10-1-1982

An Estimable Dynamic Stochastic Model of Fertility and Child Mortality

Kenneth Wolpin

ECONOMIC GROWTH CENTER

YALE UNIVERSITY

Box 1987, Yale Station
New Haven, Connecticut

CENTER DISCUSSION PAPER NO. 424

AN ESTIMABLE DYNAMIC STOCHASTIC MODEL OF FERTILITY AND CHILD MORTALITY

Kenneth I. Wolpin

October 1982

1. Introduction

This paper presents an estimable dynamic stochastic model of the fertility-child mortality behavioral interaction. The household is assumed to solve a discrete-time finite-horizon decision problem for a discrete fertility outcome subject to uncertain child survival. The model yields comparative dynamic and static implications about the timing and spacing of children in a world of imperfect foresight. A tractable estimation method is developed, which is intimately tied to the theoretical formulation, and is applied to data. Even though the behavioral model is extremely simple, its solution cannot be analytically represented. Estimation requires numerical solution of the dynamic programming model that forms the basis of the theory, and is based upon the integration of that algorithm into a maximum likelihood procedure. It is the computational burden of the dynamic programming algorithm that limits the complexity of the model. The estimation method, however, is general to any finite-horizon dynamic stochastic model that can be numerically solved. The overall approach is highly structured and in this sense follows the tradition of Sargent (for example, 1978) in macro time-series analysis or of Heckman and MaCurdy (1980) in the analysis of life-cycle labor supply.

Economic models of fertility have generally been restricted to static lifetime formulations (e.g., Willis 1973). Life cycle considerations of timing and spacing have essentially been ignored, with only a few exceptions (e.g., Heckman and Willis (1975), Hotz (1980), Moffitt (1981)). Estimation approaches have either not been structural, i.e., recovering parameters of the structural taste or technology relationships, or have made extensive accommodations in moving from the theory to the application.

Several of the more important features of the fertility process are incorporated into the model developed here, and by necessity many are ignored. I assume two life-cycle stages: a stage of finite length in which a woman is fecund followed by a stage (finite or infinite in length) of sterility. During the first stage, contraception is perfect so that child spacing is completely volitional. A woman is constrained to have at most one child per period during her fertile stage.[1] The choice is of a discrete nature, i.e., have a child or not, in a discrete time framework.[2] Child deaths occur randomly as an exogenous Bernoulli process which may be non-stationary. There is an exogenous fixed cost of a child birth, which may or may not vary over the mother's life cycle, and an exogenous child maintenance cost which, for tractability, is assumed to arise only in the first period of the life of each child.[3] Parents are assumed to obtain consumption value from the number of surviving children in each period, independent of their ages, and from a single composite good. There is no lending or borrowing and income is exogenous, but possibly imperfectly foreseen.

The lack of capital markets is crucial to the existence of birth spacing that does not result solely from random sources.[4] Given a time-independent fixed cost of a birth, and preferences independent of the age distribution of children, birth spacing is induced in this model by the shape of the family income profile. Even a smoothly changing (rising) life cycle income pattern can generate spacing as a resolution of the tension between the desire to have children early (given discounting and a finite fertile stage) and the economic incentive to have them when income is high.[5]

An additional feature of the model is that the fertility choice is assumed to be made in an environment in which there is significant child mortality. Microeconomic theories of fertility and child mortality are at the root of many explanations of the demographic transition. Most previous work concerning the relationship between child mortality and fertility has focused on the issue of replacement, primarily at an empirical level (Ben-Porath (1976), Olsen (1980), Schultz (1976)). The existence of a replacement effect rests upon the intuitive notion that a new child is a better substitute for a child that has died than is any other commodity. Put differently, the death of a child is a real income loss associated with a single commodity, children, and that loss will be spread out, if possible, among other consumption goods (Ben-Porath and Welch (1977)). The model presented here explicitly demonstrates the connection between the household income profile, the survival probability profile, the finite horizon, and the cost of bearing and maintaining children in determining replacement propensities at different stages in the life-cycle.[6] Statistical methods developed to assess the impact of child mortality on fertility have been uninformed by theory, and, thus do not identify any particular structural parameter of interest; the "replacement effect" is not a single parameter. The change in fertility (the entire profile) induced by an exogenous child death will depend upon the life-cycle stage of the parents, the existing stock of children, the mortality environment, and the family income profile.[7]

Direct replacement, i.e., births predicated on the realization of a death, is often contrasted to the alternative strategy of hoarding.

Families in low survival environments, it is argued, will carry an inventory of children in anticipation of future deaths (Schultz (1976), Ben-Porath (1976)). This strategy would seem most tenable where older children have non-negligible mortality rates since direct replacement may be more costly (perhaps, in a biological sense) at later maternal ages, or where desired fertility is non-negligible at older ages. The model discussed in this paper does not allow for deaths of older children, and thus, hoarding arises only due to the inability to replace children born at the very end of the fertile stage.

In the next section, I specify a discrete-time discrete-outcome dynamic model of fertility with exogenous stochastic child mortality and imperfect foresight.[8] Only a very simple framework is explored, although it is surprisingly flexible with respect to the life-cycle fertility paths that may be generated even when exogenous variables change smoothly over the life-cycle. Many extensions to more complex settings will be obvious, but are not necessarily tractable. Comparative static and dynamic results are derived for a three period model. Section 3 presents simulation results for a 20 period model to demonstrate the wide variety of fertility profiles that can be generated and to illustrate the sensitivity of those profiles to variation in structural parameters. Section 4 discusses the estimation strategy. Several alternative maximum likelihood estimators are presented reflecting different assumptions about the origin and character of the error which up to that point has not been introduced into the model. The approach is to numerically solve the dynamic programming problem to determine the critical value of a particular random preference parameter which would make a household indifferent

between having or not having a child. This leads to a probit-like solution in the case of normal errors.[9] Section 5 describes the data and section 6 presents estimates of the model under the assumption that women know the stochastic processes of the exogenous variables. The model is estimated using Malaysian survey data containing households with complete retrospective information on births, child deaths and husband's income. Section 7 provides a summary and a discussion of extensions to more complicated models.

2. A Discrete-Time Discrete-Outcome Dynamic Stochastic Model of Fertility and Exogenous Child Mortality

The household is assumed to maximize an intertemporably separable utility function over a finite horizon that has as arguments the number of surviving children and a single composite consumption good. There is no direct consumption value to spacing births. A period is defined to be that length of time within which a single birth may occur. Births may occur only during the fertile life-cycle stage, the length of which is known with certainty to be T periods. There are τ-T following periods of sterility. Contraception is perfect during the fertile stage. There is a fixed cost of bearing a live child whether or not the child survives the first period and there is a fixed cost of maintaining a live child for the first period of its life that is technologically determined. Child mortality and household income are exogenous and each follows a stochastic process known to the household.

The household's decision problem for any period t is described as follows:

$$(1) \qquad \max E_t \left\{ \sum_{\ell=t}^{\tau} \delta^{\ell-1} U(M_\ell, X_\ell \right.$$

with respect to the vector $(n_t$

where $n_t = 1$ indicates a birth at t and $n_t = 0$ no birth, X_t is the level of goods consumption at t, M_t is the stock of surviving children at time t governed by the law of motion

$$(2) \qquad \begin{aligned} M_t &= M_{t-1} + n_t - d_t; \quad t = 1,\ldots,T \\ M_t &= M_T \qquad\qquad\quad t = T+1,\ldots,\tau \\ M_o &= 0 , \end{aligned}$$

$d_t = 1$ indicates a death at t and $d_t = 0$ no death of a child born at t, T is the final period in which n_t is subject to control (and therefore for which X_t is subject to control as well), τ is the exogenously given end of life, δ is a discount factor, and E_t is the expectations operator conditional on information available at t. Period 1 is the first period in which the woman makes an independent decision about fertility. It is not necessarily the date of marriage, if marriage is itself a decision about fertility. Barring illegitimate births, the decision not to marry is viewed as a decision not to have children.

Child deaths are random with exogenous environmental determinants G. Thus, at time t

$$(3) \qquad \begin{aligned} d_t &= 0 \quad \text{iff} \quad G_t \gamma_1 + u_t \geq 0 \\ &= 1 \quad \text{iff} \quad G_t \gamma_1 + u_t < 0. \end{aligned}$$

The vector G_t may contain time varying variables, e.g., calendar time, to capture exogenous forces causing a decline (or rise) in infant mortality through time. All elements of G_t and γ_1 are in the information set as of period 1.

The latent error u_t has a known joint stationary distribution $g(u_1,\ldots,u_t)$ with $Eu_t = 0$ for all t. However, the conditional mean need not be zero, i.e., $E_t u_t \neq 0$.[10]

The budget constraint must be satisfied each period and is given by

$$(4) \qquad Y_t = X_t + b(n_t - d_t) + cn_t$$

with Y_t household income at t, c the fixed cost of a birth and b the maintenance cost of a child who does not die in the first period of life. Household income at time t has exogenous determinants H_t, e.g., age. It is assumed that women do no market work so that husband's income is identical to household income. The income relationship is described by

$$(5) \qquad Y_t = H_t\gamma_2 + v_t$$

The random component v_t has joint stationary distribution $f(v_1,\ldots,v_\tau)$ with $Ev_t = 0$.[11] All elements of H_t and γ_2 are in the information set as of period 1. The mortality and income random components need not be independently distributed of each other.

It is important to specify precisely the content of the information set before discussing the solution method. As of time t the household knows how many children have survived to period t (M_{t-1}). In addition, as has already been assumed, the systematic components of income and child mortality in the past, present, and future are known as of t ($H_1\gamma_2,\ldots,H_\tau\gamma_2$, $G_1\gamma_1,\ldots,G_T\gamma_1$). Past realizations of income are obviously known, but neither the current realization (v_t) nor future realizations ($v_{t+1},\ldots,v_\tau$) is assumed known at t. Similarly, past deaths are assumed known, but not the fate of the current (prospective at t) child or future children.

Substituting the budget constraint into (1) and delineating the two life-cycle stages yields the equivalent problem:

$$(6) \qquad \max_{t} E \left\{ \sum_{\ell=t}^{T} \delta^{\ell-1} U(M_\ell, Y_\ell - b(n_\ell - d_\ell) - cn_\ell) + \sum_{\ell=T+1}^{\tau} \delta^{\ell-1} U(M_T, Y_\ell) \right\}$$

The maximization problem may be solved utilizing Bellman's principle. There is only a single (discrete) control variable in each period given optimal behavior in the future contingent on realizations, n_t, and a single state variable M_{t-1}, the number of children surviving to t.[12] The solution is obtained by backwards recursion. For each period the fertility decision will be conditional on past decisions and on future optimal behavior.

It is not possible to solve for the decision rule analytically for any arbitrary period. However, the general solution can be illustrated and some results deduced by working backwards for two periods. Formalities are relegated to Appendix A. In addition, several simplifying assumptions are adopted in order to facilitate the presentation. Appendix A presents the more general case. Here it is assumed that the conditional (on prior information) distributions of the mortality and income errors are statistically independent and that each of the error processes is, in addition, independently identically distributed over the life cycle.

Consider period T, the last decision period. Given all past decisions and realizations, the household chooses n_T to maximize

$$(7) \qquad E_T \left\{ U(M_{T-1} + n_T - d_T,\ Y_T - b(n_T - d_T) - cn_T) + \sum_{t=T+1}^{\tau} \delta^{t-T} U(M_{T-1} + n_T - d_T,\ Y_t) \right\}$$

Define $P_T = \int_{-G_T\gamma_1}^{\infty} g(u_T)du_T$ to be the probability of survival of a child born at T. The expected lifetime utility at T conditional upon a child being born at T and conditional upon the available information as of T is

$$(8)\quad E_T(LU_T/n_T = 1) = P_T E_T \Big\{ U(M_{T-1} + 1, H_T\gamma_2 + v_T - b - c) + \sum_{t=T+1}^{\tau} \delta^{t-T} U(M_{T-1} + 1, H_t\gamma_2 + v_t) \Big\}$$

$$+ (1 - P_T)E_T \Big\{ U(M_{T-1}, H_T\gamma_2 + v_t - c) + \sum_{t=T+1}^{\tau} \delta^{t-T} U(M_{T-1}, H_t\gamma_2 + v_t) \Big\}$$

Similarly, the expected lifetime utility given no birth at T is

$$(9)\quad E_T(LU_T/n_T = 0) = E_T \Big\{ U(M_{T-1}, H_T\gamma_2 + v_T) + \sum_{t=T+1}^{\tau} \delta^{t-T} U(H_{T-1}, H_T\gamma_2 + v_t) \Big\}$$

Defining J_T to be the difference between (8) and (9) the decision to have a child at time T is governed by

$$(10)\quad n_T = 1 \quad \text{iff} \quad J_T = E_T(LU_T/n_T = 1) - E_T(LU_T/n_T = 0) \geq 0$$
$$= 0 \quad \text{iff} \quad J_T < 0$$

where J_T can be written as

$$(11)\quad J_T = P_T E_T \Big\{ U(M_{T-1} + 1, H_T\gamma_2 + v_T - b - c) - U(M_{T-1}, H_T\gamma_2 + v_T - c) + \sum_{t=T+1}^{\tau} \delta^{t-T} [U(M_{T-1} + 1, H_t\gamma_2 + v_t) - U(M_{T-1}, H_t\gamma_2 + v_t)] \Big\} + E_T \Big\{ U(M_{T-1}, H_T\gamma_2 + v_T - c) - U(M_{T-1}, H_T\gamma_2 + v_T) \Big\}$$

In order to obtain concrete, though only illustrative implications, it is convenient to impose strong contemporaneous separability on the utility function (this is not necessary in estimation). Together with concavity of the utility function in income and children, the latter in the sense that increments in utility decline with discrete additions to the stock of surviving children, the following comparative static statements are easily demonstrated:

$$(12)\quad \frac{\partial J_T}{\partial M_{T-1}} < 0 \qquad \frac{\partial J_T}{\partial b} < 0 \qquad \frac{\partial J_T}{\partial c} < 0$$

$$\frac{\partial J_T}{\partial H_T\gamma_2} > 0 \qquad \frac{\partial J_T}{\partial H_t\gamma_2} = 0 \quad t=T+1,\ldots,\tau \qquad \frac{\partial J_T}{\partial G_T\gamma_1} \gtreqless 0$$

The greater the number of surviving children as of period T, M_{T-1}, the smaller the relative payoff to another child at T. Thus, for example, the death of a child born at T-1 will increase the differential gain to a child at T, the replacement child. This does not imply certain replacement, only a larger differential payoff of a new child. An increase in the birth or maintenance cost also decreases the relative payoff, while an increase in an exogenous component of income at time T ($H_T\gamma_2$) increases the payoff. The payoff, however, is uneffected by anticipated future income, a result that is, in particular, special to the contemporaneous separability assumption.

Changes in the probability of survival have an ambiguous effect on the payoff; for example, if a birth was desired given certain survival and zero birth cost ($P_T = 1$ and $c = 0$), increasing the survival probability from some value less than unity would increase the differential payoff, while if under the same circumstances the child was not desired, increasing the survival probability would decrease the differential payoff.

The decision to have a child at T-1, conditional on information available at T-1, depends not only on the events and decisions of prior periods but also on the optimal future course viewed as of T-1. Given a child is born at T-1, expected lifetime utility is given by

$$
\begin{aligned}
(13)\quad & E_{T-1}(LU_{T-1} \mid n_{T-1} = 1) \\
&= P_{T-1}\, E_{T-1} U(M_{T-2} + 1, H_{T-1}\gamma_2 + v_{T-1} - b - c) \\
&+ (1 - P_{T-1})\, E_{T-1} U(M_{T-2}, H_{T-1}\gamma_2 + v_{T-1} - c) \\
&+ \delta \max \left\{ \begin{array}{l} P_{T-1} \left[E_{T-1} \left[E_T(LU_T \mid M_{T-1} = M_{T-2} + 1, n_T = 1) \right] \right], \\ P_{T-1} \left[E_{T-1} \left[E_T(LU_T \mid M_{T-1} = M_{T-2} + 1, n_T = 0) \right] \right] \end{array} \right. \\
&+ \delta \max \left\{ \begin{array}{l} (1 - P_{T-1}) \left[E_{T-1} \left[E_T(LU_T \mid M_{T-1} = M_{T-2}, n_T = 1) \right] \right], \\ (1 - P_{T-1})\, E_{T-1} \left[E_T(LU_T \mid M_{T-1} = M_{T-2}, n_T = 0) \right] \end{array} \right.
\end{aligned}
$$

where P_{T-1} is defined similarly to P_T and where $E_T(LU_T \mid M_{T-1} = M_{T-2} + 1, n_T = 1)$ is given by (8) with $M_{T-1} = M_{T-2} + 1$ and with similar definitions for

the other expressions in (13).[13] If a child is not born at T-1, expected lifetime utility is given by

$$(14)\quad E_{T-1}(LU_{T-1}|n_{T-1}=0) = E_{T-1}\Big\{U(M_{T-2},\ H_{T-1}\gamma_2 + v_{T-1})\Big\}$$
$$+ \delta \max\Big\{E_{T-1}\Big[E_T(LU_T|M_{T-1}=M_{T-2},n_T=1)\Big],$$
$$E_{T-1}\Big[E_T(LL_T|M_{T-1}=M_{T-2},n_T=0\Big\}$$

Denoting J_{T-1} as the difference in expected utilities with and without a birth at T-1, the decision to have a child at T-1 is fully described by

$$(15)\quad n_{T-1}=1 \quad \text{iff} \quad J_{T-1} = E_{T-1}(LU_{T-1}|n_{T-1}=1) - E_{T-1}(LU_{T-1}|n_{T-1}=0) \geq 0$$
$$n_{T-1}=0 \quad \text{iff} \quad J_{T-1} < 0$$

where J_{T-1} can be written as

$$(16)\quad J_{T-1} = P_{T-1}\,E_{T-1}\Big\{U(M_{T-2}+1,H_{T-1}\gamma_2 + v_{T-1} - b - c)$$
$$-U(M_{T-2},H_{T-1}\gamma_2 + v_{T-1} - c$$
$$+ E_{T-1}\ U(M_{T-2},H_{T-1}\gamma_2 + v_{T-1} - c) - U(M_{T-2},H_{T-1}\gamma_2 + v_{T-1})$$
$$+ \delta \max\ P_{T-1}\Big[E_{T-1}E_T(LU_T|M_{T-1}=M_{T-2}+1,n_T=1)\Big],$$
$$P_{T-1}\Big[E_{T-1}E_T(LU_T|M_{T-1}=M_{T-2}+1,n_T=0)\Big]$$

$$- \delta \max \Big\{ P_{T-1} \Big[E_{T-1} \big\{ E_T (LU_T | M_{T-1} = M_{T-2}, n_T = 1) \big\} \Big],$$

$$P_{T-1} \Big[E_{T-1} \big\{ E_T (LU_T | M_{T-1} = M_{T-2}, n_T = 0) \big\} \Big] \Big\},$$

In order to derive comparative static and dynamic results for the T-1 choice, it is necessary to consider each possible optimal future decision separately. There are three cases: it is optimal to have a child in period T under all contingencies, i.e., $n_T = 1$ is optimal both when $M_{T-1} = M_{T-2} + 1$ and when $M_{T-1} = M_{T-2}$; it is optimal to have a child at T only if a child was not born in T-1 or if born did not survive, i.e., $n_T = 1$ is optimal when $M_{T-1} = M_{T-2}$ but not when $M_{T-1} = M_{T-2} + 1$; and n_T is not optimal under either alternative value of M_{T-1}. Note that if $n_T = 1$ is optimal when $M_{T-1} = M_{T-2} + 1$, then it must also be optimal when $M_{T-1} = M_{T-2}$ since it was shown that the differential payoff of a child at $T(J_T)$ is inversely related to M_{T-1}. I will consider each case in turn.

It is convenient to let the first two terms in (16) be labelled D_{T-1} as it recurs in all three cases.

Case 1: $n_T = 1$ with either $M_{T-1} = M_{T-2} + 1$ or $M_{T-1} = M_{T-2}$.

The differential payoff to a birth, in this case, becomes

$$(17)\qquad J_{T-1} = D_{T-1} + \delta\Big\{P_{T-1}P_T\Big\{E_{T-1}[U(M_{T-2}+2,\ H_T\gamma_2 + v_T - b - c) - \sum_{t=T+1}^{\tau}\delta^{t-T}U(M_{T-2}+2,\ H_t\gamma_2 + v_t)]$$

$$-E_{T-1}[U(M_{T-2}+1,\ H_T\gamma_2 + v_T - b - c) - \sum_{t=T+1}^{\tau}\delta^{t-T}U(M_{T-2}+1,\ H_t\gamma_2 + v_t)]\Big\}$$

$$+ (P_{T-1})(1 - P_T)\Big\{E_{T-1}[U(M_{T-2}+1,\ H_T\gamma_2 + v_T - c) - \sum_{t=T+1}^{\tau}\delta^{t-T}U(M_{T-2}+1,\ H_t\gamma_2 + v_t)]$$

$$-E_{T-1}[U(M_{T-2},\ H_T\gamma_2 + v_T - c) - \sum_{t=T+1}^{\tau}\delta^{t-T}U(M_{T-2},\ H_t\gamma_2 + v_t)]\Big\}\Big\}$$

Under the assumption that U is strongly contemporaneously separable, it is easily shown that

$$(18)\qquad \frac{\partial J_{T-1}}{\partial M_{T-2}} < 0 \qquad \frac{\partial J_{T-1}}{\partial b} < 0 \qquad \frac{\partial J_{T-1}}{\partial c} < 0$$

$$\frac{\partial J_{T-1}}{\partial H_{T-1}\gamma_2} > 0 \qquad \frac{\partial J_{T-1}}{\partial H_t\,\gamma_2} = 0 \qquad t = T,\ldots,\tau$$

$$\frac{\partial J_{T-1}}{\partial G_T\,\gamma_1} < 0 \qquad \frac{\partial J_{T-1}}{\partial G_{T-1}\gamma_1} \gtreqless 0$$

The implications for n_{T-1}, given $n_T = 1$ is optimal, are similar to those for n_T. A birth is more "likely" in T-1 if there are fewer surviving children at T-1, if the cost of a birth and of maintenance are lower, and if income at T-1 is higher. As before, anticipated future income has no effect on the expected payoff and the probability of survival of the T-1 child has an ambiguous effect. However, if the probability of survival of the T period child (given that it is always optimal to have a T period birth) is higher, the payoff to a T-1 period birth is lower.

Case 2: $n_T = 1$ only if $M_{T-1} = M_{T-2}$

The expected payoff differential is now given by

$$
\begin{aligned}
(19)\quad J_{T-1} = D_{T-1} + \delta \Big\{ & P_{T-1} \Big\{ E_{T-1}[U(M_{T-2} + 1,\ H_T\gamma_2 + v_T) \\
& \qquad + \sum_{t=T+1}^{\tau} \delta^{t-T} U(M_{T-2} + 1, H_t\gamma_2 + v_t)] \Big\} \\
& - P_{T-1}P_T \Big\{ E_{T-1}[U(M_{T-2} + 1,\ H_T\gamma_2 + v_T - b - c) \\
& \qquad + \sum_{t=T+1}^{\tau} \delta^{t-T} U(M_{T-2} + 1,\ H_t\gamma_2 + v_t)] \Big\} \\
& - P_{T-1}(1 - P_T) \Big\{ E_{T-1}[U(M_{T-2},\ H_T\gamma_2 + v_T - c) \\
& \qquad + \sum_{t=T+1}^{\tau} \delta^{t-T} U(M_{T-2}, H_t\gamma_2 + v_t)] \Big\} \Big\}
\end{aligned}
$$

from which the following may be derived under separability:

$$(20)\qquad \frac{\partial J_{T-1}}{\partial M_{T-2}} < 0 \qquad \frac{\partial J_{T-1}}{\partial b} \gtreqless 0 \qquad \frac{\partial J_{T-1}}{\partial c} \lesseqgtr 0,$$

$$\frac{\partial J_{T-1}}{\partial H_{T-1}\,\gamma_2} > 0 \qquad \frac{\partial J_{T-1}}{\partial H_T\,\gamma_2} < 0 \qquad \frac{\partial J_{T-1}}{\partial H_t\,\gamma_2} = 0 \qquad t = T+1,\ldots,\tau.$$

$$\frac{\partial J_{T-1}}{\partial G_T\,\gamma_1} \lesseqgtr 0 \qquad \frac{\partial J_{T-1}}{\partial G_{T-1}\,\gamma_1} \lesseqgtr 0$$

These results are substantially different from those in case 1. The replacement tendency (M_{T-2}) is still apparent as is the positive T-1 income response ($H_{T-1}\gamma_2$). However, income in period T is inversely related to the T-1 payoff. This results from the fact that as income in period T rises, the expected payoff from a birth in period T also rises. Since, in this case, a T period birth is optimal only if a T-1 period birth is not, the rise in T period income will reduce the payoff from a T-1 period birth. A similar intertemporal interaction explains the ambiguity concerning the effects of birth and maintenance costs. The direct effect of an increase in b or c is to reduce the payoff in T-1. But, the indirect effect is to reduce the payoff in T as well, which increases the payoff of a T-1 birth given that a T period birth is desired only when a T-1 period child does not exist. Analogous reasoning applies to the rest

of the results and, therefore, need not be detailed.

Case 3: $n_T = 0$ when $M_{T-1} = M_{T-2}$ or $M_{T-1} = M_{T-2} + 1$

In case 3, the expected payoff is

$$(21)\quad J_{T-1} = D_{T-1} + P_{T-1}\Big\{ E_{T-1}[U(M_{T-2}+1,\ H_T\gamma_2 + v_T) + \sum_{t=T+1}^{\tau} \delta^{t-T} U(M_{T-2}+1,\ H_t\gamma_2 + v_t) - U(M_{T-2},\ H_T\gamma_2 + v_t) - \sum_{t=T+1}^{\tau} \delta^{t-T} U(M_{T-2},\ H_t\gamma_2 + v_t)]\Big\}$$

with comparative static results given by

$$(22)\quad \frac{\partial J_{T-1}}{\partial M_{T-2}} < 0 \qquad \frac{\partial J_{T-1}}{\partial b} < 0 \qquad \frac{\partial J_{T-1}}{\partial c} < 0$$

$$\frac{\partial J_{T-1}}{\partial H_{T-1}\gamma_2} > 0 \qquad \frac{\partial J_{T-1}}{\partial H_t \gamma_2} = 0 \qquad t = T,..,\tau.$$

$$\frac{\partial J_{T-1}}{\partial G_{T-1}\gamma_1'} \gtrless 0$$

These results are exactly as in case 1 with the exception that G_T does not enter into the decision given that a birth in period T is ruled

out. There are no intertemporal interactions because the decision at T is independent of the decision at T-1, i.e, $n_T = 0$ is always optimal.

The period T-2 decision is of the same form as the period T-1 decision that appears in (15) and (16) except that T-2 replaces T-1 and T-3 replaces T-2. However, the number of separate cases to be considered grows by a factor of 3 since the set of possible future paths consists of all feasible combinations of n_{T-1}, n_T, d_{T-1} and d_T. Each period adds three times as many possibilities so that the problem quickly explodes. It is, therefore, unwise to pursue analytically the implications of this model. It is sufficient to note from the two period solutions that there may be important interactions between the current decision and past realizations, and anticipated future variables. This is true even though child deaths are restricted to occur only in the birth period and maintenance costs are assumed negligible after the first period of life. If either of these assumptions is relaxed, more complex interactions would be induced.[14]

3. Simulations

Since it is difficult to gain further insights into the dynamic structure of the model in an analytical fashion, I present in this section a 20 period simulation of the model for a quadratic utility function. These simulations are intended to illustrate the flexibility of the model in generating fertility patterns and replacement responses. Note that at this point the researcher is assumed to know as much as the individual decision-maker.

Table 1 presents simulations for a quadratic utility function over several sets o parameters. Simulations A.1 to A.8 and B.1 to B.4 assume a flat income profile. The su vival probability is assumed to be constant in all simulations. When the income profile is flat (or declining) it would be optimal if unconstrained to have all children at once. However, given a maximum of one birth per period the optimum birth sequence is one without spacing. An asterisk indicates a child death so that cor trasting rows with different numbers of asterisks implies something about replacement behavior. For example, simulations A.1 to A.4 imply that there is no replacement if the first child dies, replacement of one child if the first and fifth child dies, and replacement of one child if the first, fifth, and tenth child dies. With the probability of survival reduced to .8 (from .9), replacement behavior is different; the first child is replaced as is both the first and fifth if they die, but only two children are replaced if the first, fifth and tenth die. Simulations B.1 to B.4, on the other hand, indicate full replacement.[15]

All of the other simulations posit a rising income profile. In each simulation a fertility pattern with spacing emerges. Rising income creates an incentive to have children later. Spacing emerges because of the tension this creates with the underlying desire to accumulate children rapidly. From the simulations reported, it would appear that most any

Table 1

Simulations

Model: $\max_{\{d_t\}} \sum_{t=1} \delta^{t-1}(\alpha_1 M_t - \alpha_2 M_t^2 + \beta_1 X_t - \beta_2 X_t^2 + \gamma M_t X_t)$

s.t. $Y_t = X_t + b(n_t - d_t) + cn_t$

$M_t = M_{t-1} + n_t - d_t$

$Y_t = \theta_0 + \theta_1 t \quad , t = 1,\ldots,20$

$Y_t = \theta_0 + \theta_1 (21) \quad , t = 21,\ldots,40$

$P_t = \text{prob}\ (d_t = 0)$

		δ	α_1	α_2	β_1	β_2	b	c	γ	θ_o	θ_1	P	1 2 3 4 5	6 7 8 9 10	11 12 13 14 15	16 17 18 19 20
A.	1.	.95	1500	15	25	.0009	400	500	0	5000	0	.9	1 1 1 1 1	1 1 1 1 1	1 1 0 0 0	0 0 0 0 0
	2.												1*1 1 1 1	1 1 1 1 1	1 1 0 0 0	0 0 0 0 0
	3.												1*1 1 1 1*	1 1 1 1 1	1 1 1 0 0	0 0 0 0 0
	4.												1*1 1 1 1*	1 1 1 1 1*	1 1 1 0 0	0 0 0 0 0
	5.											.8	1 1 1 1 1	1 1 1 1 1	0 0 0 0 0	0 0 0 0 0
	6.												1*1 1 1 1	1 1 1 1 1	1 0 0 0 0	0 0 0 0 0
	7.												1*1 1 1 1*	1 1 1 1 1	1 1 0 0 0	0 0 0 0 0
	8.												1*1 1 1 1*	1 1 1 1 1*	1 1 0 0 0	0 0 0 0 0
	9.									1500	250	.9	0 1 1 1 0	1 1 1 0 1	1 0 1 1 0	1 0 0 0 0
	10.											.8	0 0 0 0 0	1 1 1 0 1	1 1 0 1 1	0 0 0 0 0
	11.											.7	0 0 0 0 0	0 0 0 0 1	0 1 1 0 1	0 0 0 0 0
	12.									1000	200	.9	1 1 1 0 1	0 1 1 0 0	0 0 0 0 0	0 0 0 0 0
	13.											.8	0 0 0 0 1	0 1 0 1 0	0 0 0 0 0	0 0 0 0 0
B.	1.	.95	2750	70	35	.002	400	200	0	4000	0	.9	1 1 1 1 1	1 1 1 1 1	1 1 1 0 0	0 0 0 0 0
	2.												1*1 1 1 1	1 1 1 1 1	1 1 1 1 0	0 0 0 0 0
	3.												1*1 1 1 1*	1 1 1 1 1	1 1 1 1 1	0 0 0 0 0
	4.												1*1 1 1 1*	1 1 1 1 1*	1 1 1 1 1	0 0 0 0 0
	5.						200	800		2500	200		1 1 1 1 1	0 0 1 0 0	1 0 0 1 0	1 0 0 0 0
	6.											.8	1 1 1 0 0	1 0 0 1 0	0 1 0 1 0	1 0 1 1 1
	7.										150	.9	1 1 1 1 1	1 0 1 0 0	0 1 0 0 0	1 1 0 1 1
	8.						500						1 0 0 1 0	0 0 1 0 0	1 0 0 1 0	0 0 1 0 0

spacing pattern can be generated simply by varying the income profile and the probability of survival. Although not shown, it should be clear that a rising survival probability profile would generate the same qualitative result with respect to spacing.

4. Method of Estimation

In general, the decision rules will be functions of higher order moments of the random variables (see Appendix A). This is true even with the simplifying assumptions used in the derivation of the decisions rules in section two. An important simplification is achieved if it is assumed that the utility function in quadratic. Decision rules turn out to be linear so that only conditional first moments enter the solution and is analogous to the certainty equivalence result for continuous variables. Therefore, let

$$(24) \quad U(M_{it},X_{it}) = \alpha_{1it}M_{it} - \alpha_2 M_{it}^2 + \beta_1 X_{it} - \beta_2 X_{it}^2 + \gamma M_{it}X_{it},$$

where i denotes the household and t denotes life-cycle period. All parameters are presumed to be positive other than γ which may be of either sign. I also assume that households differ in their value of α_1, according to

$$(25) \quad \alpha_{1it} = \bar{\alpha}_1 + \xi_{it},$$

for reasons that are developed below. In order to present the estimation strategy in an orderly fashion, two cases are considered, the first where ξ_{it} is distributed independently and identically over time for each household, and the second where ξ_{it} follows a permanent-transitory scheme. I do not pursue more general formulations since they would be significantly more expensive to implement, although not necessarily more difficult to characterize.

If the utility function is quadratic, and if the ξ_{it}'s (and as before the v_{it}'s) are conditionally (given the information set at t) distributed independently of the u_{it}'s (the mortality error process) the decision rule is linear.[16] The decision rules at any time t will depend only on the conditional expectations of unknown income shocks $(v_{it},\ldots,v_{iT})$, the future preference shocks $(\xi_{it+1},\ldots,\xi_{iT})$, and the conditional distribution of the unknown mortality shocks $(g_t(u_{it},\ldots u_{iT}))$.

The decision rule for any arbitrary period t can, therefore, be written as

$$(26)\quad n_t = 1 \text{ iff } \quad J_t(\Omega_t, E_t\xi_{it+1}, \ldots, E_t\xi_{iT}, g_t(u_t,\ldots,u_T), E_t v_t, \ldots, E_t v_T; \theta) \geq 0$$

$$= 0 \text{ iff } \quad J_t < 0$$

where $\Omega_t = (\xi_{it}, M_{t-1}, G_1\gamma_1, \ldots G_T\gamma_1, H_1\gamma_2, \ldots, H_\tau\gamma_2, d_{t-1}, \ldots, d_1, v_{t-1}, \ldots, v_1)$ is the household's information set at t and θ consists of the parameters $\bar{\alpha}_1$, $\alpha_2, \beta_1, \beta_2, \gamma, c_1, c_2, \delta$. The exact form of J_t is not tractable to derive since it involve solving the complete dynamic programming problem back to t.[17] Everything in Ω_t is assumed to be available to the researcher except ξ_{it}. The parameters γ_1 and γ_2 are assumed to be estimable from data.

Estimation Method with Time Independent Preferences

Consider first the case where preferences are independently drawn each period from the same distribution with zero mean. Thus, $E_t\xi_{it+j} = 0$ for all $j>0$. In this formulation households are alike in their underlying preferences for children except that in each decision period the household draws a

random child preference parameter (α_{1it}). In calculating the optimal fertility path as of any particular time, the household assumes, given that $E_t\xi_{it+j} = 0$ for $j > 1$ that it will behave in each future period as if its child preference parameter was at the average value ($\bar{\alpha}_1$). Thus, for any given set of parameter values, θ, the exact desired fertility path from period t+1 to T as of t is determinate. The choice at t, although deterministic for the household since ξ_{it} is known at t, is stochastic as viewed by the researcher.

Now, given θ and assuming that the mortality and income processes are known by the researcher, one can find for each t a value of ξ_{it}, say ξ^*_{it}, for which $J_t = 0$. Notice that since analytical solutions for decision rules are not available, the solution for ξ^*_{it} must be numerically obtained. Note that the value of ξ^*_{it} depends on the choice of parameters θ and on the past history of births and deaths as summarized by M_{t-1}. As long as the decision rule is monotonic in α_{1it}, the ξ^*_{it} computed in this way must exist and be unique. It is fairly obvious that the relative payoff of a child at time t rises with the marginal utility of a child as reflected in α_{1it}. In fact, J_t is linear in α_1 in the quadratic formulation.[18]

The probability that individual i is observed to have a child in period t conditional on the number of surviving children, M_{t-1}, is given by

$$(27) \quad \Pr(n_t = 1 | M_{t-1}) = \int_{\frac{\xi^*_{it}}{\sigma}}^{\infty} f(\xi_{it})\, d\xi_{it} = 1 - F\left(\frac{\xi^*_{it}}{\sigma}\right)$$

where $f(\xi_{it})$ is the density of ξ_{it}, with mean zero and unit standard deviation, σ is the standard deviation of the unstandardized density, and $F(\varepsilon_{it})$ is the cumulative distribution. Similarly,

$$(28)\quad \Pr(n_t = 0|M_{t-1}) = \int_{-\infty}^{\frac{\xi^*_{it}}{\sigma}} f(\xi_{it})d\xi_{it} = F(\underline{\xi^*_{it}}).$$
$$\sigma$$

It is straightforward to form the likelihood function over any number of periods, that is for any sequence of births, and over individuals. Consider household i which has a birth sequence $(\bar{n}_{i1}, \bar{n}_{i2},\ldots\bar{n}_{iT})$ for periods 1 to T where n_{it} equals one or zero. For each time t a ξ^*_{it} is calculated from the dynamic programming solution algorithm from which the likelihood of that sequence can be obtained as the product of the appropriate probabilities as given by (27) and (28). The full likelihood over the sample of I women is given by $\prod_{i=1}^{I}\prod_{t=1}^{T} \Pr(n_{it} = \bar{n}_{it}|M_{it-1})$. Choosing alternative values of θleads to different values of ξ^*_{it} for each individual and period, and therefore, to different sample likelihood values.[19] A derivative free maximization routine is required given the absence of analytical derivatives. The utility function parameters, $\bar{\alpha}_1$, α_2, β_1, β_2. and γ are identified only up to a factor of proportionality, σ.[20] This is so because ξ^* is homogeneous of degree one in those parameters in the quadratic case. However, b, c and δ are identified without normalization since J_t is not homogeneous in these parameters.[21]

Since ξ_{it} is assumed to be independently distributed over time, any set of periods may be used to calculate consistent maximum likelihood estimates. The computational burden lies in the fact that for each set of parameter values, the dynamic programming algorithm must be used to calculate the new set of ξ^*_t's for each individual. It is possible that a single year of data may be sufficient to identify all of the structural parameters of the model. Identification rests upon there being a sufficient number of future period observations on expected income and survival probabilities from which th structural parameters are implicitly retrieved (there are eight in the model presented above). However, as already noted, the decision rule is, in general, interactive in the exogenous variables so that there are, in essence, more "reduced form"

parameters than is immediately apparent. As the decision period is moved back in time, the number of "reduced form" parameters increases rapidly. Notice that since the functional forms of these decision rules are "unknown" it is difficult to make a priori statements about identification.

Estimation Method with Household Specific Preferences

A more general strategy would combine an unobserved permanent taste component, ν_i, with a purely transitory component, η_{it}. Thus, $\alpha_{lit} = \bar{\alpha}_l + \nu_i + \eta_{it}$, where $\xi_{it} = \nu_i + \eta_{it}$. Since there is a persistent unobserved error, it is not immaterial as to which periods are used in the estimation. Since the decision rule for any intermediate period t is conditioned on the previous fertility decisions (M_{t-1}) which are themselves related to the household specific component (ν_i), estimates will be consistent only if the decision is conditioned upon the initial state since the initial state variable, M_0, is zero for all women and so is itself non-stochastic.[22] It is, therefore, crucial that the data contain complete life cycle information. It is assumed that the household knows the permanent taste factor ν_i at the beginning of the life-cycle. The recursive solution of the maximization problem therefore requires that α_{lit} be equal to $\bar{\alpha}_l + \nu_i$ from the final decision period (T) back to the initial period. In addition, as in the first case, the household draws another random term η_{it} at the time the decision is made, i.e., at t.

To simplify the discussion, let ν_i take on only two values $\pm \nu$ with probability q and 1-q respectively, i.e., there are only two types of households. Extensions to more heterogeneous populations will be obvious. The estimation strategy can be heuristically described by the following steps. Choose a set of structural parameters P. For a given value of ν compute the value of η_{i1} for each i which sets $J_1 = 0$ as was done in the first part of this section. Denote this value as $\eta_{i1}^{*(+)}$. Do the same for $-\nu$ and denote the critical value of η_{i1} as $\eta_{i1}^{*(-)}$. Repeat for the second period, third period, etc., denoting $\eta_{it}^{*(+)}$ as the critical value at t for $+\nu$ and $\eta_{it}^{*(-)}$ as the critical value at t for $-\nu$, noting that each value is conditional on the past birth and death sequence, M_{t-1}. The probability that a household has a particular birth sequence as of any period t is the probability that the household has $\nu_i = +\nu$ (q in this example) times the probability of the birth sequence plus the probability that the household has $\nu_i = -\nu$ (1-q in this example) times the probability of the birth sequence. Thus, for example,

$$\Pr(n_{i1} = 1, n_{i2} = 1) = q[\Pr(n_{i2} = 1 | n_{i1} = 1, d_{i1}, +\nu_i)\Pr(n_{i1} = 1 | +\nu_1)]$$

$$+ (1 - q)[\Pr(n_{i2} = 1 | n_{i1} = 1, d_{i1}, -\nu_i)\Pr(n_{i1} = 1 | -\nu_i)]$$

$$= q\left[\int_{\eta_{i2}^{*(+)}}^{\infty} f(\eta_{it})d\eta_{it} \int_{\eta_{i1}^{*(+)}}^{\infty} f(\eta_{it})d\eta_{it}\right]$$

$$+ (1 - q)\left[\int_{-\infty}^{\eta_{i2}^{*(-)}} f(\eta_{it})d\eta_{it} \int_{-\infty}^{\eta_{i1}^{*(-)}} f(\eta_{it})d\eta_{it}\right]$$

Similar expressions are easily derived for other birth sequences, with the sample likelihood given by products of birth sequence probabilities over households.

One can iterate on the values of ν and q until the likelihood of observing the sample is maximized. Choose another set of structural parameters (P), and again find the maximum likelihood estimate of ν and q. This procedure is repeated until the sample likelihood is maximized over ν and q, and the set of structural parameters. Of course, the actual procedure need not proceed in this manner; ν and q are treated simply as additional parameters and optimized jointly with the other parameters.

5. Data

The model can be estimated with data containing life-cycle information on births, child deaths, and husbands' income. The 1976 Malaysian family Life Survey is ideal for this purpose. It contains 1262 households consisting of at least one ever married woman under fifty years of age as of the survey data. The sample is from Peninsular Malaysia with some slight oversampling of fishing communities and Indian families. The essential feature of the survey is that it contains a retrospective life history of each woman to the earlier of age fifteen or age at marriage with all of the necessary information.

The sample actually used in the estimation consists of 188 women. From the approximately 50% of the households that are Malay (the other 50% consist of ethnic Chinese and Indians), the sample was further restricted to currently married women over the age of thirty who have been married only once and for whom there is no missing information. Since perfect foresight is n assumed about life-cycle income of the husband, this age restriction insured sufficie observations on husband's income to permit a reasonable forecasting equation to be estimated.

The length of the period was chosen to be eighteen months as a compromise between the necessity to have only a single birth within a period and the computational benefit from having fewer periods in the life-cycle. Of the total of 3086 periods, in only 30 periods (27 women) were there two births within eighteen months. In those cases, a birth was moved to the

next period in which there was no birth. Each woman's initial period was set at age fifteen or age at marriage, whichever is first. The implicit assumption is that marriage is not subject to choice prior to age fifteen but is an independent decision, not unrelated to the fertility decision, after age fifteen. The final decision period is assumed to occur after twenty periods or approximately at age forty-five, given that each period is eighteen months in length.

Husband's income (earnings) is reported continuously beginning at age fifteen. That is, labor supply and wage rates are reported retrospectively at each moment in time that either of them changes, beginning at husband's age fifteen. Thus, starting from the initial period of the woman, her husband's income, possibly prospective if the woman was not married at that time, can be computed for each eighteen month period. There are potentially as many data points on husband's income as there are eighteen month periods in the husband's life, aligned according to the wife's life cycle. A semi-log (price deflated) earnings function was estimated for each husband individually with period and period squared as regressors, where period is simply age of husband suitably transformed to the wife's life cycle.[23] The log of earnings was used as the dependent variable so that earnings predictions would always be positive. A predicted earnings profile was then generated and is the basis for the earnings observations used in the empirical analysis.[24] The woman is assumed to live for ten periods (fifteen years) after her child bearing period, and husband's income is assumed to be constant and equal thereafter to predicted income in the twenty-first period. It is thus assumed that each woman knows with certainty the parameters of her husband's earnings function at the beginning of her life cycle,

Notice that log earnings is not given an autoregressive structure; new information acquired during the life cycle requires updating the future component of the dynamic programming decision rule each period, which greatly increases the computational burden. What is optimal in the future, say at t+1, as of t is not necessarily optimal as of t-1 when less information is available.

The survival probability was obtained from state data on survival rates.[25] The idea is that individuals use information about the mortality experience of those in similar environments. Perhaps less aggregated data would be more appropriate, but such information is not available. The predicted survival probabilities are based upon a log odds regression on time, its square and a constant for each state. The actual data spanned the period from 1948 to 1975 while predicted survival rates were needed from 1940 to 1990 given the cohort span of the women in the sample. Details of the entire procedure are provided in Appendix B.

The data therefore consists of the actual occurrences of births and deaths for each woman, a predicted husband's income profile based on a semi-log quadratic in time formulation, and a predicted survival probability profile based on a quadratic in time logistic formulation. At each moment in time, the probability of a birth

depends in a complicated way upon the number of surviving children to that time, the predicted future income of her husband, and the predicted future survival probabilities.

The actual mean birth probability profile is presented in Table 2 for the complete twenty periods. Note that sample size falls after eleven periods from 188 women for periods one through eleven to 136 women by period fifteen to only 44 women by period twenty. The birth probability is very low in period one, rises to a peak in period four, and remains roughly constant until period ten when it begins a slow decline. The very low probability in period one is not due solely to a low marriage rate. In period one 58.5% of the women were married while by the beginning of period four 91% of the women were married. The same birth probability in period four applied only to married women would have yielded a birth probability of 34.5% in period one rather than the observed 13.8%.

In the twenty periods, there are 2^{20} possible birth sequences. With only 188 women it is impossible to determine with any accuracy the sample likelihood of observing all of these different paths. However, the state variable in the model in the preceding section is the number of surviving children rather than their arrival sequence. Recall that for a single individual the probability of a birth at t will be a declining function of the number of surviving children to t. Looking across women is not appropriate since they differ in observed and unobserved characteristics that presumably affect the fertility outcomes. Nevertheless, such a tabulation is presented below as a descriptive device and to obtain some indication of the existence of permanent heterogeneity.

Table 2

Sample Birth Probability Profile

Period	Probability	Period	Probability
1	.138	11	.410
2	.335	12	.425
3	.489	13	.321
4	.537	14	.360
5	.511	15	.310
6	.505	16	.262
7	.511	17	.200
8	.516	19	.140
9	.516	19	.140
10	.394	20	.105

Given the small number of women that form the sample, the contingency table showing the probability of births conditional on the number of surviving children is presented in Table 3 with a moving average of surviving children. This smooths out the relationship, hopefully making whatever pattern does exist more discernable. In addition, Table 3 presents only the first fourteen periods as the number of births becomes too few after that time. Table 3 does not reveal a strong tendency for women with more surviving children to have a larger probability of a birth as would be expected if there was significant observed or unobserved heterogeneity in the underlying probability of a birth. Such a tendency is only revealed beginning at period eight, with a more pronounced trend appearing after period eleven. In several periods, the pattern is of an inverted U shape ignoring cells with few observations. For most periods, one cannot reject the hypothesis at conventional significance levels that there is no relationship between the number of surviving children and the birth propensity, as revealed by the χ^2 statistics presented in Table 3. Of course, there may be countervailing forces which, if they exist, should be revealed upon estimating the model.

As another crude measure of the importance of heterogeneity, the observed variance in the number of children born over all women and all periods was computed and compared to the variance that would be expected if the birth process was simply Bernoulli. The former is obtained from $\Sigma(N_j - \bar{P}H_j)^2$ where N_j is the number of children born to woman j, $\bar{P}$ is the sample probability of a birth (.316), and H_j is the number of periods for woman j; the observed variance is equal to 1394. The expected variance is obtained from the conventional variance formula for a binomial distribution $\bar{P}(1 - \bar{P})\Sigma H_j$ and is equal to 733.

Table 3

Probability of a Birth Conditional on the Number of Surviving Children

Number of Surviving Children	Period											
	3	4	5	6	7	8	9	10	11	12	13	14
0 or 1	.486	.533	.477	.507	.538	.389	.435	.063*	.438	.300*	.400*	.000*
1 or 2	.549	.604	.589	.509	.558	.500	.408	.379	.367	.292	.211*	.133
2 or 3		.556	.544	.514	.495	.529	.470	.375	.275	.209	.256	.241
3 or 4			.555	.532	.429	.505	.570	.451	.377	.320	.225	.289
4 or 5				.400*	.500	.543	.594	.442	.430	.574	.220	.369
5 or 6					.800*	.867*	.519	.415	.435	.574	.323	.400
6 or 7						1.000*	.583*	.550	.552	.459	.379	.340
7 or 8							.500*	.444	.533	.350	.483	.333
8 or 9								.500*	.677*	.250*	.500	.474
9 or 10									1.000*	.677*	.333*	.556*
10 or 11										1.000*	1.000*	1.000*
11 or 12											1.000*	1.000*
12 or 13												1.000*
Number of children born	92	101	96	95	96	97	97	74	78	71	51	50
Number of women	188	188	188	188	188	188	188	188	188	167	159	147
χ^2	.813	1.30	2.95	1.52	4.77	16.7	6.63	14.5	11.1	28.9	16.3	18.8
d.f.	1	2	3	4	5	6	7	8	9	10	11	12
Probability	.38	.53	.41	.82	.45	.01	.47	.07	.27	.01	.14	.09

*indicates 5 or fewer observations in the appropriate cell

In the most extreme case where women either had a child each period or no children at all, the variance in total number of children would be 11,984. Clearly, homogeneity is a more reasonable interpretation, and particularly if important observed determinants of fertility are included in the analysis.

6. Estimation Results

The model was estimated for two samples, one using the birth outcomes of the initial ten periods of each woman's life cycle and the other using birth outcomes over each woman's entire life cycle. The ten period sample was used primarily to test for unobserved heterogeneity since the composition of the sample changes as more periods are added. Given computational cost limitations only a model with two types of individuals was estimated as discussed in section 4. Since a likelihood ratio test clearly rejected the existence of unobserved heterogeneity, results from the ten period sample are not reported.[26] It is assumed, therefore, that unobserved heterogeneity will not, if ignored, seriously contaminate estimates using the entire life cycle of each woman. The estimates presented below therefore are based on the assumption that the random component of the model (ξ_t) follows a serially independent stochastic process.

The model is formulated slightly more generally than in previous sections. In particular, the cost of birth is permitted to differ in the first and second periods from each other and from subsequent periods and the cost profile is permitted to have a quadratic shape. It is unlikely a priori that income and survival probability profiles can by themselves trace the life-cycle fertility profile shown in Table 2. Moreover, presumed biological constraints early and late in the life-cycle associated with a reduced propensity to conceive can be formally expressed as a reduced contraceptive cost or an increased net cost of a birth.

The structure of the model is, therefore, as follows:

$$LU = \sum_{t=1}^{\tau} \delta^{t-1}(\alpha_{1t}M_t - \alpha_2 M_t^2 + \beta_1 X_t - \beta_2 X_t^2 + \gamma_1 M_t X_t + \gamma_2 M_t S)$$

$$Y_t = X_t + b(n_t - d_t) + (c_1 m_1 + c_2 m_2 + c_{30} + c_{31} t + c_{32} t^2) n_t$$

$$\alpha_{1t} = \bar{\alpha}_1 + \xi_t, \quad \xi_t \sim N(0,\sigma)$$

where, in addition to previous definitions, m_1 is a dummy variable equal to unity if the period is the first and zero otherwise, m_2 is a dummy variable equal to unity if the period is the second and zero otherwise, and S is the woman's years of schooling. Thus $c_1 + c_{30} + c_{31} + c_{32}$ measures the birth cost in period one, $c_2 + c_{30} + 2c_{31} + 4c_{32}$ the birth cost in period two, and $c_{30} + c_{31} t + c_{32} t^2$ the birth costs in all periods after the second. Notice that schooling is arbitrarily introduced as affecting the incremental utility of surviving children. Obviously, schooling could be modelled as affecting any or all of the cost parameters or any or all of the other utility function parameters. Since it is unlikely that the model could distinguish between these alternatives, the simplest approach was adopted.[27]

Table 4 presents the estimates of the model under the assumption that ξ_t is normally distributed.[28] It should be noted that the model was also estimated under the restriction that $c_{31} = c_{32} = 0$. Not only was that joint hypothesis resoundingly rejected, but the parameter estimates were not reasonable.[29] Some of the utility function and cost parameters were negative as was the discount factor. Those results are, therefore, not specifically described.

Although individual parameters cannot be translated into the experiments of interests, they do provide a simple check on the plausibility of the estimates, and thus, of the method. All of the signs in Table 4

Table 4

Maximum Likelihood Structural Parameter Estimates

Parameter	Estimate	Asymptotic Standard Error
$\bar{\alpha}_1$	3.854×10^{-2}	2.274×10^{-3}
α_2	5.761×10^{-7}	2.187×10^{-6}
β_1	5.376×10^{-5}	1.863×10^{-5}
β_2	4.130×10^{-15}	1.198×10^{-14}
γ_1	4.038×10^{-8}	1.195×10^{-8}
γ_2	-4.614×10^{-3}	1.328×10^{-3}
b	3.433×10^{2}	2.883×10^{2}
c_1	2.228×10^{4}	8.736×10^{3}
c_2	8.121×10^{3}	4.701×10^{3}
c_{30}	2.884×10^{3}	5.230×10^{2}
c_{31}	-3.025×10^{1}	1.136×10^{1}
c_{32}	5.782×10^{1}	2.530×10^{1}
δ	9.155×10^{-1}	6.565×10^{-3}

ln Likelihood = -1923.7

conform to rather strong priors; incremental or marginal utilities are positive and diminishing in surviving children and in the composite good, the cost of bearing a child is positive in all life cycle periods as is the time invariant maintenance cost, and the discount factor translates into a rate of time preference of approximately .093. Magnitudes of the cost parameters are also not unreasonable. The cost of a live birth in period one (c_1) is 22,280 Malaysian dollars and in period two (c_2) 8,121 Malaysian dollars. The cost of a birth is lowest in period three, 3,314 dollars, and rises at an increasing rate reaching 25,407 dollars by period twenty. These costs may seem slightly high given that the Malaysian dollar was worth about 33% of the U.S. dollar in 1960. The child maintenance cost, on the other hand, is rather small, only 343 dollars.[30] Notice also that the incremental utility of an additional surviving child rises with goods consumption ($\gamma_1 > 0$) and falls with mother's schooling ($\gamma_2 < 0$). However, quantitatively the non-linear terms in the utility function other than schooling are quite small and this feature of the results is crucial to the comparative dynamics of the model discussed below.

There are several ways to assess the fit of the estimated model. The ℓn likelihood value as shown in Table 4 is -1923.7. The model degenerates to a simple Bernoulli process for births given that all of the parameters except $\bar{\alpha}_1$ are set to zero; $\bar{\alpha}_1$ must be estimated to retrieve the sample fraction of births. The ℓn likelihood value for the pure chance model is -2059.9. Twice the difference in the likelihoods of the two models is 272.4 which is sufficient to reject the pure chance model at almost any significance level; $\chi^2(.01) = 26.2$ with 12 degrees of freedom. In addition, most of the individual parameters are statistically significant at conventional levels (9 out of 13).

Table 5 compares the actual and predicted birth probability profiles while Table 6 uses the estimated parameters to predict birth probabilities conditional on the number of surviving children, analogous to Table 3. In both tables, it appears that the model does fairly well. Indeed, the χ^2 statistics for the hypothesis that Table 6 and Table 3 are the same, i.e., that the predicted and actual probabilities do not differ, is only rejected in one period as shown in Table 7. Overall, then, the model seems to capture some significant features of the data.

The estimated parameters may be used to explore the sensitivity of fertility over the life cycle to changes in income and survival probability profiles, to child deaths, and to mother's schooling. All of the results will be presented first with intuitive explanations offered subsequently.

Income Effects

Table 8 presents fertility responses to alternative income profiles evaluated at the mean survival probability (flat profile) and at mean mother's schooling (two years). To summarize the impact on the entire fertility profile, the table depicts the expected number of children born over periods one to five, six to ten, eleven to fifteen and sixteen to twenty where the expectations are sums of the unconditional probabilities of births within the subperiods, i.e., integrating over all possible values of the state variable (the number of surviving children). Since income was predicted from a semi-log income function, the experiments are performed by changing the income function parameters.

The first row shows the fertility profile for an individual with a flat income profile at the average sample income. In the next four rows

Table 5

Actual and Predicted Mean Birth Probabilities

Period	Actual Probability	Predicted Probability
1	.138	.138
2	.335	.360
3	.489	.552
4	.537	.543
5	.511	.530
6	.505	.515
7	.511	.498
8	.516	.477
9	.516	.454
10	.394	.428
11	.410	.400
12	.425	.376
13	.321	.345
14	.360	.312
15	.310	.281
16	.262	.250
17	.200	.216
18	.140	.182
19	.140	.149
20	.105	.122

Table 6

Probability of a Birth Conditional on the Number of Surviving Children: Predicted

Number of Surviving Children	Period											
	3	4	5	6	7	8	9	10	11	12	13	14
0 or 1	.552	.539	.523	.507	.481	.472	.391	.412	.375	.400	.429	.333
1 or 2	.563	.622	.540	.519	.500	.471	.408	.400	.400	.375	.368	.333
2 or 3		.555	.558	.513	.505	.480	.446	.422	.392	.372	.333	.310
3 or 4			.550	.511	.506	.484	.460	.423	.406	.380	.325	.316
4 or 5				.500	.500	.478	.464	.430	.407	.382	.341	.316
5 or 6					.600	.533	.444	.434	.406	.382	.354	.320
6 or 7						.750	.417	.444	.414	.378	.362	.321
7 or 8							.000	.444	.400	.233	.379	.333
8 or 9								.333	.333	.375	.417	.369
9 or 10									.000	.333	.333	.333
10 or 11										.000	.000	.000
11 or 12											.000	.000
12 or 13												.000

Table 7

Tests of Equality between Predicted and Actual Birth Probabilities

Conditional on Surviving Children

	Period											
	3	4	5	6	7	8	9	10	11	12	13	14
χ^2	3.24	1.34	2.11	.522	4.60	10.95	15.07	13.15	8.08	29.22	13.84	15.57
$\chi^2_k(.05)^*$	3.84	5.99	7.82	9.49	11.07	12.59	14.07	15.51	16.92	18.31	19.68	21.03

*k equals the period number minus two.

Table 8

The Effect of Income on Fertility

$\ln Y = a_0 + a_1 t$ a_0	a_1	$\bar{Y}$	N_{1-5}	N_{6-10}	N_{11-15}	N_{16-20}	N
9.35	0	11,500	2.124	2.369	1.687	.867	7.047
8.52	0	5,000	2.117	2.362	1.682	.863	7.024
9.21	0	10,000	2.121	2.367	1.685	.866	7.039
10.13	0	25,000	2.134	2.380	1.699	.873	7.086
10.82	0	50,000	2.154	2.402	1.717	.885	7.158
8.52	.088	11,500	2.126	2.376	1.769	.875	7.146

the mean income level is changed while maintaining the shape of the profile. Income effects are obviously quite small. Income elasticities tend to rise with the level of income with the arc elasticities ranging from .003 for income between 5,000 and 10,000 dollars to .015 for income between 25,000 and 50,000 dollars. The last row shows the impact on the fertility profile of altering the shape of the income profile while maintaining the mean at approximately 11,500 dollars. The intercept parameter corresponds to an income level of 5,000 dollars and income rises by almost 9% per period (about 6% per year). The fertility profile now is skewed slightly toward later births, although, given the quite substantial alteration in the income profile the impact is negligible.

Survival Probability Effects

Table 9 presents the fertility response to changes in survival probability profiles in a fashion analogous to the previous table. Since survival probabilities were predicted from a logistic formulation, experiments are performed by changing the parameters of that equation. The first row shows the fertility profile for a flat survival probability profile at the sample average survival probability (.94), assuming that income is at the sample average (flat profile) and that mother's schooling is approximately at the mean level. Reducing the survival probability by .05 percentage points (next four rows) reduces the number of children by about one-quarter of a child. Thus, women residing in environments with high survival propensities have more children. Notice also that as the survival probability drops, there is a tendency to have children earlier in the life cycle.

Table 9

The Effect of the Survival Probability on Fertility

$\ln \frac{p}{1-p} = a_0 + a_1 t$			N_{1-5}	N_{6-10}	N_{11-15}	N_{16-20}	N
a_0	a_1	$\bar{p}$					
2.78	0	.94	2.124	2.369	1.687	.867	7.047
2.09	0	.89	2.078	2.326	1.614	.783	6.801
1.67	0	.84	2.032	2.282	1.540	.701	6.555
1.33	0	.79	1.979	2.230	1.453	.612	6.274
1.05	0	.74	1.927	2.175	1.367	.532	6.001
1.05	1.81	.94	2.003	2.356	1.722	.937	7.018

Altering the shape of the profile so as to have a rising trend in the survival rate while maintaining the mean rate (last row) clearly skews the fertility profile toward delayed childbearing.

Replacement Effects

Table 10 shows the responsiveness of fertility to the occurrence of child deaths. Each entry gives the expected number of children ever born given that there are M_{t-1} surviving children as of period t for t = 2, ...10. Replacement responses can easily be calculated by differencing any two columns for a given row. Average income and survival probabilities, and average schooling are assumed in the calculations. The last column averages the replacement effects within each row. As Table 10 reveals, replacement responses are trivial for this sample. A child death induces an increase in the number of children ever born by at most .005.

Mother's Schooling Effects

Since mother's schooling merely shifts the marginal utility of a child in an additive manner, its impact on fertility is to change the level without changing the shape of the profile. Each additional year of schooling reduces the expected number of children ever born approximately by .35 evaluated at the mean of the other exogenous variables. Schooling therefore plays an important role in accounting for differences in completed fertility across households.

To recapitulate, the fertility level and profile respond negligibly to income levels and profiles and to child deaths. There is a noticeable response to the level and profile of infant survival probabilities and a

Table 10

Expected Number of Births by Number of Surviving Children and Period

Period	Number of Surviving Children										Average Replacement
	0	1	2	3	4	5	6	7	8	9	
2	6.916	6.911									.005
3	6.545	6.541	6.536								.005
4	5.997	5.993	5.987	5.982							.005
5	5.456	5.454	5.451	5.416	5.412						.003
6	4.897	4.897	4.893	4.889	4.884	4.884					.002
7	4.385	4.383	4.381	4.379	4.373	4.370	4.370				.003
8	3.890	3.886	3.886	3.882	3.879	3.874	3.874	3.874			.002
9	3.414	3.411	3.409	3.408	3.406	3.402	3.399	3.399	3.398		.002
10	2.964	2.960	2.956	2.956	2.953	2.951	2.946	2.946	2.946	2.943	.002

substantial impact of mother's schooling on completed fertility. Partial insights into the relationship of these experimental outcomes to the parameter estimates can be obtained by looking at the decision rule at period T for the quadratic utility function as given in footnote sixteen. Rearranging terms, the income effects in terms of the change in the differential payoff of a birth in period T are given by

$$\frac{\Delta J_T}{\Delta E_T Y_T} = 2\beta_2(bP_T + c), \quad \frac{\Delta J_T}{\Delta E_T Y_t} = \delta^{t-T}\gamma_1 P_T \qquad t > T$$

The magnitude of this change in the intensity with which a child is desired determines the change in the critical value of the taste parameter ξ_T^* and thus the change in the probability of a birth. Clearly, if β_2 is very small, as is evident in Table 4, the current income effect will be very small and if γ_1 is very small as is also evident in Table 4, future income effects will be small. Simulations reveal that large increases in β_2 and/or γ_1, of several orders of magnitude, given all other parameters, can generate non-negligible income effects. However, these increases are many standard deviations from the point estimate.

The replacement effect in period T is related to

$$\frac{\Delta J_T}{\Delta M_{T-1}} = -2\alpha_2 P_T\left(1 + \sum_{t=T+1}^{\tau} \delta^{t-T}\right) - \gamma_1(bP_T + c).$$

Replacement effects are obviously larger the greater are α_2 and γ_1. Our results reveal α_2 (and γ_1) to be indeed very small which accounts for the

trivial replacement response. Again, simulations reveal that orders of magnitude increases in α_2 or γ_1 can generate substantial replacement effects, changes which are far outside reasonable confidence intervals for the estimates.

The survival probability effect in period T, unlike the income or replacement effect, is related to the linear utility function parameters. In particular, letting $\alpha_2 = \beta_2 = \gamma_1 = 0$,

$$\frac{\Delta J_T}{\Delta P_T} = (\alpha_1 - b\beta_1)$$

which, given the parameter estimates, is not of negligible magnitude. Notice that if $b = 0$, an increase in the survival probability must increase the payoff to a birth. Indeed, even if all costs are zero, an increase in b will increase the intensity with which a child is desired (J_T) and thus will increase the probability of a birth, although in a deterministic world the decision to have a birth, i.e., the sign of J_T will be unaffected.

It would, thus, appear that, in large part, it is the higher order utility function parameters that are crucial in the determination of fertility profiles and their responsiveness to exogenous variables. In the extreme case where the utility function is linear in its arguments, and where the cost of birth and the survival probability do not change over the life cycle, it is intuitive that the fertility decision at any period would be independent of current or anticipated future income and of the stock of surviving children. Child spacing, in a probabilistic sense, would arise only from the random taste component of the model (ξ_t). Hence, given that

the utility function is estimated to be essentially linear for this sample of women (ignoring schooling effects), it is primarily variation in life cycle birth costs and survival probabilities that determine fertility profiles.

7. Conclusions

This paper is the first attempt to model sequential life cycle individual fertility behavior in an uncertain environment in a way that is conducive to estimation. As such, I believe that it demonstrates the empirical feasibility of the dynamic programming approach when the problem at least has a tractable numerical solution. However, many major simplifying assumptions were required and the sensitivity of the results to those assumptions is an important issue yet to be addressed. Extensions of the basic framework to incorporate endogenous female labor supply, savings, and imperfect contraception would comprise a ratherfull agenda for future research.

The particular problem addressed in this paper, that of the timing and spacing of children given significant infant mortality, has been of concern to many social science researchers. Specifically, "the replacement effect" estimated from this model, within a unified framework of the life cycle fertility decision, is much smaller than that obtained using other methods. At this stage, these estimates should be taken no more or less seriously than those based upon looser theoretical considerations.

FOOTNOTES

*Research support was received from the William and Flora Hewlett Foundation. Zvi Eckstein contributed many useful comments and substantially enhanced my understanding and appreciation of dynamic economic models. Randall Olsen provided extremely valuable comments particularly, though not exclusively, with respect to the estimation methodology. Comments from T. Paul Schultz and Paul McGuire are also gratefully acknowledged. Paul McGuire most ably performed the rather complicated calculations.

[1]As in all discrete time models, the length of a period is somewhat arbitrary. In this model, biological limitations provide some guidance.

[2]It is not apparent how to formulate and estimate a structural continuous time model of fertility. Vijverberg (1981) develops a model in which individuals choose the points during the life cycle at which to bear children. However, that methodology requires that the number of children be predetermined.

[3]Child mortality may, in part, be affected by the parental choice of the levels of birth and maintenance costs. However, allowing for endogeneity would greatly complicate the dynamics and the estimation. In addition, allowing maintenance costs to vary with the age of the child introduces overwhelming computational, though not conceptual, difficulties.

[4]Given the discrete nature of the outcome, spacing involves a fertility pattern in which births are not all consecutive, i.e., if one indicates a birth and zero not, then a 1, 0, 1 pattern is a birth sequence with spacing. Spacing is different than timing in that the latter relates to life-cycle phase. Thus, births may come predominately early or late in the childbearing stage, and they may be spaced or not. The model presented here solves for both aspects of fertility.

[5]If capital markets are perfect, child spacing could only be generated

if additional complementary or substitute commodities for children are introduced in the model, e.g., parental leisure. In addition, the introduction of savings opportunities adds another choice variable to the model.

[6]Explicit models of child mortality and fertility (e.g., Ben-Porath and Welch (1977)) have been static .

[7]The assumption that there is only a one period maintenance cost restricts the range of replacement behavior considerably. Replacement is more likely to depend upon life-cycle stage if the total maintenance cost at any time depended upon the age distribution of the living children.

[8]The model presented below is similar in structure to that of Heckman and Willis (1975). Their focus is on contraceptive uncertainty while I focus on survival uncertainty. They do not, however, attempt to solve the dynamic programming problem nor do they estimate structural parameters.

[9]Some of the statistical issues that arise in this kind of problem are discussed in Heckman (1980), although the statistical solution to concrete dynamic programming problems is not discussed there.

[10]Learning based on a woman's actual previous mortality experience is not considered. It is theoretically feasible to model, but it is not empirically tractable in the context of this model. Indeed, serial correlation in the u_t process will also be ruled out in the empirical application for similar reasons.

[11]As is the case with u_t, in principle, $Ev_t v_{t'}$, need not be zero for $t \neq t'$.

[12]In the terminology of Heckman (1980), there is true state dependence. As will become clear, however, there is no simple structural state dependence parameter. If a single state dependence parameter even exists, it will be a function of the underlying structural parameters of the model and so there are no additional identification issues raised in the application.

[13]There is a substantial simplification achieved by the assumptions concerning the mortality and income error processes. Notice, for example, that $\int_{-G_T\gamma_1}^{\infty} g_{T-1}(u_T)du_T = \int_{-G_T\gamma_1}^{\infty} g(u_T)du_T$. See Appendix A.

[14]The birth cost is permitted to change systematically over the life cycle in the empirical work that follows and introduces dynamic considerations that could mirror biologically related age variation in the cost of contraception which is not itself explicitly modeled.

[15]It is very difficult to simulate a hoarding strategy given that a child can die only in its first period of life.

[16]For example, with the utility function as in (24), the kernel of the decision rule at T is given by

$$J_T = P_T E_T(\alpha_{1it}) + P_T \sum_{t=T+1}^{\tau} \delta^{t-T} E_T(\alpha_{1it}) - \alpha_2(1 + 2M_{T-1})\left(1 + \sum_{t=T+1}^{\tau} \delta^{t-T}\right)P_T$$
$$-[\beta_1 b - (2\beta_2 b + \gamma)(b + c) + \gamma b M_{T-1}]P_T$$
$$+ 2\beta_2 b P_T E_T Y_T + (P_T)\gamma \sum_{t=T}^{\tau} \delta^{t-T} E_T Y_t$$
$$- (\beta_1 + \beta_2 c)c - \gamma c M_{T-1} + 2\beta_2 c E_T Y_T$$

Decision rules at times prior to T are simply sums of expressions that are in a form similar to, though more complex than, J_T; there are no second-order or higher terms in random variables. The independence assumption concerning the conditional distributions of the ξ_{it}'s and u_t's is necessary in order that no covariance terms enter the expression. Note also that $E_T(\alpha_{1iT}) = \bar{\alpha}_1 + \xi_{iT}$, i.e., ξ_{iT} is known to the women at T.

[17]In general, J_t is a function that is highly interactive in observables, $G_1\gamma_1, \ldots, G_T\gamma_1$, $g_t(u_t, \ldots, u_T)$, $E_t Y_t \ldots, E_t Y_\tau$, M_{t-1}, and highly non-linear in para-

meters, $\bar{\alpha}_1, \alpha_2, \beta_1, \beta_2, \gamma, c, b$. Note that c and b are treated as parameters because it is unlikely that any data set would contain them as variables. Of course, the form of J_t depends upon the optimal future path of births from t + 1 to T which cannot be determined without information about the utility function and cost parameter values.

[18] From footnote (16), it is easily seen that the only term in α_{liT} is $\alpha_{liT}P_T$. For the general case, at time t, the term in α_{lit} will be multiplicative in sums of products of probability of survival terms.

[19] From footnote 18, the critical value of ξ_{iT}, ξ^*_{iT} is given by

$$\xi^*_{iT} = [\alpha_2(1 + \sum_{t=T-1}^{\tau}\delta^{t-T}) + \beta_1 b - (2\beta_2 b + \gamma)(b + c) - \bar{\alpha}_1]$$
$$+ [2\alpha_2(1 + \sum_{t=T+1}^{\tau}\delta^{t-T}) + \gamma b]M_{T-1} - 2\beta_2 C\frac{E_T Y_T}{P_T}$$
$$- 2\beta_2 b E_T Y_T - \gamma \sum_{t=T+1}^{\tau}\delta^{t-T}\frac{E_T Y_t}{P_T} + \gamma c\frac{M_{T-1}}{P_T} + \frac{c(\beta_1 + \beta_2 c)}{P_T}$$

[20] Viewed differently, utility function parameters can only be determined relative to each other.

[21] For example, it is easy to see that doubling all of the utility function parameters in the expression in footnote [16] doubles J_T, while performing a similar operation for the rest of the parameters does not lead to a general effect on J_T.

[22] See Heckman (1980) for a discussion of the problem of initial conditions in discrete models, and for suggested solutions.

[23] Earnings are deflated by the price index that prevailed at the beginning of each eighteen month period with the base year 1960. The price index is country wide and was obtained from National Accounts of West Malaysia 1947-1976.

[24] The expected value of income was formed without the variance correction $e^{\frac{\sigma^2}{2}}$ that is strictly required under the assumption of normal errors in the log equation. Given the retrospective nature of the data, it was felt that measurement error would dominate the estimate of σ^2.

[25]There are eleven states in Peninsular Malaysia.

[26]The χ^2 value (twice the difference in the ln likelihood values of the two models) was .2 which falls far short of the critical χ^2 value with two degrees of freedom, $\chi^2(.05) = 5.99$.

[27]The average schooling level of the Malay women in this sample is low, only about two years. Over fifty percent have no schooling and of those who have any, the average level is only four and one-half years. Thus, schooling is usually completed prior to age fifteen. Schooling is, therefore, treated as a parental decision, exogenous with respect to fertility.

[28]The maximization routine used was the DFP component of GQOPT. Standard errors were calculated by inverting the matrix of second partials which were obtained by one percent perturbations of the parameters around the maximum.

[29]The χ^2 value for the likelihood ratio test was 204.2 which greatly exceeds the critical χ^2 value with two degrees of freedom, $\chi^2(.05) = 5.99$.

[30]It is possible, given actual income levels, that for some individuals at some time periods consumption will be negative. There is no simple way to impose a positivity constraint on consumption in the estimation, nor would that necessarily be desirable since the cost of a child would be dictated by the smallest income level of any woman in any period in which a child was born.

APPENDIX A

Begin with the T period decision in which the household chooses n_T to maximize

$$(A.1) \quad E_T \Big\{ U(M_{T-1} + n_T - d_T \, , \, Y_T - b(n_T - d_T) - cn_T) + \sum_{t=T+1}^{\tau} \delta^{t-T} U(M_{T-1} + n_T - d_T, \, Y_t) \Big\}$$

If (5) is substituted into (7), and (3) is utilized, the expected lifetime utility associated with having a child at time T (LU_T) conditional on the information set at time T can be written as

(A.2)

$$\begin{aligned} & E_T(LU_T / n_T = 1) \\ & = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \int_{-G_T\gamma_1}^{\infty} \Big[U(M_{T-1} + 1, \, H_T\gamma_2 + v_T - b - c \\ & \qquad + \sum_{t=T+1}^{\tau} \delta^{t-T} U(M_{T-1} + 1, \, H_t \gamma_2 + v_t) \Big] \\ & \qquad h_T(u_T, v_T, \cdots, v_\tau) \, du_T \, dv_T \cdots dv_\tau \\ & + \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{-G_T\gamma_1} \Big[U(M_{T-1}, \, H_T\gamma_2 + v_T - c) \\ & \qquad + \sum_{t=T+1}^{\tau} \delta^{t-T} U(M_{T-1}, \, H_t\gamma_2 + v_t) \Big] \\ & \qquad h_T(u_T, v_T, \cdots, v_\tau) \, du_T \, dv_T \cdots dv_\tau \end{aligned}$$

where $h_T(u_T, v_T, \ldots, v_\tau)$ is the joint conditional (on information at time T) density of the current and future random errors associated with child mortality and income. Recall that v_T and u_T are not known at T.

Similarly, the expected lifetime utility at T given that a child is not born is

$$(A.3) \quad E_T(LU_T/n_T = 0)$$

$$= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Big[U(M_{T-1}, H_T\gamma_2 + v_T) + \sum_{t=T+1}^{\tau} \delta^{t-T} U(M_{T-1}, H_T\gamma_2 + v_t) \Big] h_T(u_T, v_T, \ldots, v_\tau)\, du_T\, dv_T \ldots dv_\tau$$

Note that u_T is integrated out since a child that was not born cannot die.

Defining J_T to be the difference between (A.2) and (A.3), the decision to have a child at T is governed by

$$(A.4) \quad n_T = 1 \text{ if } J_T = E_T(LU_T/n_T = 1) - E_T(LU_T/n_T = 0) \geq 0$$

$$= 0 \text{ if } J_T < 0$$

If $h_T(u_T, v_T, \ldots, v_\tau) = g_T(u_T)\, f_T(v_T, \ldots, v_\tau)$, i.e., if u_T and the v_t's $t = T, \ldots, \tau$ are independently distributed given past information, then J_T may be simplified to

$$(A.5)\qquad J_T = \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\int_{-G_T\gamma_1}^{\infty}\Big\{U(M_{T-1}+1,\ H_T\gamma_2+v_T-b-c)-U(M_{T-1},\ H_T\gamma_2+v_T-c)$$

$$+\sum_{t=T+1}^{\tau}\delta^{t-T}\left[U(M_{T-1}+1,\ H_t\gamma_2+v_t)-U(M_{T-1},\ H_t\gamma_2+v_t)\right]$$

$$g_T(u_T)\ f_T(v_T,\ \ldots,\ v_\tau)\ du_T\ dv_T\ \ldots\ dv_\tau$$

$$+\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\Big\{U(M_{T-1},\ H_T\gamma_2+v_T-c)-U(M_{T-1},\ H_T\gamma_2+v_T)\Big\}$$

$$f_T(v_T,\ \ldots,\ v_\tau)\ dv_T\ \ldots\ dv_\tau$$

For simplicity only, this independence assumption is adopted throughout the rest of the discussion. In the text $\int_{-G_T\gamma_1}^{\infty} g_T(u_T)\ du_T = \int_{-G_T\gamma_1}^{\infty} g(u_T)\ u_T = P_T$ and the integrals over the income density are collapsed to an expectational representation. The implications derived in the text assume that the utility function is contemporaneously separable.

The expected lifetime utility given that a child is born at T-1 is much more complicated than its counterpart at T. It is given by

(A.6) $E_{T-1}(LU_{T-1} | n_{T-1} = 1)$

$$= \int_{-\infty}^{\infty} \int_{-G_{T-1}\gamma_1}^{\infty} U(M_{T-2} + 1,\ H_{T-1}\gamma_2 + v_{T-1} - b - c)\, g_{T-1}(u_{T-1})\, f_{T-1}(v_{T-1})\, du_{T-1}\, dv_{T-1}$$

$$+ \int_{-\infty}^{\infty} \int_{-\infty}^{-G_{T-1}\gamma_1} U(M_{T-2},\ H_{T-1}\gamma_2 + v_{T-1} - c)\, g_{T-1}(u_{T-1})\, f_{T-1}(v_{T-1})\, du_{T-1}\, dv_{T-1}$$

$$+ \delta \max \left\{ \begin{array}{l} \displaystyle\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \int_{-G_T\gamma_1}^{\infty} \int_{-G_{T-1}\gamma_1}^{\infty} \Big[U(M_{T-2} + 2,\ H_T\gamma_2 + v_T - b - c) \\ \qquad + \displaystyle\sum_{t=T+1}^{\tau} \delta^{t-T-1}\, U(M_{T-2} + 2,\ H_t\gamma_2 + v_t) \Big] \\ \qquad g_{T-1}(u_{T-1},\ u_T)\, f_{T-1}(v_T, \ldots, v_\tau)\, du_{T-1}\, du_T \\ \qquad dv_T \ldots dv_\tau \\ \\ \displaystyle\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \int_{-\infty}^{-G_T\gamma_1} \int_{-G_{T-1}\gamma_1}^{\infty} \Big[U(M_{T-2} + 1,\ H_T\gamma_2 + v_T - c) \\ \qquad + \displaystyle\sum_{t=T+1}^{\tau} \delta^{t-T}\, U(M_{T-2} + 1,\ H_t\gamma_2 + v_t) \Big] \\ \qquad g_{T-1}(u_{T-1}, u_T)\, f_{T-1}(v_T, \ldots, v_\tau)\, du_{T-1}\, du_T \\ \qquad dv_T \ldots dv_\tau \\ \\ \displaystyle\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \int_{-G_{T-1}\delta_1}^{\infty} \Big[U(M_{T-2} + 1,\ H_T\gamma_2 + v_T) \\ \qquad + \displaystyle\sum_{t=T+1}^{\tau} \delta^{t-T}\, U(M_{T-2} + 1,\ H_T\gamma_2 + v_t) \Big] \\ \qquad g_{T-1}(u_{T-1})\, f_{T-1}(v_T, \ldots, v_\tau)\, du_{T-1} \\ \qquad dv_T \ldots dv_\tau \end{array} \right.$$

$$
\begin{aligned}
&\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\int_{-G_T\gamma_1}^{\infty}\int_{-\infty}^{-G_{T-1}\gamma_1} [U(M_{T-2}+1,\ H_T\gamma_2 + v_T - b - c) \\
&\qquad + \sum_{t=T+1}^{\tau} \delta^{t-T} U(M_{T-2}+1,\ H_t\gamma_2 + v_t)] \\
&\qquad g_{T-1}(u_{T-1},\ u_T) f_{T-1}(v_T, .., v_\tau)\, du_{T-1}\, du_T \\
&\qquad dv_T, .. dv_\tau \\
&+ \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{-G_T\gamma_1}\int_{-\infty}^{-G_{T-1}\gamma_1} [U(M_{T-2},\ H_T\gamma_2 + v_T - c) \\
&\qquad + \sum_{t=T+1}^{\tau} \delta^{t-T} U(M_{T-2},\ H_T\gamma_2 + v_t)] \\
&\qquad g_{T-1}(u_{T-1},\ u_T) f_{T-1}(v_T, .\ ,\ v_\tau)\, du_{T-1}\, du_T \\
&\qquad dv_T, .. dv_\tau, \\
&\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\int_{-\infty}^{-G_{T-1}\gamma_1} [U(M_{T-2},\ H_T\delta_2 + v_T) \\
&\qquad + \sum_{t=T+1}^{\tau} \delta^{t-T} U(M_{T-2},\ H_T\gamma_2 + v_t)] \\
&\qquad g_{T-1}(u_{T-1}) f_{T-1}(v_T, \cdots v_\tau)\, du_{T-1} \\
&\qquad dv_T, \ldots dv_\tau
\end{aligned}
$$

+ δ max

where $g_{T-1}(u_{T-1}) = \int_{-\infty}^{\infty} g_{T-1}(u_{T-1},\ u_T)\, du_T$ and

$$f_{T-1}(v_{T-1}) = \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f_{T-1}(v_{T-1},\ v_T,\ ..,\ v_\tau)\, dv_T .. dv_\tau .$$

It should be noted that not all four combinations of maximals are viable as discussed in the text .

If a child is not born at T-1, expected lifetime utility is given by

(A.7) $E_{T-1}(LU_{T-1} | n_{T-1} = 0)$

$$= \int_{-\infty}^{\infty} [U(M_{T-2},\ H_{T-1}\gamma_2 + v_{T-1})] f_{T-1}(v_{T-1})\, dv_{T-1}$$

$$+ \delta \max \Bigg[\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \int_{-G_T\gamma_1}^{\infty} [U(M_{T-2}+1,\ H_T\gamma_2 + v_T - b - c) + \sum_{t=T+1}^{\tau} \delta^{t-T} U(M_{T-2}+1,\ H_t\gamma_2 + v_t)]\, g_{T-1}(u_T) f_{T-1}(v_T, .., v_\tau)\, du_T dv_T .. dv_\tau$$

$$+ \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \int_{-\infty}^{-G_T\gamma_1} [U(M_{T-2},\ H_T\gamma_2 + v_T - C) + \sum_{t=T+1}^{\tau} \delta^{t-T} U(M_{T-2},\ H_t\gamma_2 + v_t)]\, g_{T-1}(u_T) f_{T-1}(v_T, .., v_\tau)\, du_T dv_T .. dv_\tau ,$$

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} [U(M_{T-2},\ H_T\gamma_2 + v_T) + \sum_{t=T+1}^{\tau} \delta^{t-T} U(M_{T-2}, H_T\gamma_2 + v_t)]\, f_{T-1}(v_T, .., u_\tau)\, dv_T .. dv_\tau \Bigg]$$

where $g_{T-1}(u_T) = \int_{-\infty}^{\infty} g_{T-1}(u_{T-1},\ u_T)\, du_{T-1}$.

Denoting J_{T-1} as the difference in expected utilities with and without a birth at T-1, the decision rule at T-1 is

$$(A.8)\quad n_{T-1} = 1 \text{ iff } J_{T-1} = E_{T-1}(LU_{T-1} | n_{T-1}=1) - E_{T-1}(LU_{T-1} | n_{T-1} = 0) \geq 0$$
$$= 0 \text{ iff } J_{T-1} < 0$$

It may be verified that the conclusions reached in the text as to the comparative statics of the T-1 decision carry over directly to the more general case presented above.

Appendix B

Information on survival rates of infants (aged less than one) was available by state for the Malay population only from 1967 to 1972, 1974 and 1975. On a national basis, survival rates of infants were available from 1957 to 1970 for Malays while for the entire Malaysian population they were available back to 1948. Predicted survival rates were needed for Malays by state from 1940 to 1992, given the cohort span of the women in the sample. The following procedure was employed to form these estimates. Log national survival rates for Malays were regressed on log national survival rates for all Malaysia, time and its square, and a constant from 1957-1970. Log national survival rates for Malays were then predicted from this regression for the period 1948-1956. The log survival rate predicted from 1948-1956 and the log of the actual rate from 1957-1970 were then converted to log odds and regressed on time, its square and a constant. The log odds of the survival rate was then predicted for the period 1940-1947 and 1976-1992. The logistic formulation seemed to give more credible predictions both backward and forward in time. Each state level log survival rate was then regressed on the national level log survival rate for Malays, a time trend and a constant over the overlapping period, and predicted survival rates for each state were obtained for the period 1940-1964, 1973, and 1976-1992. The predicted survival rates combined with the actual survival rates for the available years were combined to form a state level Malay-specific survival rate series for the period 1940-1992. In the last stage, the log odds for each state based on this series was regressed on time, its square, and a constant. The survival probabilities predicted from these regressions were assumed to be

those predicted by the woman from the beginning of her decision making period. Since each woman's state of residence was known at every moment during the life-cycle the assignment was made on the basis of state of residence at the beginning of each eighteen month period. Perfect foresight about future residence is assumed. In the actual sample, child deaths within the first eighteen months occurred as follows: forty-seven women had one child death, eleven women had two child deaths, two women had three deaths and two women had four deaths. The actual mortality rate is about seven percent which is close to the same figure calculated from the state data for this sample, about six percent.

References

Ben-Porath, Yoram, "Fertility Response to Child Mortality: Micro Data from Israel," J.P.E. 84,No. 4, part 2 (August 1976).

_____ and Finis Welch, "Chance, Child Traits, and Choice of Family Size," publication no. R-117-NIH/RF. Santa Monica, Calif.: Rand, 1972.

Heckman, James J. "Statistical Models for Discrete Panel Data," mimeo, University of Chicago (January 1979).

_____ and Thomas MaCurdy, "A Life Cycle Model of Female Labor Supply," Review of Economic Studies 47,no. 146 (January 1980).

_____ and Robert J. Willis, "Estimation of a Stochastic Model of Reproduction: An Econometric Approach," in Household Production and Consumption, (ed.) N.E. Terleckyj, New York: Columbia University Press, 1975.

Hotz, V. Joseph, "A Life Cycle Model of Fertility and Married Women's Labor Supply," mimeo, Carnegie-Mellon University, October 1980.

Moffitt, Robert, "Life Cycle Profiles of Fertility: A State Dependent Multinomial Probit Model," mimeo, Rutgers College, August 1980.

Olsen, Randall J., "Estimating the Effects of Child Mortality on the Number of Births," Demography vol. 17, no. 4 (November 1980).

Sargent, Thomas J., "Estimation of Dynamic Labor Demand Schedules under Rational Expectations," J.P.E., 86, no. 6 (December 1978).

Schultz, T. Paul, "Interrelationships between Mortality and Fertility," in Population and Development, (ed.) Ronald G. Ridker, Baltimore: Johns Hopkins University Press, 1976.

Vallin, Jacques and Alain Lery, "Estimating the Increase in Fertility Consecutive to the Death of a Young Child," in The Effects of Infant and Child Mortality on Fertility, (ed.) Samuel H. Preston, New York: Academic Press, 1978.

Vijverberg, Wim, "Labor Supply and Fertility Decisions: A Dynamic Model of the Economic Behavior of Married Women," unpublished Ph.D. thesis, University of Pittsburgh, 1981.

Willis, Robert J., "A New Approach to the Economic Theory of Fertility Behavior," J.P.E. 81, no. 2, pt. 2 (March/April 1973).