all decisions are

made at the beginning of the lifetime and are unalterable as is assumed
in the static model of the first section, then children who died prior to
receiving formal schooling should be assigned a zero level of schooling.
Of course, even if this is correct, it is the schooling level of surviving
children that is of interest for testing the demographic transition hypothe-
sis. Therefore, we also perform the analysis excluding those children who died prior
to entering school. It would also be correct to exclude those same children if house-
holds adapted to the death of a child by altering their household level allocation of
resources to schooling. But, as already noted, such a dynamic response
may not be compatible with the procedure used to estimate the mortality
fixed effect.

The regressions in each case consist of the family level fixed effects
estimated from the waiting time regression shown in Table 3 and of all
other family level variables which might influence life-expectancy, namely
parental schooling levels and parental ages (cohort). It is easy to show
that if the estimated fixed effect is the true fixed effect plus the
effects of these other variables on duration of life, incorporating these
other variables in the fertility and schooling waiting time models permits
consistent estimation of the impact of the fixed effect on fertility and
schooling.[14] There is the additional problem that the estimated v_j measures
the true v_j with error because families have only a finite number of children.
However, the sampling variances of the v_j's appear to be very small
relative to the magnitude of the fixed effect. Consequently, correction
for measurement error makes little difference.

Results are reported in Table 5.[15] The first column reports the effect of altering the inherent or exogenous family component of the monthly probability of child death (v_j). The second column uses the same waiting time formulations but substitutes the actual mean family monthly probability of death, i.e., it assumes that mortality is exogenous in the sense that it is uninfluenced by input choices.

With respect to fertility, it is estimated that a family with a 1% per month higher probability of death of children within the first 24 months will have the first birth 2.9 months earlier (given a maximum age at first birth of 300 months). Since the average interval between births is slightly more than 30 months, the family would have .1 more children over its life-time if no other intervals were changed.

The schooling results imply that each live-born child in a family with a .01 higher monthly death probability would, on average, have .4 fewer years of schooling; each surviving child, however, would receive almost the identical average level of schooling. The schooling of surviving children, therefore, appears invariant to exogenous mortality, although total family resources devoted to schooling is lower in families with higher inherent mortality. It is possible that the mortality fixed effect calculated only over the first 24 months of life is not closely related to human capital investments that don't begin until age six.

Table 5

The Effect of Exogenous and Endogenous Mortality

on Fertility and Child Schooling[1]

(t-statistics in parentheses)

	Fixed Effect (v_j) (exogenous)	Mean Monthly Probability of Death (endogenous)
Waiting Time to First Birth	.0063 (1.64)	.0144 (1.20)
Duration of Schooling (all children)	2.254 (4.48)	14.87 (11.87)
Duration of Schooling (surviving children)	-.1858 (.55)	1.55 (1.22)

[1]See Table 3, note 1.

The results, therefore, imply that exogenous increases in mortality increase fertility but do not affect the schooling of surviving children. The former is consistent with the hypothesis that exogenous improvements in medical technology may, in part, account for the positive correlation between fertility and mortality observed in aggregate time-series observations for many countries. Increased schooling levels, however, may have other causes. Results are quite different if these models are estimated using the mean monthly probability of death which presumably includes preferences as well as endowments. Both the fertility and schooling responses are greatly magnified as compared to the influence of exogenous mortality. Families who "choose" to have higher mortality, also choose to have higher fertility and lower schooling, even of surviving children. The hypothesis concerning the relationship between medical technology and the demographic transition is much more strongly supported in the cross-section when preferences are not purged from the measure of mortality.

V. Conclusions

In this paper we have devised a regression method for estimating a family-specific endowment of child survival which incorporates time varying regressors in a static lifetime decision model. In estimating the production function for life expectancy we found important connections between child endowments, such as sex and parity, and survival. Parental choices about birth spacing, child care time and breastfeeding also were found to influence life expectancies. In the second stage of the analysis, we found that families with higher endowed mortality rates began childbearing earlier and presumably would ultimately have more children, but purchased equal schooling for their surviving children. These fertility results are consistent with the interpretation of the demographic transition which rests on exogenous medical improvements as a contributing cause.

Footnotes

[1]The compensated price effects are those obtained from a model which omits the production function (2), and thus reflects only properties of the preference function. If additional structure is imposed on the model such as the interactive prices of Becker and Lewis (1973), assumptions about the preference structure would still be required to obtain predictions from this model (see Rosenzweig and Wolpin 1980a). We, therefore, ignore this added complexity.

[2]This lower case v is opposite in sign to the upper case V in the previous section.

[3]It is important to notice T_{ij} is not estimated and is not a function of the unknown variables.

[4]Since T_{ij}^2 is a random variable, we also require for consistency that $E(u_{ij}/T_{ij}^2$, other regressors) $= 0$.

[5]This standard transformation, which eliminates the across group variation and leaves only within group variation is of the form

$$R_{ij}^* = R_{ij} - \bar{R}._j$$

where $\bar{R}._j$ is the sample mean of the R_{ij} for group j. Because we must weight each observation by W_{ij}, we must use

$$R_{ij}^* = w_{ij}(R_{ij} - \bar{R}._j^w)$$

where

$$\bar{R}._j^w = \sum_i w_{ij}^2 R_{ij} / \sum_i w_{ij}^2 .$$

[6]If fixed effects are not present the procedure is greatly simplified to simple two stage weighted least squares.

[7]Note that the behavior of the hazard function through time can always be inferred from the p.d.f.

[8]Transformations of time have been used before in waiting time models. For example, the pdf for the Weibull waiting time model is

$$f(t) = r\theta \; t^{r-1} \exp[-\theta t^r]$$

and if we measure time in τ units where $\tau = t^r$, we obtain the simple exponential model

$$f(\tau) = \theta \; \exp[-\theta \tau].$$

By transforming time a Weibul model can be converted to an exponential model where the "forgetfulness" property is restored.

[9]Some household level inputs may be more akin to public goods in which case the number of siblings is irrelevant.

[10]This created some difficulty in deciding upon whether a child died while being breastfed when actual dates of birth and death were missing. For children breastfed less than 3 months, a separate question was asked about the reason for weaning one of the answers to which was that the child had died. For those breastfed longer than three months, we followed the convention of assuming the child died prior to desired weaning only in cases where our estimated month of death occurred at the month of weaning. Deleting the 19 households for which this problem arose did not alter the results significantly.

[11]This effect is quite large and at least suggests the possibility that the random death of a child may induce a shortening of the interval to the next child. This may be particularly true when a child dies while being breastfed

since it is thought that breastfeeding provides natural contraception. Our assumption is that the family, even in this case, maintains its optimal lifetime plan, i.e., the next birth occurs at the same date as if the child had not died. Note that the fact that we consider only the first 24 months of life mitigates the problem if adaptation to a child's death influences decisions by a lag of more than 24 months. Had we chosen a potential life of, say, 6 months the replacement interpretation would not arise at all. To check robustness, we deleted the variable number of live siblings under one, and found no substantial differences in results.

[12] It is possible that they merely start earlier in order to have a longer period in which to replace children when they die. As already noted, this dynamic behavior may not be consistent with our estimation strategy for the fixed effect (v_j).

[13] Omitting these children may not be innocuous if schooling is related to parity, and if, as we postulate, the mortality fixed effect is related to age at first birth.

[14] Consider the general model

$$y = \alpha_1 v_j^* + \alpha_2 x + \varepsilon$$

Where y is some transformation of fertility or schooling, v_j^* is the true

family specific component of child mortality and x is a representative
family variable such as mother's schooling. We estimate from the produc-
tion function only v_j, where

$$v_j = v_j^* + \beta_1 x.$$

Thus, in the second stage, we estimate

$$y = \alpha_1 v_j + (\alpha_2 - \alpha_1 \beta_1) x + \epsilon,$$

from which α_1 is retrieved. Without knowledge of β_1, however, α_2 cannot
be estimated.

[15]There are 288 households and 1018 children used in the age at first birth
regression and the first schooling regression respectively. The same households
are used in both; the reduction in sample size from the original 311 households and 19
children occurred primarily because some households had no children old enough
to have begun school and because children not currently in the household had
no schooling information.

References

Becker, Gary and H. Greg Lewis,"On the Interaction Between Quantity and Quality of Children," JPE, March/April 1973.

Cantrelle, P. and H. Leridon, "Breast-feeding, Mortality in Childhood and Fertility in a Rural Zone of Senegal," Population Studies, vol. 25, no. 3, November 1971.

Coleman, James S., Introduction to Mathematical Sociology, Collier-MacMillan, London, 1964.

_____, The Mathematics of Collective Action, Aldine Publishing Company, Chicago, 1973.

Eggenberger,F.,"Die Wahrscheinlichkeit Sansteckung," Mitteilungen der Vereinigung Schweizerischer Versicherungs – Mathematiker, 1924.

_____, and G. Polya, "Über die Statistik Verketteter Vorgänge," Zeit-schrift für Angewandte Mathematik und Mechanik, vol. 1 1923.

Feller, W.,"On a General Class of 'Contagious Distributions'," Annals of Mathematical Statistics, 14, #4, December 1943.

Flinn, Christopher J. and James J. Heckman, "Models for the Analysis of Labor Force Dynamics," mimeographed, 1980.

Greenwood, Major and G. Udny Yule, "An Inquiry into the Nature of Frequency Distributions Representative of Multiple Happenings with Particular Reference to the Occurrence of Multiple Attacks of Disease or Repeated Accidents, Journal of the Royal Statistical Society, vol. 83, 1920.

Heckman, James J. and George J. Borjas, "Does Unemployment Cause Future Unemployment? Definitions, Questions and Answers from a Continuous Time Model of Heterogeneity and State Dependence," mimeographed, Chicago, September 1979 (forthcoming Economica).

Higham, John A., "On the Value of Selection as Exercised by the Policy Holder Against the Company ," The Assurance Magazine, vol. 1, No. 3, April 1851.

Hoem, Jan M., "In Homogeneous Semi-Markov Processes, Select Acturarial Tables, and Duration Dependence in Demography," in Population Dynamics, T.N.E. Greville (ed.), Academic Press, New York, 1972.

Mundlak, Yair, "On the Pooling of Time Series and Cross Section Data", Econometrica, 46, January 1978.

Newbold, Ethel M., "Practical Applications of the Statistics of Repeated Events, Particularly to Industrial Accidents," JRSS, vol. 90, 1927.

O'Hara, Donald J., "Microeconomic Aspects of the Demographic Transition," JPE, December 1975.

Rosenzweig, Mark R. and Kenneth I. Wolpin, "Testing the Quality-Quantity Fertility Model: Results from a Natural Experiment Using Twins," Econometrica, January 1980a.

Rosenzweig, Mark R. and Kenneth I. Wolpin, "Governmental Interventions and Household Behavior in a Developing Country: Anticipating the Unanticipated Consequences of Social Programs," Economic Growth Center Discussion Paper No. 354, July 1980b.

Salant, Stephen W., "Search Theory and Duration Data: A Theory of Sorts," QJE, February 1977, vol. 91, No. 1

Schultz, T. Paul, "An Economic Perspective on Population Growth," in Rapid Population Growth, edited by National Academy of Sciences, John Hopkins Press, 1971.

Silcock, H., "The Phenomenon of Labor Turnover," JRSS (A), 117, No. 4, 1954.

Singer, Burton and Seymour Spilerman, "Some Methodological Issues in the Analysis of Longitudinal Surveys," Annals of Econ. and Social Measurement, 1976, Fall, vol. 5, No. 4.

Tuma, Nancy Brandon, Michael T. Hannan and Lyle P. Groenveld, "Dynamic History of Even Histories," American Journal of Sociology, 84, No. 4, January 1979.

Tuma, Nancy Brandon and Michael T. Hannan, "Approaches to the Censoring Problem in Analysis of Event Histories," Sociological Methodology 1979, Karl F. Schuessler (ed.), Jossey-Bass, Inc., San Franciso, 1978.