EC 262P V. Joseph Hotz
Population Economics

Issues in Life Cycle Modeling of Fertility Behavior

1. Features of Life Cycle Models of Fertility and The Optimal Solution

Models include components from:

- (i) models of optimal life cycle consumption,
- (ii) models of life cycle labor supply decisions,
- (iii) models of human capital investment and accumulation,
- (iv) stochastic models of human reproduction.

1.1 Parental Preferences

Lifetime parental preferences considered in the literature takes the following form:

(1) $$U = \sum_{t=0}^{T} \beta^t u\left(a_t, \ell_t, s_t\right),$$

where

ℓ_t is mother's non-market, non-child care (leisure) time at age t,
s_t is parental consumption,
β is the couple's rate of time preference ($0 \leq \beta \leq 1$),

a_t, the flow of child services parents receive at age t from their stock of children, is governed by

(2) $$a_t = g(b_0, b_1, \ldots, b_{t-1}),$$

where $b_\tau = 1$ if the parents gave birth to a child when they were age τ ($\tau = 0,\ldots,t$-1) and $b_\tau = 0$ otherwise. Note a simplified version of a_t would be the couples' stock of children:

$$a_t = n_t = \sum_{\tau=0}^{t-1} b_\tau .$$

or could entail age-dependent service flow from "quality-adjusted" index of children, such as:

$$a_t = \sum_{\tau=0}^{t-1} \alpha_\tau b_\tau .$$

1.2 Maternal Time Constraints

Period-by-period constraints on the mother's time:

(3) $$\ell_t + h_t + t_{ct} = 1$$

where h_t is mother's labor supply and t_{ct} is time spent rearing children.

1.3 The Household's Budget Constraint

Existing life cycle models of fertility either assume capital markets either *perfect (PCM)* —i.e., parents are able to borrow and lend across time periods at a real interest rate, r_t—or *perfectly-imperfect (PICM)*, in which case no borrowing or saving is possible.

PCM: Parents' age t budget constraint can be expressed as:

(4) $$s_t + c_t + \boldsymbol{p}'_{et}\boldsymbol{e}_t = Y_{ht} + w_t h_t + S_t,$$

where S_t ($\equiv A_t$ - $A_{t\text{-}1}$) is savings, Y_{ht} denotes husband's income at age t, w_t is the wife's market wage rate, and $\boldsymbol{e}_t$ and $\boldsymbol{p}_{et}$ are vectors of contraceptive methods and their associated prices.

PICM:

(5) $$s_t + c_t + \boldsymbol{p}'_{et}\boldsymbol{e}_t = Y_{ht} + w_t h_t$$

Parents may face uncertainty about future income and prices. See model of Hotz and Miller (1986, 1993), where future realizations of husband's income, Y_{ht}, and the wife's wage rate, w_t, are treated as stochastic.

1.4 Production of Child Services or "Maintenance" of Children

Production of child services or maintenance of children involves parental time and market goods. Let c_t denote parental "costs" of maintaining existing stock of children:

(6) $$c_t = c(b_0, b_1, \ldots, b_{t-1}),$$

and parental (mother's) time:

(7) $$t_{ct} = t(b_0, b_1, \ldots, b_{t-1}),$$

Hotz and Miller (1989, 1993) use following specifications

(8) $$t_{ct} = \gamma \sum_{r=1}^{t-t_0} \delta^{r-1} b_{t-r}$$

(9) $$c_t = \psi \sum_{r=1}^{t-t_0} b_{t-r} = \psi n_t .$$

where

$$\begin{array}{l} t_{cr} = \gamma\delta^{r-1},\ \gamma>0,\ 0<\delta<1 \\ c_r = \psi,\ \psi>0 \end{array},$$

Note that this implies that the ratio of inputs,

$$c_r / t_{cr} = \psi / \gamma\delta^{r-1},$$

increases in a child's age.

1.5 Production of Children and Control of Fertility

"Reproduction" function given by:

$$b_t = R(\mathrm{e}_t, \phi_t) \tag{10}$$

where $\boldsymbol{e}_t$ denotes a K-dimensional vector of whose typical element, e_k, denotes whether or not the k^{th} contraceptive method is used, $k = 1,...,K$, and ϕ_t denotes the stochastic component governing the likelihood that a birth is produced with an unprotected sexual act.

Parents' *birth probability function* is given by:

$$p_{bt}\left(\mathrm{e}_t, \mu, \sigma_\phi^2\right) \equiv \Pr\left(b_t = 1 \middle| t, \mathrm{e}_t, \mu, \sigma_\phi^2\right) = E_\phi\left(R(\mathrm{e}_t, \phi_t)\right) \tag{11}$$

where $E_\phi(\cdot)$ denotes the expectations operator over the random variable ϕ_t, and μ and σ_ϕ^2 are the mean and variance, respectively, of ϕ_t.

The couple's μ can be interpreted as a couple's *fecundity*, where $\partial p_{bt}(\boldsymbol{e}_t,\mu,\sigma_\phi^2)/\partial\mu \geq 0$.

In addition to the pill, condoms, etc., three contraceptive methods are of particular analytic interest: (i) the use of no protection (which we indicate by e_{1t}); (ii) permanent sterilization, which precludes any (further) births; and (iii) an induced abortion (which we indicate by e_{Kt}) which, for our purposes, can be thought of as an *ex post* contraceptive method.

1.6 Maternal Investments in Human Capital

Treatment of the mother's wages: either exogenous or human capital production process, where wages are function of human capital and human capital is produced as function of past time investments.

Learning-by-Doing Specification:

$$w_t = H(w_{t-1}, h_t) - \delta_1 w_{t-1} - \delta_2 w_{t-1} 1[h_t = 0] \tag{12}$$

where $H(\cdot,\cdot)$ is the human capital production function, δ_1 and δ_2 are rates of depreciation ($0 \leq \delta_i \leq 1$, $i = 1,2$), and $1[\cdot]$ is the indicator function.

1.7 The Structure of the Solution to the Parents' Intertemporal Optimization Problem

Parents sequentially making choices over: (i) childbearing, or contraceptive methods, (ii) parental consumption and (iii) the allocation of the mother's time across labor market and childrearing activities so as to maximize (1) subject to the constraints implied by the relationships in (2), (3), (6), (7), and (10) [if reproduction is stochastic], (4) or (5) and, possibly (12).

The parents' optimization problem can be solved using techniques in the dynamic programming literature. See notes by Stern on structure of these models.

3. General Dynamic Structural Models

3.1. General Economic Structures

The primitives of a dynamic model are choices d_t, observed state variables X_t, unobserved state variables ε_t, taste shifters Z_t, a discount factor β, a utility function $U(\bullet)$, laws of motion, and an information set Ω_t. Let $U(X_t, d_t, \varepsilon_t; Z_t)$ be the utility function. The agent chooses $d^t = \{d_t, d_{t+1}, ..., d_T\}$ at time t given information set Ω_t to maximize

$$V_t(X_t, d_t, \varepsilon_t; Z_t) = E\left[\sum_{s=t}^{T} \beta^{s-t} U(X_s, d_s, \varepsilon_s; Z_s) \mid \Omega_t\right]$$

such that

$$\begin{aligned} X_{s+1} &= g(X_s, d_{s+1}, e_{s+1}); \\ Z_{s+1} &= h(Z_s, e_{s+1}); \\ \varepsilon_{s+1} &= k(Z_s, X_s, \varepsilon_s, e_{s+1}); \end{aligned}$$

and possibly a terminal condition $X_T = c$.

3.2. Examples

3.2.1. Berkovec and Stern

$X_t = \{$type of job, tenure$\}$
$d_t =$ job choice
$\varepsilon_t = \varepsilon$'s and μ's in the model
$Z_t =$ demographic characteristics
e_t are implicit in how ε's and μ's in the model change over time
$U(\bullet)$ is the flow each period
$g(\bullet)$ is deterministic (e_s has no effect) and depends upon where you were X_s and your choice d_s
$h(\bullet)$ is deterministic
$k(\bullet)$ is independent (ε_s has no effect), exogenous (X_s has no effect), and time-invariant (Z_s has no effect).

3.2.2. Dynamic Structural Fertility Models

(See, for example Wolpin or Hotz and Miller)
$X_t = \{$ages and number of children$\}$
$d_t =$ birth control choice
$\varepsilon_t =$ unobserved disutility of each birth control choice
$Z_t =$ demographic characteristics
e_t are implicit in realizations of when children are born
$U(\bullet)$ is the flow each period associated with number and ages of children and choice of birth control method
$g(\bullet)$ is random (birth control methods are not sure things) and depends upon where you were X_s and your choice d_s
$h(\bullet)$ is deterministic
$k(\bullet)$ is independent (ε_s has no effect) (is this reasonable?), exogenous (X_s has no effect), and time-invariant (Z_s has no effect).

3.2.3. Search Model

$X_t = \{$present offer, (possibly) search duration$\}$
$d_t =$ accept or reject
$\varepsilon_t =$ unobserved component of offer
$Z_t =$ demographic characteristics
e_t are implicit in realizations of new wage offers
$U(\bullet)$ is the negative of the cost of search plus the wage if it is accepted
$g(\bullet)$ is deterministic
$h(\bullet)$ is deterministic
$k(\bullet)$ is independent (ε_s has no effect) and exogenous (X_s has no effect).

3.2.4. Wolpin

$X_t = \{$SK (specific capital), GK (general capital), two wages$\}$
$d_t =$ accept or reject offer, stay at job or quit
$\varepsilon_t = u$ (period specific shock)
$Z_t =$ demographic characteristics
e_t are implicit in realizations of new wage offers and layoff events
$U(\bullet)$ is the value of leisure plus unemployment benefits if not working or the wage if it is accepted
$g(\bullet)$ is random because there are random layoffs and recalls

$h(\bullet)$ is deterministic
$k(\bullet)$ is independent (ε_s has no effect) and exogenous (X_s has no effect).

3.3. Value Function

The value function can be written as

$$\begin{aligned}
V_t(X_t, d_t, \varepsilon_t; Z_t) &= E\left[\sum_{s=t}^{T} \beta^{s-t} U(X_s, d_s, \varepsilon_s; Z_s) \mid \Omega_t\right] \\
&= U(X_t, d_t, \varepsilon_t; Z_t) + E\left[\max_{d^{t+1}} \sum_{s=t+1}^{T} \beta^{s-t} U(X_s, d_s, \varepsilon_s; Z_s) \mid \Omega_t\right] \\
&= U(X_t, d_t, \varepsilon_t; Z_t) + \beta E\left[\max_{d_{t+1}} V_{t+1}(X_{t+1}, d_{t+1}, \varepsilon_{t+1}; Z_{t+1}) \mid \Omega_t\right].
\end{aligned}$$

This is Bellman's equation.

3.4. Solution Techniques

There are three computational issues involved with solving stochastic dynamic programming models: a) evaluating $E_t V_{t+1}$, b) solving the value function for all possible paths, and c) solving the value functions for each sample person.

There are three approaches in the literature to the first problem. First, one can make a fortuitous assumption about the distribution of ε (as in Berkovec and Stern) so that $E_t V_{t+1}$ has a nice functional form. This either means having a very small number of choices (≤ 3) or using the Extreme Value distribution. Second, one can try to evaluate $E_t V_{t+1}$ numerically. Keane and Wolpin (1994), Holt (1997), and Brien, Lillard and Stern (1997) are three examples of this approach. Third, one can approximate $E\left[\max_{d_{t+1}} V_{t+1}\right]$ with $\max_{d_{t+1}} EV_{t+1}$. This is done in Stock and Wise (1990) and work by Manski and coauthors. Stern (1997) shows why this approach has poor properties.

There are two issues involved with the second problem. The first is dealing with the time dimension of the problem, and the second is dealing with the state space for each time period. For the first issue, there are two approaches. Rust (?) shows how to use a fixed point algorithm to solve infinite horizon problems. This is generally infeasible and unnecessary because, in fact, life is finite. The second approach is that described in the Berkovec and Stern notes: solve backwards recursively.[2]

The second issue occurs because the value function needs to be evaluated for each possible combination of state variables. Consider the fertility models. A minimal set of state variables would be duration of marriage and number and age of each child. Assuming marriage duration can take on, let's say, 20 values and we allow for a maximum of, let's say, 6 children up to 18 years of age, there are approximately[3] $20 \bullet 18^6 = 680,244,480$ combinations of state variables for each time period. If you add to that any continuous state variable such as an autoregressive unobserved component to choice specific utilities, then, in theory, it goes to ∞. One might reduce the number of state variable combinations by a) assuming marriage duration effects are constant after some duration, let's say 4 years; b) number of children effects are constant after, let's say, 3 children; and c) child age effects are constant after age 6. Then the number of combinations reduces to approximately $4 \bullet 6^3 = 864$. One could allow for a continuous state variable by discretizing it.

A second approach, used in Keane and Wolpin (1994), Smith (?), Stinebrickner (1997), and Holt (1997) is to evaluate the value function at only a subset of state variable combinations and approximate it at other points. This approach has promise but needs to be evaluated more formally.

A third approach, suggested by Hotz, Miller, Sanders, Smith (?), uses a non-linear transformation of the model and simulation to solve the problem. While clever and useful, it is limited to cases where there is no unobserved heterogeneity.

The last problem is that the value function needs to be evaluated for each person in the sample for each guess of the model parameters. Some researchers bite the bullet and do this. Others make very strong homegeneity assumptions so the value functions need be computed only once (or a very small number of times) for each guess of the parameters. Brien, Lillard and Stern (1997) use a weighting scheme to gain the benefits of the homogeneity assumptions without losing the benefits of allowing for heterogeneity across people.

3.5. Estimation

There are generically two steps involved in estimation. First, one may estimate exogenous processes g and h separately. Then one estimates the parameters asso-

[2] The finite horizon backwards recursion approach is a special case of Rust's fixed point approach.

[3] It is actually less if you assume no multiple births.

ciated with the choice process. Berkovec and Stern estimate death probabilities first, and the fertility models estimate fertility probabilities conditional on birth control method first. This two-step procedure is inefficient but usually very worthwhile in terms of computational cost.

All models estimate either $\Pr[d_t \mid \Omega_t]$, $\Pr[d^1 \mid \Omega_1]$, or $E[d_t \mid \Omega_t]$. Berkovec and Stern estimate $E[d_t \mid \Omega_t]$, the fertility models estimate $\Pr[d^1 \mid \Omega_1]$, and search models estimate $\Pr[d^1 \mid \Omega_1]$ and g simultaneously.

The two methods of estimation are MLE or (MSLE) and MOM (or MSM). In the generic problem, consider using MLE with $\Pr[d^1 \mid \Omega_1]$ to estimate the model. Then the MLE of θ is

$$\hat{\theta} = \arg\max_{\theta} \sum_{i=1}^{N} \log \Pr\left[d_i, X_i \mid Z_i, \Omega_{1_i}\right]$$

where d_i is the vector of observed choices made by observation i, X_i is vector of state variables, Z_i is the vector of (exogenous) taste shifters, and Ω_{1_i} is any initial conditions. The MOM estimator θ of is

$$\hat{\theta} = \arg\min_{\theta} \sum_{i=1}^{N} \left[\begin{pmatrix} d_i \\ X_i \end{pmatrix} - E\left(\begin{pmatrix} d_i \\ X_i \end{pmatrix} \mid \Omega_i\right)\right]' W_i W_i' \left[\begin{pmatrix} d_i \\ X_i \end{pmatrix} - E\left(\begin{pmatrix} d_i \\ X_i \end{pmatrix} \mid \Omega_i\right)\right]$$

where $E\left(\binom{d_i}{X_i} \mid \Omega_i\right)$ is the expected value of each element of $\binom{d_i}{X_i}$ conditional on the information available at the time $\binom{d_i}{X_i}$ is determined.

1.8 Comparative Dynamic Predictions from Models

Hard to come by. See Notes by Hotz and Sanders (2000), for some discussion of this, as well as Rosenzweig and Schultz (1985).

Predictions of interest are how wages, income and past fertility, contraceptive methods and contraceptive failures, affect

- optimal timing of first and subsequent births

  See references in Hotz, Klerman and Willis (1997).

- women's labor supply and human capital investment decisions

  See references in Hotz, Klerman and Willis (1997).

- relationship between fertility and labor supply and other decisions

  Later topic is focus of paper by Rosenzweig and Wolpin (JPE, 1980), Angrist and Evans (1998) and Rosenzweig and Wolpin (2001) for a critique of latter paper's strategy.

  See also papers in teenage childbearing literature, including Hotz, Sanders and McElroy (1997, forthcoming), Hotz, Mullin and Sanders (1997), and Mullin (2002). Also see notes by Hotz and Sanders (2000).