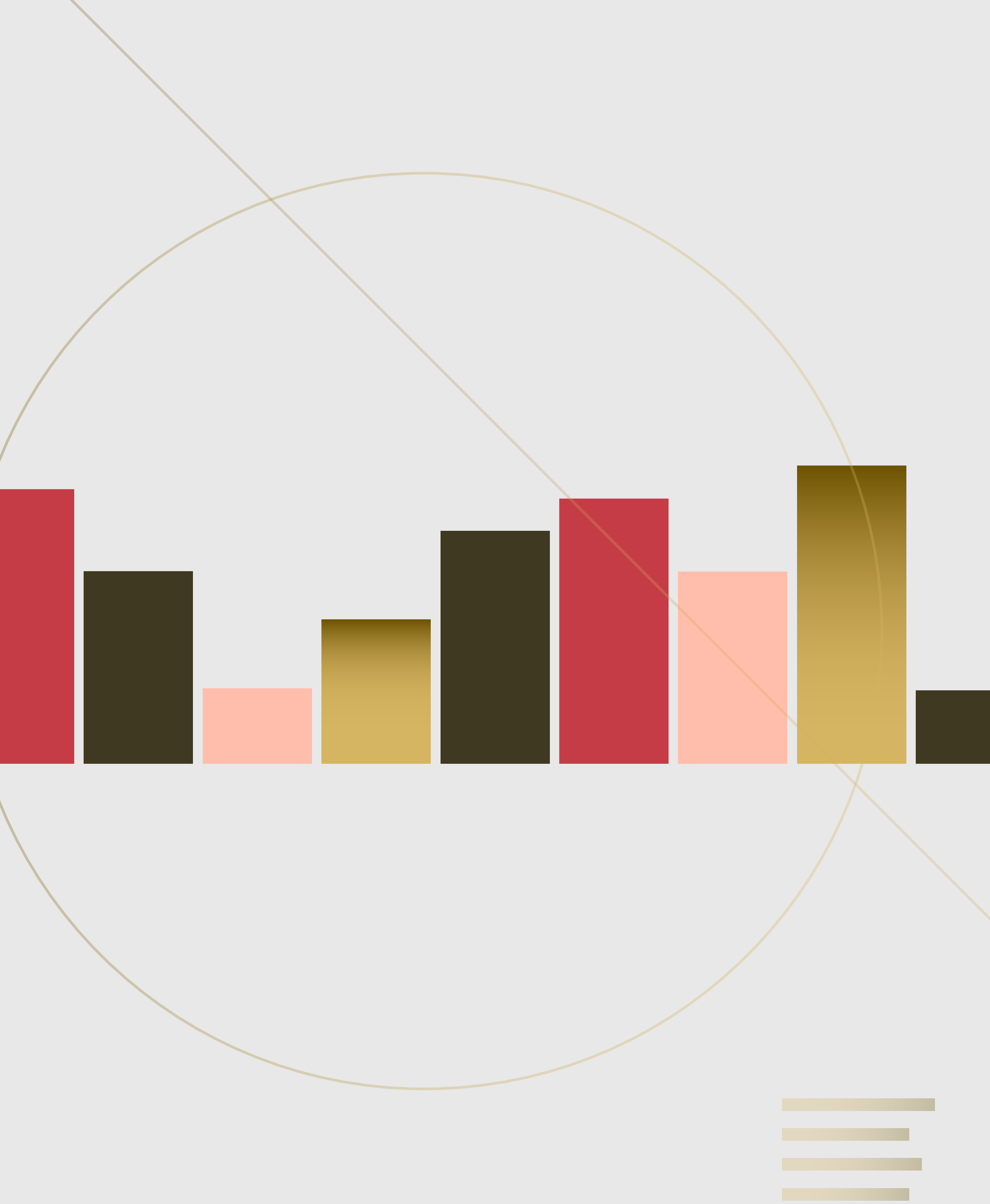

Inference in semi parametric Dynamic models for binary longitudanal data

- Chib and Jeliazkov (2006)

Journal of American Statistical Association

$$y_{it} = 1\{\tilde{x}_{it}'\delta + w_{it}'\beta_i + g(s_{it}) + \phi_1 y_{i,t-1} + \phi_2 y_{i,t-2} \cdots \phi_j y_{i,t-j} + \varepsilon_{it} > 0\}$$

Prior for non parametric function $g(\cdot)$

Let $S_{N\times 1}$, m is the unique numbers of elements in S i.e $m \leq N$.

order the unique elements in ascending order $v_1 < v_2 \cdots < v_m$.

$\therefore$ $g = (g(v_1), g(v_2), \cdots g(v_m)') = (g_1, g_2 \cdots g_m)$

Second order random walk specification

$$g_t = \left(1 + \frac{h_t}{h_{t-1}}\right) g_{t-1} - \frac{h_t}{h_{t-1}} g_{t-2} + u_t \qquad u_t \sim N(0, \tau^2 \underbrace{h_t}_{}) \quad \underbrace{h_t = v_t - v_{t-1}}$$

variance is dependence upon the distance b/w 2 unique values of s.

Let $\begin{pmatrix} g_1 \\ g_2 \end{pmatrix} \Big| \tau^2 \sim N\left(\begin{pmatrix} g_{10} \\ g_{20} \end{pmatrix}, \tau^2 G_0\right)$

- initializing condition for the prior

Stacking up over t, the sequence is defined as

$$
\begin{aligned}
g_{10} & \\
g_{20} & \\
g_3 &= \left(1+\frac{h_3}{h_2}\right) g_{20} - \frac{h_3}{h_2} g_{10} + u_3 \\
\vdots & \qquad\qquad \vdots \qquad\qquad\qquad \vdots \\
g_T &= \left(1+\frac{h_T}{h_{T-1}}\right) g_{T-1} - \frac{h_T}{h_{T-1}} g_{T-2} + u_T
\end{aligned}
$$

$$
\begin{bmatrix}
1 & 0 & \cdots & \cdots & \cdots & 0 \\
0 & 1 & \cdots & \cdots & \cdots & 0 \\
\frac{h_3}{h_2} & -\left(1+\frac{h_3}{h_2}\right) & 1 & 0 & \cdots & 0 \\
\vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\
\cdots & \cdots & \cdots & \frac{h_T}{h_{T-1}} & -\left(1+\frac{h_T}{h_{T-1}}\right) & 1
\end{bmatrix}
\begin{bmatrix} g_1 \\ g_2 \\ g_3 \\ \vdots \\ g_T \end{bmatrix}
=
\begin{bmatrix} g_{10} \\ g_{20} \\ u_3 \\ \vdots \\ u_T \end{bmatrix}
$$

$$
\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \sim N\left(\begin{matrix} 0 \\ 0 \end{matrix}, G\tau^2 \right)
$$

$$
\begin{aligned}
H g &= u \\
g &= H^{-1} u
\end{aligned}
$$

$$
E(g) = H^{-1} \begin{pmatrix} g_{10} \\ g_{20} \\ 0 \\ \vdots \\ 0 \end{pmatrix}
$$

$$Var(g) = \tau^2 H^{-1} \underbrace{\begin{pmatrix} G_0 & & & \\ & h_3 & & \\ & & \ddots & \\ & & & h_m \end{pmatrix}}_{\Sigma} H^{-1}$$

$$g \mid \tau^2 \sim N\left(H^{-1} \begin{pmatrix} g_{10} \\ g_{20} \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \underbrace{\tau^2 H^{-1} \Sigma H^{-1}}_{K} \right)$$

K is banded to easy for computation.

Prior for $\tau^2 \sim IG(v_0/2, \delta_0/2)$

Priors on the linear effect

Now we well focus on parametric facet.

$$z_{it} = x_{it}'\delta + w_{it}'\beta_i + g(s_{it}) + \phi_1 1\{z_{i,t-1} > 0\} + \dots + \phi_J 1\{z_{i,t-J} > 0\} + \varepsilon_{it}$$

$$y_{it} = 1\{z_{it} > 0\}$$

$$\varepsilon_{it} = \rho_1 \varepsilon_{i,t-1} + \dots + \rho_p \varepsilon_{i,t-p} + v_{it}$$

Stacking over t

$$\underset{T\times 1}{z_i} = \underset{T\times M}{\tilde{X}_i} \underset{M\times 1}{\delta} + \underset{T\times S}{W_i} \underset{S\times 1}{\beta_i} + \underset{T\times 1}{g_i} + \underset{T\times 1}{L_i}\phi + \underset{T\times 1}{\varepsilon_i}$$

Assume $\underset{5\times 1}{\beta_i} = \underset{5\times 3}{A_i}\underset{3\times 5}{\gamma} + b_i$, $b_i \sim N(0, D)$

In simplest case where W_i does not have intercept and β_i is independent of other independent variables, we assume that β_i depends on initial values y_{i0}

$$A_i = \begin{bmatrix} 1 \ \bar{y}_{i0}' & & \\ & 1 \ \bar{y}_{i0}' & \ddots \\ & & 1 \ \bar{y}_{i0}' \end{bmatrix}$$

we can include other independent variables in A_i if we suspect that β_i is related to them.

$$A_i = \begin{bmatrix} 1 \ \underbrace{\bar{y}_{i0}}_{\text{mean}} \ \underbrace{\bar{r}_{i1}'}_{\text{mean}} & & \\ & \ddots & \\ & & 1 \ \bar{y}_{i0} \ \bar{r}_{iq} \end{bmatrix}$$

Substituting assumption of β_i in z_i

$$z_i = \tilde{X}_i \delta + W_i \beta_i + g_i + L_i \phi + \varepsilon_i$$

$$z_i = \tilde{X}_i \delta + w_i (A_i \gamma + b_i) + g_i + L_i \phi + \varepsilon_i$$
$$z_i = \tilde{X}_i \delta + w_i A_i \gamma + g_i + L_i \phi + w_i b_i + \varepsilon_i$$

$$z_i = [\tilde{X}_i \quad w_i A_i \quad L_i] \begin{bmatrix} \delta \\ \gamma \\ \phi \end{bmatrix} + g_i + [w_i b_i + \varepsilon_i]$$

$$z_i = X_i\beta + g_i + W_i b_i + \varepsilon_i$$

- Basis for likelihood

$$\beta \sim N(\beta, B_0)$$
$$D^{-1} \sim W(r_0, R_0)$$
$$\rho \sim N(\rho_0, P_0)\, I_{S_\rho}$$

Estimation Algorithm

$$\varepsilon_i \sim N(0, \Omega_i)$$

Discussing the construction of Ω_i

Let $\varphi_j = E(\varepsilon_{it}, \varepsilon_{it+j}) = \varphi_{-j}$ (j^{th} autocovariance)

$$\varphi_j = \rho_1 \varphi_{j-1} + \rho_2 \varphi_{j-2} \cdots \rho_p \varphi_{j-p}$$ [same p^{th} order difference equation as the process itself]

The first p values $(\varphi_0, \cdots \varphi_{p-1})$ are given by the first column of

$$[I - F \otimes F]^{-1}_{p^2 \times p^2}$$

where
$$F = \underset{p \times p}{\begin{bmatrix} \rho'_{1\times p} & \\ \underset{(p-1)\times(p-1)}{I_{p-1}} & \underset{(p-1)\times 1}{0_{(p-1)\times 1}} \end{bmatrix}} = \begin{bmatrix} \rho_1 & \rho_2 & \cdots & \rho_p \\ 1 & & 0 & 0 \\ & 1 & \ddots & 0 \\ & & \ddots & \vdots \\ 0 & & 1 & 0 \end{bmatrix}_{p \times p}$$

$\therefore \quad \Omega[j,k] = \psi_{j-k}$

For example AR(1)

$$\varepsilon_t = \rho\, \varepsilon_{t-1} + v_t$$

$$E(\varepsilon_t, \varepsilon_{t-j}) = E(\rho^j \varepsilon_{t-j} + \rho^{j-1} v_{t-j+1} \cdots \rho v_{t-1} + v_t, \varepsilon_{t-j})$$

$$= \rho^j E(\varepsilon_{t-j})^2$$

$$\operatorname{var} \varepsilon_t = \operatorname{var}(\rho\, \varepsilon_{t-1}) + \operatorname{var}(v_t)$$

$$\sigma^2_\varepsilon = \rho^2 \sigma^2_\varepsilon + 1$$

$$\sigma^2_\varepsilon = \frac{1}{1-\rho^2}$$

$$E(\varepsilon_t, \varepsilon_{t-j}) = \rho^j / 1-\rho^2$$

Now let's focus on the main equation

$$z_i = X_i \beta + g_i + u_i \qquad (u_i = W_i b_i + \varepsilon_i)$$

$$z_i \sim N(X_i \beta + g_i, V_i) \text{ where } V_i = \Omega_i + W_i D W_i'$$

$$\Pr(y_i \mid \beta, g_i, D, \rho) = \int_{B_{iT}} \cdots \int_{B_{i1}} N(z_i \mid X\beta + g_i, V_i)\, dz_i$$

Now moving back to the earlier representation of z_i

$$z_i = x_i \beta + g_i + W_i b_i + \varepsilon_i \quad , \varepsilon_i \sim N(0, \Omega_i)$$

Stacking over i

$$z = X\beta + Qg + Wb + \varepsilon \quad , \varepsilon \sim N(0, \Omega)$$

$$X = \begin{bmatrix} X_1 \\ \vdots \\ X_N \end{bmatrix} , \quad W = \begin{bmatrix} W_1 & & \\ & \ddots & \\ & & W_n \end{bmatrix}$$

$$\Omega = \begin{bmatrix} \Omega_1 & & \\ & \ddots & \\ & & \Omega_n \end{bmatrix} , \quad Q = \text{incidence matrix}$$

Case 1 : Uncorrelated errors. $(\Omega = I)$

$$g \mid y, \beta \cdots z \sim N(\hat{g}, G)$$

$$\hat{g} = G\left(K g_0 / \tau^2 + Q'(z - X\beta - Wb)\right)$$
$$G = (K/\tau^2 + Q'Q)^{-1}$$

G^{-1} is banded

Let $P'P = G^{-1}$ (cholesky decomposition)

1) Sample $u \sim N(0, I)$
2) Solve $Px = u$ (back substitution)

$\therefore x \sim N(0, G)$

3) $x + \hat{g} \sim N(\hat{g}, G)$

$\hat{g}$ can be obtained by solving

$P'P\hat{g} = (Kg_0/\tau^2 + Q'(z - X\beta - Wb))$

by forward & back substitution

Case 2: Correlated errors ($\Omega_i \neq I$)

$$g \mid y, \beta, b, \tau^2, z, \Omega \sim N(\hat{g}, G)$$

$$\hat{g} = G\,(Kg_0/\tau^2 + Q'(z - X\beta - Wb))$$
$$G = (K/\tau^2 + Q'\Omega^{-1}Q)^{-1}$$

Decompose $\Omega_i = \underbrace{R_i}_{\text{symmetric positive definite}} + \kappa I$, $\kappa = \min\{\lambda_{ij}\}/2$

λ_{ij} = eigen value

Let $C_i'C_i = R_i$ (cholesky decomposition)

$\therefore \Omega_i = C_i'C_i + \kappa I$

$$z_i = X_i\beta + W_ib_i + g_i + \varepsilon_i$$
$$= X_i\beta + W_ib_i + g_i + C_i'u_i + \xi_i$$

$u_i \sim N(0, I)$, $\xi_i = N(0, \kappa I)$

Stacking the observations

$$z = X\beta + Qg + Wb + C'u + \varepsilon, \quad \varepsilon \sim N(0, \kappa I)$$

$$u = (u_1' \cdots u_n')'$$

$$C = \begin{bmatrix} C_1 & & \\ & \ddots & \\ & & C_n \end{bmatrix}$$

Algorithm

$\rho = (\rho_1, \cdots \rho_p)'$

$e_{it} = z_{it} - X_{it}'\beta + w_{it}' b_i - g(s_{it})$

$e_i = (e_{i,p+1}, \cdots e_{i,T_i})'$

$e = (e_1' \cdots e_n')$

1) $$z_{it} \sim \begin{cases} TN_{(0,\infty)}(\mu_{it}, v_{it}) & \text{if } y_{it} = 1 \\ TN_{(-\infty,0)}(\mu_{it}, v_{it}) & \text{if } y_{it} = 0 \end{cases}$$

μ_{it} & v_{it} are full conditional mean & variance

2) Sample β, ρ & u_i in one block

a) $\beta \mid y, D\{z_{it}\}, g, \rho \sim N(\hat{\beta}, B)$

$$\hat{\beta} = B\left(B_0^{-1}\beta_0 + \sum_{i=1}^{n} X_i' V_i^{-1}(z_i - g_i)\right)$$

$$B = \left(B_0^{-1} + \sum_{i=1}^{n} X' V_i^{-1} X_i\right)^{-1}$$

b) $$b_i \mid y, D, \{z_{it}\}, \beta, g, p \sim N(\hat{b}_i, B_i)$$

$$\hat{b}_i = B_i W_i' \Omega_i^{-1} (z_i - X_i\beta - g_i)$$
$$B_i = (D^{-1} + W_i' \Omega_1^{-1} W_i)^{-1}$$

c) $$u_i \mid y, \{z_{it}\}, \{b_i\}, \beta, g, p \sim N(\hat{u}_i, U_i)$$

$$\hat{u}_i = U_i C_i (z_i - X_i\beta - g_i - W_i b_i)$$
$$U_i = (I + C_i C_i'/\kappa)^{-1}$$

3) $$D^{-1} \mid \{b_i\} \sim W_p\left[\gamma_0 + n, \left(R_0^{-1} + \sum_{i=1}^{n} b_i b_i'\right)^{-1}\right]$$

4) $$g \mid y, \beta, \{b_i\}, \tau^2, \{z_{it}\}, \{u_{it}\} \sim N(\hat{g}, G)$$

$$G = (K/\tau^2 + Q'Q/\kappa)^{-1}$$
$$\hat{g} = G(K g_0/\tau^2 + Q'(z - X\beta - Wb - C'u)/\kappa)$$

(G^{-1} is banded, proceed as case 1)

5) $$\tau^2 \mid g \sim IG\left(\frac{\nu_0 + m}{2}, \cancel{\frac{\delta_0 + (g - g_0)' K (g - g_0)}{2}}\right)$$

6) $$p \mid y, g, \beta, \{b_i\}, \{z_{it}\} \propto \Psi(p) \times N(\hat{p}, P) \times I_{S_p}$$

$$\hat{p} = P(P_0^{-1} p_0 + E'e)$$
$$P = (P_0^{-1} + E'E)^{-1}$$
$$\Psi(p) = |\Omega_p|^{-n/2} \exp\left(-\frac{1}{2} \sum_{i=1}^{n} e_{i1}' \Omega_p^{-1} e_{i1}\right)$$

Simple linear model

$$y_{it} = 1\{\tilde{x}_{it}'\delta + w_{it}'\beta_i + \phi_1 y_{i,t-1} + \phi_2 y_{i,t-2} + \varepsilon_{it} > 0\}$$

Stacking for each cluster

$$\underset{5\times 1}{z_i} = \underset{5\times x}{\tilde{x}_i'}\ \underset{x\times 1}{\delta} + \underset{5\times y}{w_i'}\ \underset{y\times 1}{\beta_i} + \underset{5\times 1}{y_{i,-1}}\ \underset{1\times 1}{\phi_1} + \underset{5\times 1}{y_{i-2}}\ \underset{1\times 1}{\phi_2} + \underset{5\times 1}{\varepsilon_i}$$

$$\varepsilon_i \sim N(0, \Omega_i)$$

$$\underset{y\times 1}{\beta_i} = \underset{y\times l}{A_i}\ \underset{l\times 1}{\gamma} + \underset{y\times 1}{b_i} \qquad b_i \sim N(0, \underset{y\times y}{D})$$

$$A_i = \begin{bmatrix} 1 & \bar{y}_{i0} & \bar{\gamma}_{i1}' & & & \\ & & & \ddots & & \\ & & & 1 & \bar{y}_{i0} & \bar{\gamma}_{i1} \end{bmatrix}$$

$$z_i = \underset{5\times 11}{X_i}\ \underset{11\times 1}{\beta} + \underset{5\times 3}{W_i}\ \underset{3\times 1}{b_i} + \underset{5\times 1}{\varepsilon_i}, \quad \varepsilon_i \sim N(0, \underset{5\times 5}{\Omega_i})$$

$$b_i \sim N(0, \underset{3\times 3}{D})$$

$$X_i = [\underset{5\times 6}{\tilde{x}_i'} \quad \underset{5\times 3}{w_i'} \quad \underset{5\times 1}{y_{i,-1}} \quad \underset{5\times 1}{y_{i-1}}]$$

$$\tilde{x}_{it} = \underset{6\times 1}{\begin{bmatrix} 1 \\ Race_i \\ Edu_{it} \\ \ln(Inc_{it}) \\ Age_{it} \\ Age^2_{it} \end{bmatrix}} \qquad w_{it} = \underset{3\times 1}{\begin{bmatrix} 1 \\ CH_{2,it} \\ CH_{5,it} \end{bmatrix}}$$

Priors

$\beta \sim N(\beta_0, B_0)$

$z_i = X_i \beta + W_i b_i + \varepsilon_i$

$z_i = X_i \beta + u_i \quad , \quad u_i \sim N(0, \underset{5\times 5}{V_i})$

$V_i = \underset{5\times 3}{W_i'} \underset{3\times 3}{D} \underset{3\times 5}{W_i} + \underset{5\times 5}{\Omega_i}$

stacking over i

$\underset{(5\times 1500)\times 1}{Z} = \underset{[(5\times 1500)\times 11]}{X} \underset{(11\times 1)}{\beta} + u \qquad u \sim N(0, V)$

$$V = \begin{bmatrix} V_1 & & \\ & \ddots & \\ & & V_{1500} \end{bmatrix}$$

$$\beta \sim |V|^{-1/2} \exp\{(z - X\beta)V^{-1}(z - X\beta)\} |B_0|^{-1/2} \exp\{(\beta - \beta_0)' B_0^{-1}(\beta - \beta_0)\}$$

$$\sim \exp\{-z'V^{-1}X\beta - \beta'XV^{-1}z + \beta'XV^{-1}X\beta + \beta'B_0^{-1}\beta - \beta'B_0^{-1}\beta_0 - \beta_0'B_0^{-1}\beta\}$$

$$= \exp\{\beta'[XV^{-1}X + B_0^{-1}]\beta - \beta'[XV^{-1}z + B_0^{-1}\beta_0] - [z'V^{-1}X + \beta_0'B_0^{-1}]\beta$$

$$\sigma_i^2 \sim IG\left(\frac{v_0 + n}{2}, \frac{\delta_0 + (z_i - X_i\beta)'(z_i - X_i\beta)}{2}\right)$$