Dynamic Microeconomic Models of Fertility Choice: A Survey

Cristino R. Arroyo and Junsen Zhang

Journal of Population Economics, 1997, 10:23-65

James J. Heckman

THE UNIVERSITY OF CHICAGO

Econ 350, Winter 2021

1. Introduction and Motivation

2. Structural Models

2.1 General Framework

- A general framework nesting several well-known dynamic models may be given as follows.
- An individual *i* maximizes the (expected, discounted) value of utility over her life-cycle $t = \tau, \dots, T$:

$$\max E_\tau \sum_{t=\tau}^{T} \beta^t U\left(N_t,\, M_t,\, X_t,\, H_t,\, u_t,\, \theta\right) \qquad (1)$$

using all available information known as of τ.

Table 1. Specifications of period utility, various structural models

Heckman and Willis (1976):

$U = W(\psi N_t, X_t) - f(u_t)$

Wolpin (1984):

$U = W(M_t, X_t, \theta)$

$= (a_1 + \theta) M_t - a_2 M_t^2 + \beta_1 X_t - \beta_2 X_t^2 + \gamma M_t X_t$; γ any sign

Hotz and Miller (1984):

$U = W(M_t, Z_t)$; $Z_t = Z(H_t, X_t, \zeta_t)$

where Z_t is household production

ζ_t is a random error

Rosenzweig and Schultz (1985):

$U = W(N_t, M_t, X_t, H_t, \theta)$

$= \phi_1 N_t - \phi_2 N_t^2 + a_1 M_t - a_2 M_t^2 + \beta_1(\theta) X_t - \beta_2 X_t^2 + \delta_1 H_t - \delta_2 H_t^2 + \gamma H_t M_t$; γ any sign

Newman (1988):

$U = W(M_t, X_t, u_t)$

$= a_1 M_t - a_2 M_t^2 + \beta_1 X_t - \beta_2 X_t^2 + \gamma M_t X_t + \rho_1 u_t - \rho_2 u_t^2$; γ any sign

Leung (1991):

$U = W(M_t, X_t) - f(\pi(u_t))$

where $\pi(.)$ is the probability of a birth

Note: All parameters above, unless specified, are positive. All variables are described in the text of Sect. 2.1.

Maximization is subject to a sequence of budget constraints for each period t:

$$I_t + w_t(\bar{H} - H_t) = X_t + p_t^M M_t + p_t^u u_t, \tag{2}$$

where I_t is current (husband's) income, w_t is the individual's market wage, $\bar{H}$ is the total amount of time available for work, so that $(\bar{H} - H_t)$ is the mother's labor market time.

The individual's choice variables are, generally, consumption X_t, non-work time H_t, and contraceptive efficiency u_t, while family size evolves according to

$$M_{t+1} = M_t + N_t, \tag{3}$$

where N_t is the number (zero or one) of surviving newborn children in period t.

- Whether a net birth occurs or not typically depends on some stochastic birth and death processes, as well as the level of contraceptive efficiency.

- In general, one has

$$N_t = N\left(\pi^{b}(1 - u_t), \pi^{\mathrm{m}}\right), \tag{4}$$

where π^{b} is the probability of a birth assuming no contraceptive control (i.e., an individual's natural fecundity) and π^{m} is infant mortality risk.

2.2 Predictions of the General Framework

- That is, if N_t^*, X_t^*, H_t^* and u_t^* are the solution values for the maximization, then

$$V\left(M_t;\, t,\, \pi^b, \pi^{\mathrm{m}},\, I_t,\, w_t,\, p_t^M,\, p_t^u,\, \beta,\, \theta\right)$$

$$= E_\tau \sum_{t=\tau}^{T} \beta^t U\left(N_t^*,\, M_t,\, X_t^*,\, H_t^*,\, u_t^*,\, \theta\right). \quad (5)$$

Suppressing other arguments in *V* save *M* and *t*, Bellman's optimality principle allows us to rewrite *V* as

$$
\begin{aligned}
V(M_t;\, t) = \; & \max\{U(M_t,\, X_t,\, H_t,\, u_t,\, \theta) + \beta E_t V(M_{t+1};\, t+1)\} \\
= \; & \max\{U(M_t,\, X_t,\, H_t,\, u_t,\, \theta) + \beta\,[\pi_t V(M_t + 1;\, t+1) \\
& + (1-\pi_t)\, V(M_t;\, t+1)]\}. \qquad (6)
\end{aligned}
$$

Treating M_t as fixed, the partial derivative of the right-hand-side with respect to u_t gives the optimal contraception rule:

$$-U_{X_t} p_t^u + U_{u_t} - \beta \pi^b [V(M_t + 1; t + 1) - V(M_t; t + 1)] = 0$$

or

$$U_{X_t} p_t^u - U_{u_t} = \beta \pi^b [V(M_t; t + 1) - V(M_t + 1; t + 1)]. \tag{7}$$

On the other hand, the graph of the expected marginal benefits of contraception is a flat curve in u, whose exact sign and location is determined by the sign and value of

$$\Delta V\left(M_{t+1};\, t+1\right) = V\left(M_t;\, t+1\right) - V\left(M_t + 1;\, t+1\right), \tag{8}$$

which is the capitalized value (at t+1) of preventing a birth at t, given parity M_t.

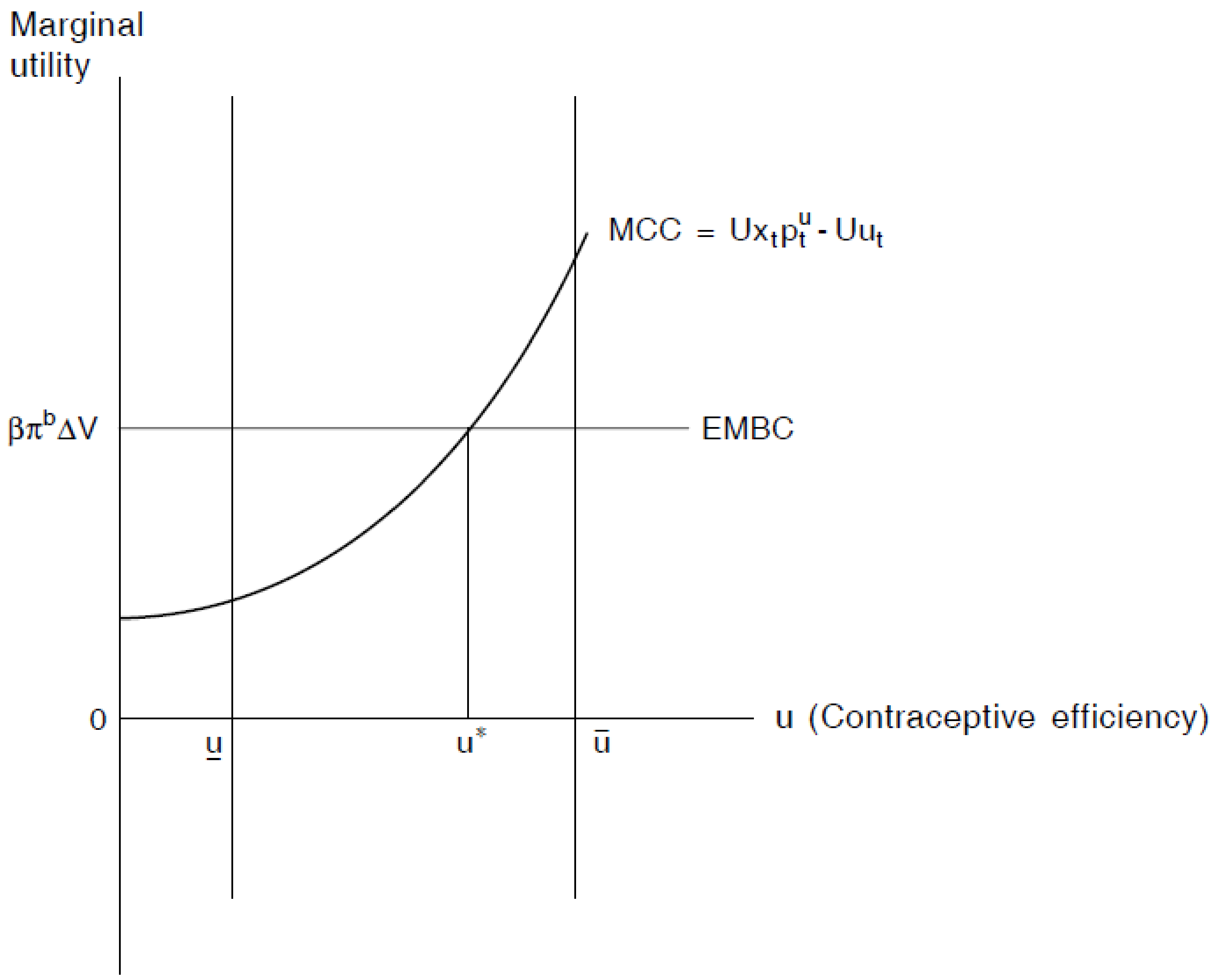

Fig. 1. Marginal costs of contraception, expected marginal benefits of contraception, and optimal contraceptive efficiency

2.3 Vijverberg's (1984) Model

2.4 Direct Estimation of Structural Parameters

Under the functional form of U assumed by Wolpin (see Table 1), the decision to add a child or not can be shown to depend upon the function:

$$J_t = E_t\left[V\left(M_t + 1;\, t+1\right) - V\left(M_t;\, t+1\right)\right] + P_t\,\theta_t$$
$$= -E_t\left[\Delta V\left(M_{t+1};\, t+1\right)\right] + P_t\,\theta_t, \qquad (9)$$

where $E_t\,\Delta V\left(M_{t+1};\, t+1\right)$ denotes the capitalized value (at t+1) of preventing a birth at t.

The optimal decision rule for births is therefore:

$$N_t^* \begin{cases} = & 1 \text{ iff } J_t > 0 \\ = & 0 \text{ iff } J_t \leq 0. \end{cases} \tag{10}$$

$$\theta_t^* = E_t\left[\Delta V\left(M_{t+1};\, t+1\right)\right]/P_t. \tag{11}$$

$$Pr\left[N_t = 1 | M_t\right] = 1 - \Phi\left(\theta_t^* / \sigma\right), \tag{12}$$

where σ is the standard error of θ_t^* and $\Phi(\mathrm{x})$ is the value at x of the standard normal cumulative density function. Analogously, the conditional probability of no birth at t is

$$Pr\left[N_t = 0 | M_t\right] = \Phi\left(\theta_t^* / \sigma\right). \tag{13}$$

Let Ω be the set of time periods where there is a birth, and Ω^c its complement. *For individual i*, the likelihood L^i of any particular birth pattern is

$$L^i = \prod_{t \in \Omega} Pr\left[N_t = 1 | M_t\right] \prod_{t \in \Omega^c} Pr\left[N_t = 0 | M_t\right], \tag{14}$$

where it is understood that Ω, Ω^c, N_t, M_t, *and* θ_t^* all depend on i, and that L^i, in general, is a function of parameters α via θ_t^*.

Given a sample of I individuals, the sample likelihood is

$$L = \prod_{i=1}^{I} L^i, \tag{15}$$

which is maximized with respect to α.

The conditional expectation at t of $V(M_{t+1}; t+1)$ is

$$E_t V(M_{t+1};\ t+1) = \int\limits_{\varepsilon} V(M_{t+1};\ t+1) f(\varepsilon_{t+1} | M_t)\, d\varepsilon_{t+1}. \tag{16}$$

Indeed, one must solve the entire dynamic program, which at the time required backward solution from period T, following Bellman's recursion:

$$
\begin{aligned}
E_t V(M_{t+1}; t+1) &= E_t \{\max [U(\cdot; M_t) + \beta E_{t+1} V(M_{t+2}; t+2)]\} \\
&= E_t \{\max [U(\cdot; M_t) + \beta E_{t+1} \{\max U(\cdot; M_{t+1}) \\
&\quad + \beta E_{t+2} [V(M_{t+3}; t+3)]\}]\} \qquad (17)
\end{aligned}
$$

and so on until $t+j=T$.

$$E_t V(M_{t+1}; t+1) = E_t \max_k V_k(M_{t+1}; t+1). \tag{18}$$

Here $k = 1, ..., K$ indexes the discrete-valued and mutually exclusive alternatives the individual chooses from

- Keane and Wolpin (1994) call these k functions the *alternative-specific value functions.*

- These functions satisfy

$$V_k(M_{t+1}; t+1) = U_k(\cdot; M_{t+1}) + \beta E_{t+1} V(M_{t+2}; t+2); \; k = 1, \ldots, K. \quad (19)$$

The general form of the approximating function proposed by Keane and Wolpin is

$$EMAX\left(M_t,\, t\right) = MAXE\left(M_t,\, t\right) + g\left(MAXE\left(M_t,\, t\right) - \bar{V}_k\right)\,, \qquad (20)$$

where $\text{MAXE}(M_{t,} t) = \max_k \quad \bar{V}_k, \quad \bar{V}_k = E_t V_k(M_{t+1}; t+1).$

In Monte Carlo experiments, Keane and Wolpin (1994) find that the following form of g worked well:

$$EMAX(M_t, t) - MAXE(M_t, t) = \delta_0 + \sum_{j=1}^{K} \delta_{1j} (MAXE - \bar{V}_j) + \sum_{j=1}^{K} \delta_{2j} (MAXE - \bar{V}_j)^{1/2} \quad (21)$$

In the context of that example, Geweke et al. propose approximating J_t by the translated logistic of a polynomial-in-M_t:

$$\tilde{J}_t = exp\left(r^p\left(M_t\right)'\delta\right)/\left[1 + exp\left(r^p\left(M_t\right)'\delta\right)\right] - \frac{1}{2}\ , \qquad (22)$$

where $r^P(M_t)$ is a vector of variables of the form $\prod_{i=1}^{p}(M_{it})^p$ so that $(r^p(M_t)'\delta)$ is the pth-order polynomial function in the state M_t with coefficient δ.

To estimate δ, the authors suggest maximum likelihood estimation of the logit model.

$$Pr\left[N_t = 1 | M - t\right] = exp\left(r^p\left(M_t\right)'\delta\right) / \left[1 + exp\left(r^p\left(M_t\right)'\delta\right)\right] \tag{23}$$

based on R simulations of $\{M_t\}_{t=1}^{T}$ and S possible values of the choice variable. Estimates $\hat{\gamma}$ can now be used to construct the approximate decision rule J

$$\hat{J}_t = exp\left(r^p\left(M_t\right)'\hat{\delta}\right) / \left[1 + exp\left(r^p\left(M_t\right)'\hat{\delta}\right)\right] - \frac{1}{2}\ . \tag{24}$$

3. Reduced-form Models

3.1 Hazard-rate Analysis (Heckman and Walker, etc.)

If a woman becomes at risk for the jth birth at time $\tau(j-1)$, the conditional hazard at duration t_j is defined to be

$$h_j\left(t_j \mid H\left(\tau\left(j-1\right)+t_j\right)\right) . \tag{25}$$

Assuming that T_j is absolutely continuous given H, we may integrate (25) to find the survivor function

$$S\left(t_j \mid H\left(\tau\left(j-1\right)+t_j\right)\right) = exp\left[-\int_0^{t_j} h_j\left(u \mid H\left(\tau\left(j-1\right)+u\right)\right) du\right] . \tag{26}$$

A woman at risk for a first birth at $\tau = 0$ continues childless a random length of time governed by the survivor function

$$Pr\left(T_1 > t_1 | H\left(\tau\left(0\right) + t_1\right)\right) = exp\left[-\int_0^{t_1} h_j(u|H(\tau(k-1)+u))\,du\right]. \qquad (27)$$

At time $\tau(1)$, the woman conceives and moves to the state $Y(\tau)=1$. In the general case, $Y(\tau)=k–1$ for $\tau(k–1)\leq\tau<\tau(k)$ and $T_k=\tau(k)–\tau(k–1)$ is governed by the conditional survivor function

$$Pr\left(T_k > t_k | H(\tau(k-1)+t_k)\right) = exp\left[-\int_0^{t_k} h_k(u|H(\tau(k-1)+u))\,du\right]. \qquad (28)$$

The conditional density function of duration $T_k=t_k$ is given by the product of the hazard and survivor functions

$$g(t_k|H(\tau(k-1)+t_k)) = h_k(t_k|H(\tau(k-1)+t_k)) \cdot S(t_k|H(\tau(k-1)+t_k)). \quad (29)$$

In their empirical specification, Heckman and Walker approximate the jth conditional hazard using the following functional form

$$h_j(t_j | H(\tau(j-1) + t_j); \theta) = exp\left[\gamma_{0j} + \sum_{k=1}^{K_j} \gamma_{kj} \left(\frac{t_j^{\lambda_{kj}} - 1}{\lambda_{kj}} \right) + \mathbf{Z}\beta_j + f_j\theta \right], \quad (32)$$

where $\mathbf{Z}$ includes all observed regressors possibly including durations from previous spells and spline functions of current durations.

Finally, Heckman and Walker's empirical framework allows for period-specific stopping behavior. The survivor function for the jth birth is

$$S_j\left(t_j \mid H\left(\tau\left(j-1\right)+t_j\right);\theta\right) = P^{(j-1)} + \left(1-P^{(j-1)}\right) exp\left[-\int_0^{t_j} h_j\left(u \mid H\left(\tau\left(j-1\right)+u\right);\theta\right) du\right], \quad (33)$$

where $P^{(j-1)}$ is the probability that a woman with j–1 children is never at risk to have the jth birth and thus captures permanent biological or behavioral sterility (i.e. a parity-specific mover-stayer mixture distribution).

The contribution to sample likelihood of a woman with fertility history $T_1 = t_1, T_2 = t_2, \ldots, T_k = t_k$ and an incomplete $k+1$ spell of length $\bar{t}_{k+1}$ is

$$\sum_{j=1}^{I} \prod_{j=1}^{k} \left[-\frac{\partial \ln S_j \left(t_j | H\left(\tau\left(j-1\right)+t_j\right); \theta_i\right)}{\partial t_j} \right] \cdot S_j \left(t_j | H\left(\tau\left(j-1\right)+t_j\right); \theta_i\right) \cdot S_{k+1}\left(\bar{t}_{k+1} | H\left(\tau\left(j-1\right)+t_j\right); \theta_i\right) p_i \,. \quad (34)$$

Table 2. Empirical results from reduced-form models

Authors (Year)	Emphasized (or primary) covariates	Other or additional covariates	Dataset
Heckman, Hotz and Walker (1985)	Age at marriage, previous birth intervals	Current spell duration, labor participation, schooling, urban, white-collar, attended university	1981 Swedish Fertility Survey
Heckman and Walker (1990 a, b, 1991)	Female wage and male income	Current spell duration, urban, white-collar, age, time trend, attended university, ever-married, unemployment rate, policy measures	1981 Swedish Fertility Survey
Tasiran (1995)	Female wage, male income, female schooling, and working experience	Current spell duration, urban, white, age at union start, birth cohorts, ever-married	1981 Swedish Fertility Survey, Swedish 1984 and 1988 Household Market and Non-Market Activities, 1985–1988 PSID

Table 2. Empirical results from reduced-form models

Authors (Year)	Emphasized (or primary) covariates	Other or additional covariates	Dataset
Newman and McCulloch (1984)	Male and female years of schooling	Female's birth year, child mortality, family planning measure	1976 Costa Rica National Fertility Survey
David and Mroz (1989)	Age of husband and wife, previous child's sex, number of boys and girls alive, number of boys and girls alive aged 10 or older	Number of prior deaths of boys and girls before the 3rd month, number of prior deaths of boys and girls aged 3–11 months, village mortality	Rural France data relating to the marriage cohorts of 1749–1789

3.2 Quasi-maximum Likelihood Estimates of Linear Decision Rules

Hotz and Miller now specify q_{it} as

$$q_{it} = v_0 + v_1 I_{it} + v_2 c_{it} + v_3 \mathrm{a}_{it} + \sum_{k=1}^{t} v_{4k} N_{it-k} + v_5 t_i + v_6 \theta_i + \varepsilon_{it}, \qquad (35)$$

where v_j, $j = 1, \ldots, 6$ are linear coefficients, some of which can be separately identified and estimated.

Hotz and Miller's (1988) framework:

$$\pi_{it} \begin{cases} = & \bar{\pi} \text{ if } q_{it} \geq 0 \\ = & \underline{\pi} \text{ if } q_{it} < 0, \end{cases} \tag{36}$$

where $\bar{\pi}$ *and* $\underline{\pi}$ are, respectively, the upper and lower values of the birth hazard.[31]

Hotz and Miller posit that total expenditure a_{it}, and total maternal care time c_{it} are linear sums of the mother's birth history $\{N_{it-k}\}_{k=1}^{t}$:

$$\mathrm{a}_{it} = \sum_{k=1}^{t} \mathrm{a}_k N_{it-k} \tag{37}$$

$$c_{it} = \sum_{k=1}^{t} c_k N_{it-k}. \tag{38}$$

Using (37)–(38), one can rewrite the index function as

$$q_{it} = \tilde{v}_0 + \tilde{v}_1 I_{it} + \sum_{k=1}^{t} (\tilde{v}_2 c_k + \tilde{v}_3 \mathrm{a}_k + \tilde{v}_{4k}) N_{it-k} + \tilde{v}_5 t_i + \tilde{v}_6 \theta_i + \tilde{\varepsilon}_{it}, \tag{39}$$

where $\tilde{\varepsilon}_{it}$ is now a $N(0,1)$ error and the parameter vector $\tilde{v}$ is just the normalization $\tilde{v} = \sqrt{\sigma} v$, where v is the vector of identifiable parameters $(v_0, v_1, \{v_2 c_k + v_3 a_k + v_{4k}\}_{k=1}^{t}, v_5, v_6)$ in the original index function (35).

This is

$$
\begin{aligned}
&\prod_t Pr\left(N_{it}|I_{it},\, \{N_{it-k}\}_{k=1}^{t},\, \theta_i\right) \\
&= \prod_t \{N_{it}\left[\underline{\pi} + (\bar{\pi} - \underline{\pi}) Pr\left(q_{it} \geq 0 | I_{it}\,,\{N_{it-k}\}_{k=1}^{t}, \theta_i\right)\right] \\
&\quad + (1 - N_{it})\left[1 - \underline{\pi} - (\bar{\pi} - \underline{\pi})\, Pr\left(q_{it} \geq 0 | I_{it},\, \{N_{it-k}\}_{k=1}^{t},\, \theta_i\right)\right]\} \\
&= \prod_t \{(1 - N_{it}) + (2\,N_{it} - 1)\left[\underline{\pi} + (\bar{\pi} - \underline{\pi})\, \Phi\left(\tilde{v}_{it}\right)\right]\}\ , \qquad (40)
\end{aligned}
$$

where ϕ is the standard normal cumulative distribution function and the product $\prod_t$ ranges over the entire lifetime of the ith individual, that is t goes from $\underline{t}_i$ to t_i .

Now form the following *quasi*-likelihood function

$$Q\left(\underline{\pi},\, \bar{\pi},\, \tilde{v},\, \tilde{v}^{+},\, \theta_i\right) = \sum_i Q_i\left(\underline{\pi},\, \bar{\pi},\, \tilde{v},\, \tilde{v}^{+},\, \theta_i\right)$$

$$= \sum_i \left[L_{i1}(\underline{\pi},\, \bar{\pi},\, \tilde{v},\, \theta_i) + \sum_p L_{ip}(\underline{\pi},\, \bar{\pi},\, \tilde{v},\, \tilde{v}^{+},\, \theta_i)\right], \quad (41)$$

where $\tilde{v}^{+}$ are any extra parameters appearing in the p equations L_p.

Given , the parameter estimates that maximize the quasi-likelihood function Q are the solutions of

$$\sum_i \mathrm{m}_i(z,\, \theta_i) = \sum_i \partial L_{i1}(z^1,\, \theta_i)/\partial z^1 + \sum_i \sum_p L_{ip}(z,\, \theta_i)/\partial z = 0 \ . \qquad (42)$$

$$J_I = \left[(1/I)\sum_i \mathrm{m}_i\,(z,\, \theta_i)\right]' W_I \left[(1/I)\sum_i \mathrm{m}_i\,(z,\, \theta_i)\right] . \qquad (43)$$

The consistent estimate $\hat{z}$ of z based on (43), is $\hat{z}$=argmin J_I.

4. Other Models and Approaches

5. Conclusions