

Mixed Nonhomogeneous Poisson Process Spline Models

for the Analysis of Recurrent Event Panel Data

Jason D. Nielsen and Charmaine Dean
`jdn@stat.sfu.ca`

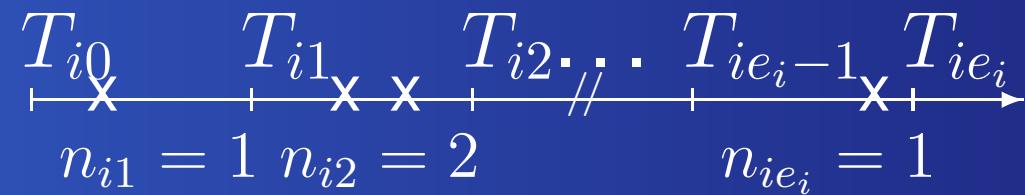
`http://www.stat.sfu.ca/~jdn`

Department of Statistics and Actuarial Science
Simon Fraser University

Recurrent Events

- Interested in how recurrence rate changes across treatments and over time
 - Focus on splines
- Within and between subject correlation
 - Clustering
- Computational challenge
 - Efficient algorithms

Panel Data



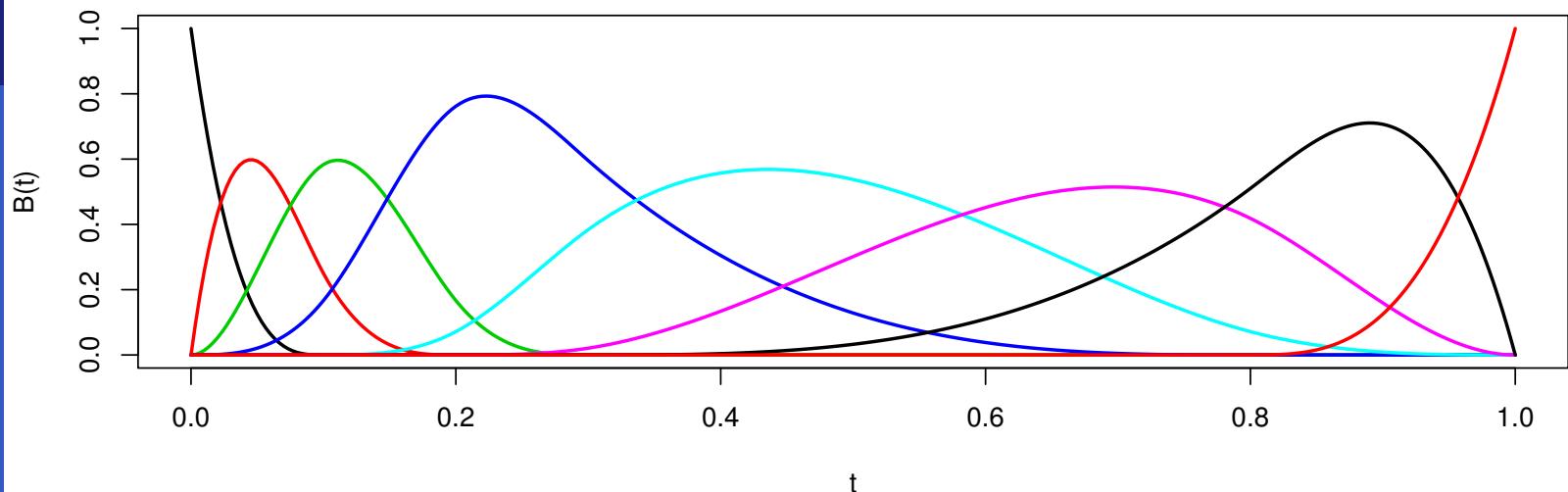
- Individuals examined at fixed set of times, $\{T_{ij}\}$
- Record number of events, n_{ij} , in each subinterval, $[T_{i(j-1)}, T_{ij})$

Splines

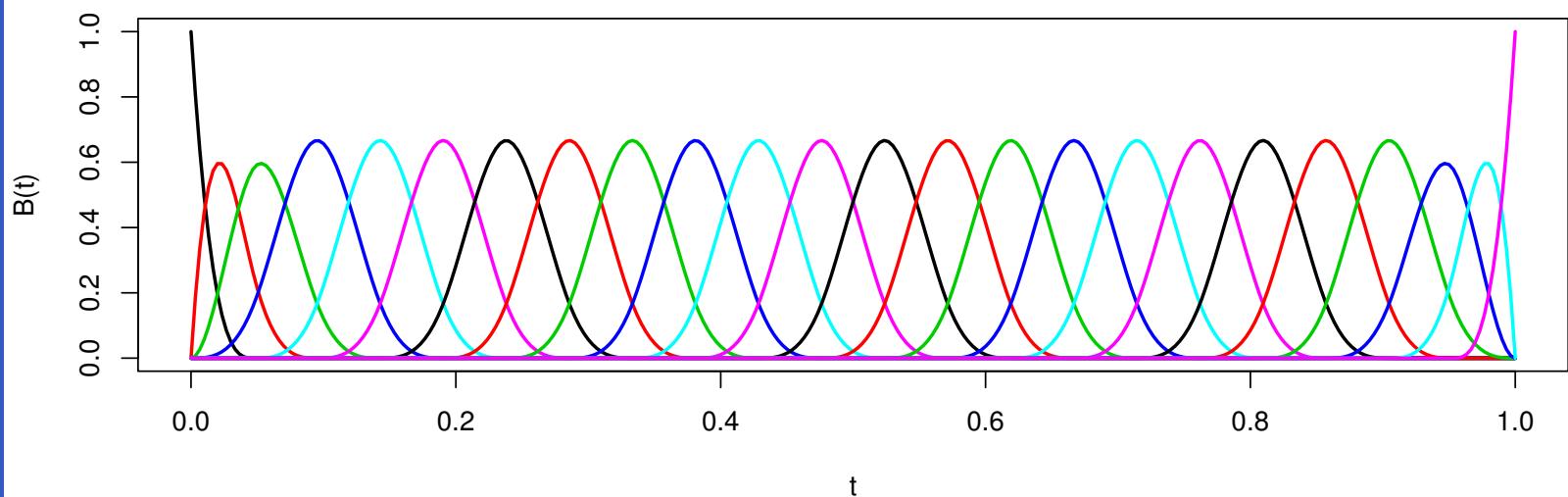
Two basic strategies:

- Penalty Methods - restrict the amount of flexibility.
- Free Knot - Select optimal position for basis functions.

Knots at [.1, .2, .3, .8]



20 equally spaced knots



Mixed NHPP Model

- Let $M(\nu)$ denote a known mixing distribution
- Conditional on ν counting process for each individual follows a NHPP with intensity with function $\lambda(\omega; \nu_i)$

Mixed NHPP Model

- Let $M(\nu)$ denote a known mixing distribution
- Conditional on ν counting process for each individual follows a NHPP with intensity with function $\lambda(\omega; \nu_i)$
- Cumulative Intensity Function

$$\Lambda_i(\omega; \nu_i) = \int_0^\omega \lambda_i(t; \nu_i) dt$$

- Distribution of $\mathbf{N}_i = \{N_i[T_{i(j-1)}, T_{ij})\}_{e_i \times 1}$ is

$$P_M(\mathbf{N}_i) = \int \prod_{j=1}^{e_i} P(N_{ij} \mid \boldsymbol{\nu}_i) dM(\boldsymbol{\nu}_i)$$

- Distribution of $\mathbf{N}_i = \{N_i[T_{i(j-1)}, T_{ij})\}_{e_i \times 1}$ is

$$P_M(\mathbf{N}_i) = \int \prod_{j=1}^{e_i} P(N_{ij} \mid \boldsymbol{\nu}_i) dM(\boldsymbol{\nu}_i)$$

with expectation

$$\boldsymbol{\mu}_i = \{\mu_{ij}\}_{e_i \times 1} = \{E[N_{ij}]\}$$

$$\mu_{ij} = E[\Lambda_i(T_{ij}; \boldsymbol{\nu}_i) - \Lambda_i(T_{i(j-1)}; \boldsymbol{\nu}_i)]$$

Spline Intensity

$$\begin{aligned}\lambda_i(\omega) &= \lambda(\omega; \boldsymbol{\nu}_i, \mathcal{X}_i) \\ &= \exp \left\{ \gamma_o(\omega; \boldsymbol{\nu}_i) + \sum_{q=1}^p x_{iq}(\omega) \gamma_q(\omega; \boldsymbol{\nu}_i) \right\}\end{aligned}$$

Spline Intensity

$$\begin{aligned}\lambda_i(\omega) &= \lambda(\omega; \boldsymbol{\nu}_i, \mathcal{X}_i) \\ &= \exp \left\{ \gamma_o(\omega; \boldsymbol{\nu}_i) + \sum_{q=1}^p x_{iq}(\omega) \gamma_q(\omega; \boldsymbol{\nu}_i) \right\}\end{aligned}$$

- Could also use constraints

Spline Intensity

$$\begin{aligned}\lambda_i(\omega) &= \lambda(\omega; \boldsymbol{\nu}_i, \mathcal{X}_i) \\ &= \exp \left\{ \gamma_o(\omega; \boldsymbol{\nu}_i) + \sum_{q=1}^p x_{iq}(\omega) \gamma_q(\omega; \boldsymbol{\nu}_i) \right\}\end{aligned}$$

$$\gamma_q(\omega; \boldsymbol{\nu}_i) = \sum_{r=1}^{k+4} \psi_{qr}(\boldsymbol{\nu}_i) B_{r-4}(\omega)$$

- Knots at panel midpoints: $\left\{ \frac{T_{i(j-1)} + T_{ij}}{2} \right\}$

Spline Intensity

$$\begin{aligned}\lambda_i(\omega) &= \lambda(\omega; \boldsymbol{\nu}_i, \mathcal{X}_i) \\ &= \exp \left\{ \gamma_o(\omega; \boldsymbol{\nu}_i) + \sum_{q=1}^p x_{iq}(\omega) \gamma_q(\omega; \boldsymbol{\nu}_i) \right\}\end{aligned}$$

- Generally p small!
- $x_{ir}(\omega)$'s often discrete

- Need to compute:

$$\int_{T_{i(j-1)}}^{T_{ij}} \lambda_i(\omega; \boldsymbol{\nu}_i) d\omega$$

- Need to compute:

$$\int_{T_{i(j-1)}}^{T_{ij}} \lambda_i(\omega; \boldsymbol{\nu}_i) d\omega$$

- No closed form
- Use q-point L-G quadrature
- Monotonicity

Finite Mixture

- Model possible hidden subpopulations
- Decomposition of intensity into components
- Between individual correlation

Heterogeneity

- Individual specific, ν_i , iid random effects
- Mean 1 and variance τ

Heterogeneity

- Individual specific, ν_i , iid random effects
- Mean 1 and variance τ
- Intensity:

$$\lambda(\omega; \boldsymbol{\nu}_i, \mathcal{X}_i) = \nu_i(\tau) \lambda(\omega; \boldsymbol{\psi}, \mathcal{X}_i)$$

Heterogeneity

- Individual specific, ν_i , iid random effects

- Mean 1 and variance τ

- Intensity:

$$\lambda(\omega; \boldsymbol{\nu}_i, \mathcal{X}_i) = \nu_i(\tau) \lambda(\omega; \boldsymbol{\psi}, \mathcal{X}_i)$$

- Mean:

$$\boldsymbol{\mu}_i = \{\mu_{ij}\}_{e_i \times 1} = \{\Lambda_i(T_{ij}) - \Lambda_i(T_{i(j-1)})\}$$

Estimation of $\psi = [\psi_o^T, \psi_1^T, \dots, \psi_p^T]^T$:

$$\mathbf{g}_\psi = \sum_{i=1}^I \mathbf{D}_i^T \mathbf{V}_i^{-1} (\mathbf{N}_i - \boldsymbol{\mu}_i) - (\boldsymbol{\delta} \otimes \mathbf{P}) \boldsymbol{\psi}$$

Estimation of $\psi = [\psi_o^T, \psi_1^T, \dots, \psi_p^T]^T$:

$$\mathbf{g}_\psi = \sum_{i=1}^I \mathbf{D}_i^T \mathbf{V}_i^{-1} (\mathbf{N}_i - \boldsymbol{\mu}_i) - (\boldsymbol{\delta} \otimes \mathbf{P}) \boldsymbol{\psi}$$

$$\mathbf{D}_i = \frac{\partial \boldsymbol{\mu}_i}{\partial \boldsymbol{\psi}^T}; \quad \mathbf{V}_i = \text{diag}\{\boldsymbol{\mu}_i\} + \tau \boldsymbol{\mu}_i \boldsymbol{\mu}_i^T$$

Estimation of $\psi = [\psi_o^T, \psi_1^T, \dots, \psi_p^T]^T$:

$$\mathbf{g}_\psi = \sum_{i=1}^I \mathbf{D}_i^T \mathbf{V}_i^{-1} (\mathbf{N}_i - \boldsymbol{\mu}_i) - (\boldsymbol{\delta} \otimes \mathbf{P}) \boldsymbol{\psi}$$

$$\mathbf{P} = \sum_{i=1}^I \int_{T_{io}}^{T_{ie_i}} \mathbf{b}'(\omega) \mathbf{b}'(\omega)^T \mathbf{d}\omega$$

- Form of penalty:

$$\delta_r \rightarrow \infty \quad \Rightarrow \quad \gamma_r(\omega) \rightarrow \beta_r$$

- Form of penalty:

$$\delta_r \rightarrow \infty \quad \Rightarrow \quad \gamma_r(\omega) \rightarrow \beta_r$$

- Proportional intensity model
- Extreme case: Poisson regression

Smoothing parameters:

$$\tilde{\delta}_r^{(j)} = \frac{e_{\tilde{\psi}_r} \left(\tilde{\delta}_r^{(j-1)} \right)}{\tilde{\psi}_r^T \mathbf{P} \tilde{\psi}_r}$$

- $e_{\tilde{\psi}_r} \left(\tilde{\delta}_r^{(j-1)} \right)$ effective df for ψ_r

Smoothing parameters:

$$\tilde{\delta}_r^{(j)} = \frac{e_{\tilde{\psi}_r} \left(\tilde{\delta}_r^{(j-1)} \right)}{\tilde{\psi}_r^T \mathbf{P} \tilde{\psi}_r}$$

- $e_{\tilde{\psi}_r} \left(\tilde{\delta}_r^{(j-1)} \right)$ effective df for ψ_r
- Empirical Bayes hybrid

Overdispersion parameter τ :

$$g_\tau = \sum_{i=1}^I \frac{(N_{i+} - \mu_{i+})^2 - \mu_{i+}(1 + \tau\mu_{i+}) + r_i}{(1 + \tau\mu_{i+})^2}$$

Overdispersion parameter τ :

$$g_\tau = \sum_{i=1}^I \frac{(N_{i+} - \mu_{i+})^2 - \mu_{i+}(1 + \tau\mu_{i+}) + r_i}{(1 + \tau\mu_{i+})^2}$$

where

$$r_i = \text{tr} \left[\left(\sum_{i=1}^I \mathbf{D}_i^T \mathbf{V}_i^{-1} \mathbf{D}_i + (\boldsymbol{\delta} \otimes \mathbf{P}) \right)^{-1} \mathbf{D}_i^T \mathbf{V}_i^{-1} \mathbf{D}_i \right]$$

Finite Mixture

- Model possible hidden subpopulations
- Decomposition of intensity into components
- Between individual correlation

- Intensity Function:

$$\lambda(\omega; \boldsymbol{\nu}_i, \mathcal{X}_i) = \sum_{g=1}^G \nu_i(\tau_g) \lambda_i(\omega; \boldsymbol{\psi}_g) \mathbf{1}\{i \in \mathcal{V}_g\}$$

- Intensity Function:

$$\lambda(\omega; \boldsymbol{\nu}_i, \mathcal{X}_i) = \sum_{g=1}^G \nu_i(\tau_g) \lambda_i(\omega; \boldsymbol{\psi}_g) \mathbf{1}\{i \in \mathcal{V}_g\}$$

- Mean:

$$\mu_{ij} = \sum_{g=1}^G \mu_{gij} = \sum_{g=1}^G p_g [\Lambda_{gi}(T_{ij}) - \Lambda_{gi}(T_{i(j-1)})]$$

Estimation:

$$\mathbf{g}\psi_g = \sum_{i=1}^I z_{gi}^* \mathbf{D}_{gi}^T \mathbf{V}_{gi}^{-1} (\mathbf{N}_i - \boldsymbol{\mu}_{gi}) - (\boldsymbol{\delta}_g \otimes \mathbf{P}) \psi_g$$

Estimation:

$$\mathbf{g}_{\psi_g} = \sum_{i=1}^I z_{gi}^* \mathbf{D}_{gi}^T \mathbf{V}_{gi}^{-1} (\mathbf{N}_i - \boldsymbol{\mu}_{gi}) - (\boldsymbol{\delta}_g \otimes \mathbf{P}) \boldsymbol{\psi}_g$$

- Similar weighted sum for τ_g 's
- Closed form for probs: $p_g = \sum_{i=1}^I z_{gi}^* / I$

Weights:

$$z_{gi}^* = \frac{p_g P(\mathbf{N}_i \mid i \in \mathcal{V}_g, \nu_{gi}^*)}{\sum_{g=1}^G p_g P(\mathbf{N}_i \mid i \in \mathcal{V}_g, \nu_{gi}^*)}$$

Weights:

$$z_{gi}^* = \frac{p_g P(\mathbf{N}_i \mid i \in \mathcal{V}_g, \nu_{gi}^*)}{\sum_{g=1}^G p_g P(\mathbf{N}_i \mid i \in \mathcal{V}_g, \nu_{gi}^*)}$$

where

$$\nu_{gi}^* = \frac{1 + \tilde{\tau}_g n_{i+}}{1 + \tilde{\tau}_g \tilde{\mu}_{i+}}$$

posterior expectation under gamma frailty.

Algorithm

Notorious optimization problem:

- 1) EM from random initialization
- 2) Max. Directional Derivative
- 3) Swap weak/strong cluster
- 4) Repeat

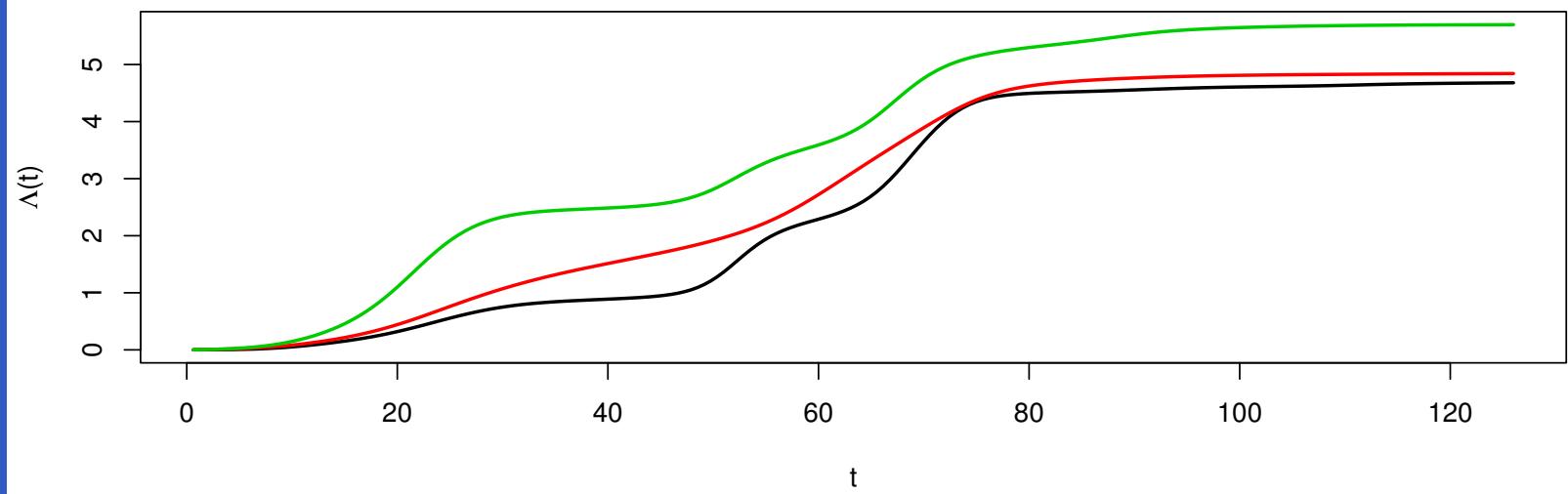
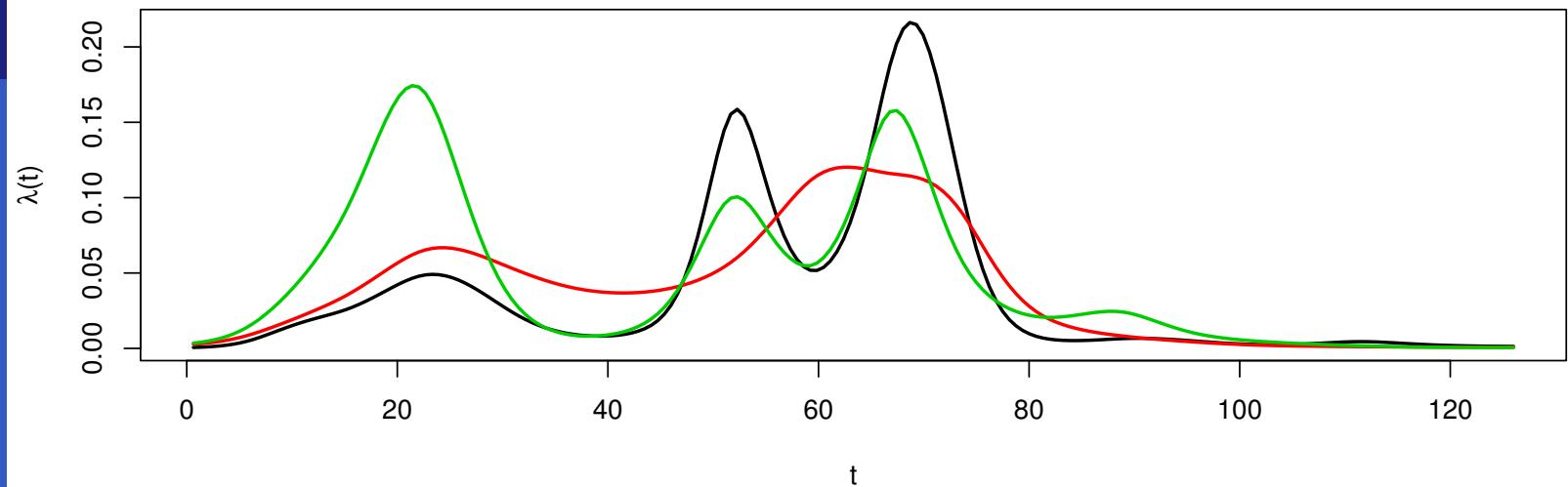
Cherry Bark Tortrix Data

Experiment to test effectiveness of pheromone-based mating disruption of Cherry Bark Tortrix moths.

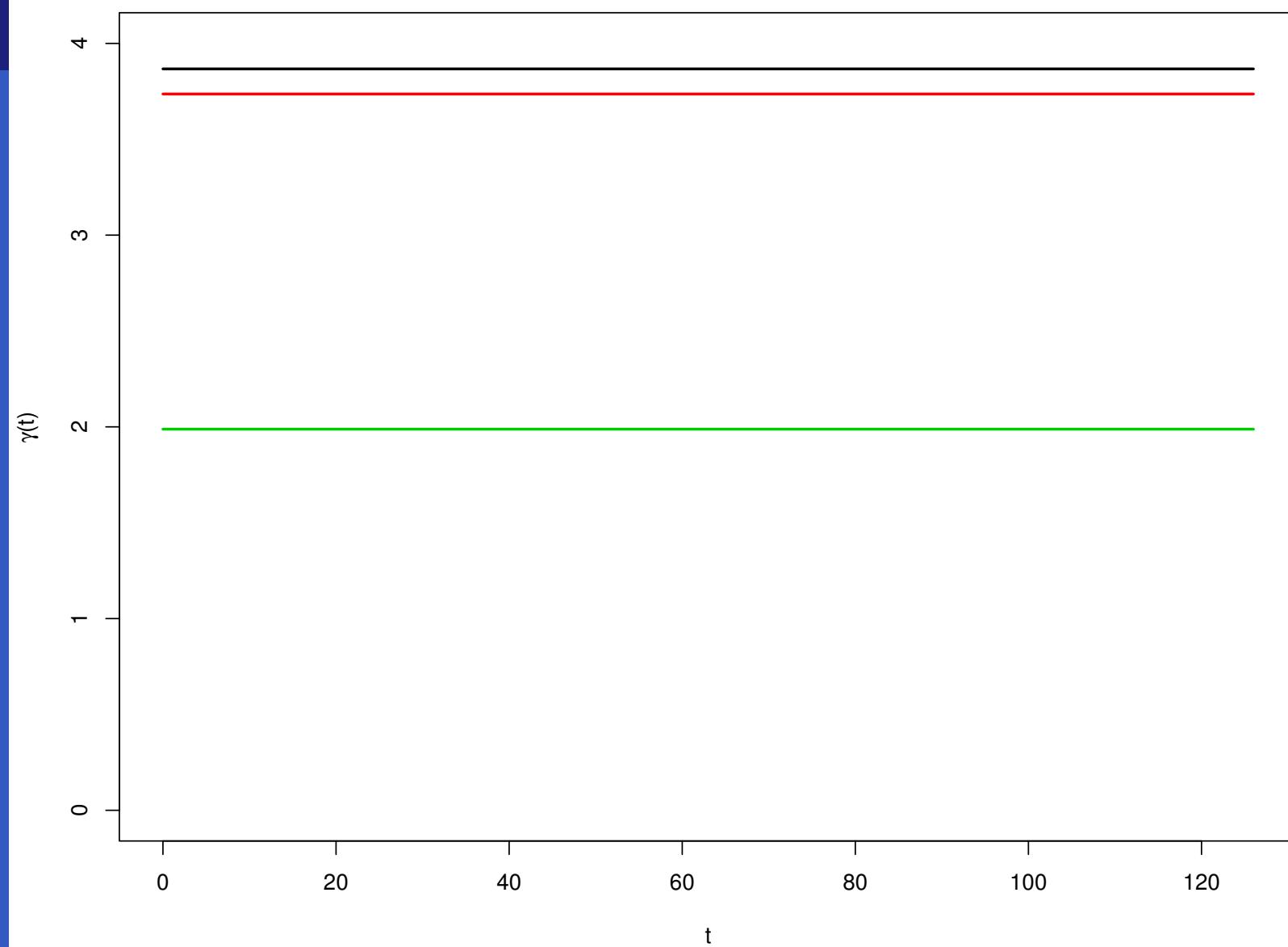
Design:

- 20 trees outfitted with trap & dispenser
- 10 randomly chosen, dispensers filled
- Traps checked weekly for 19 follow-ups
- Baits refreshed every 3 weeks

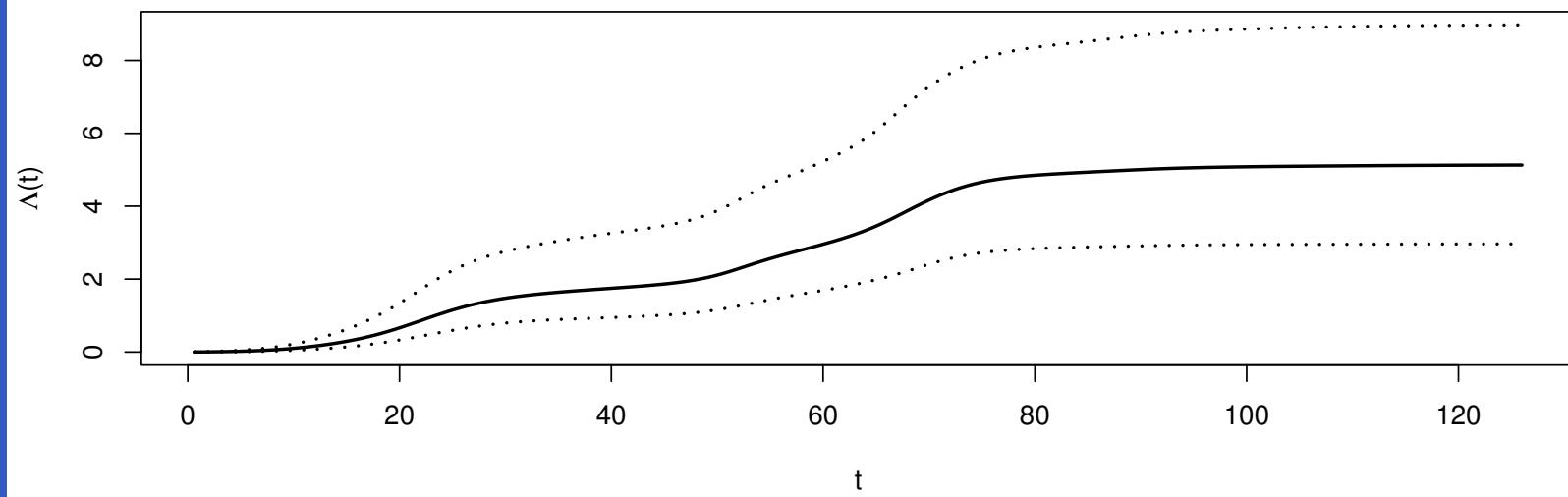
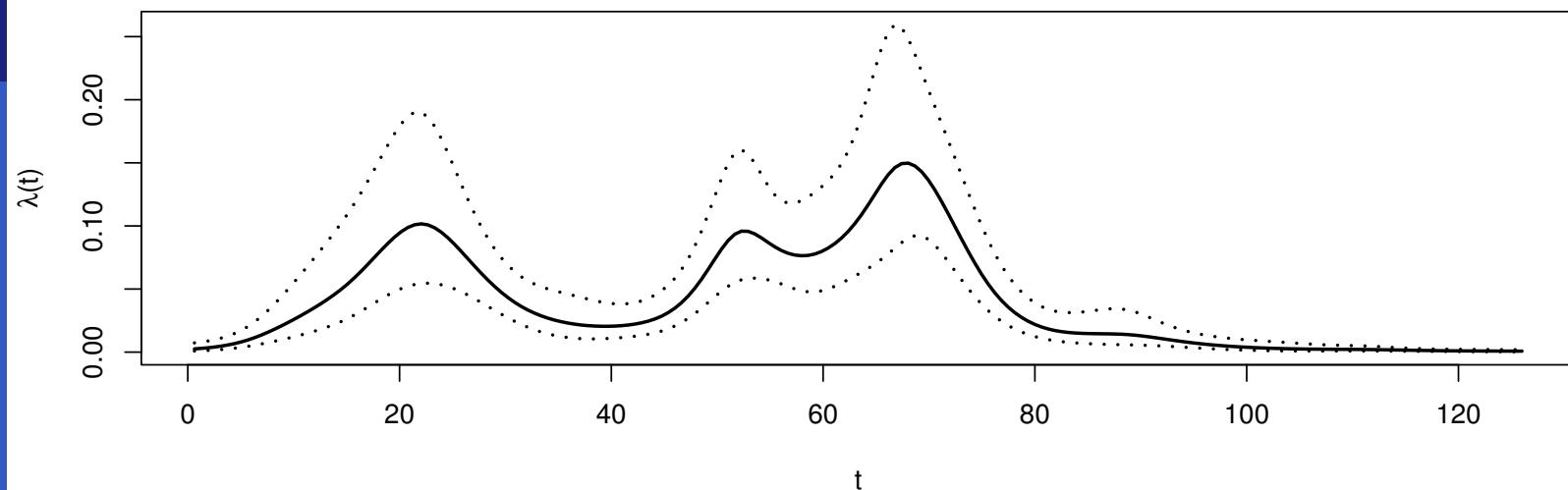
Baseline – Group Level



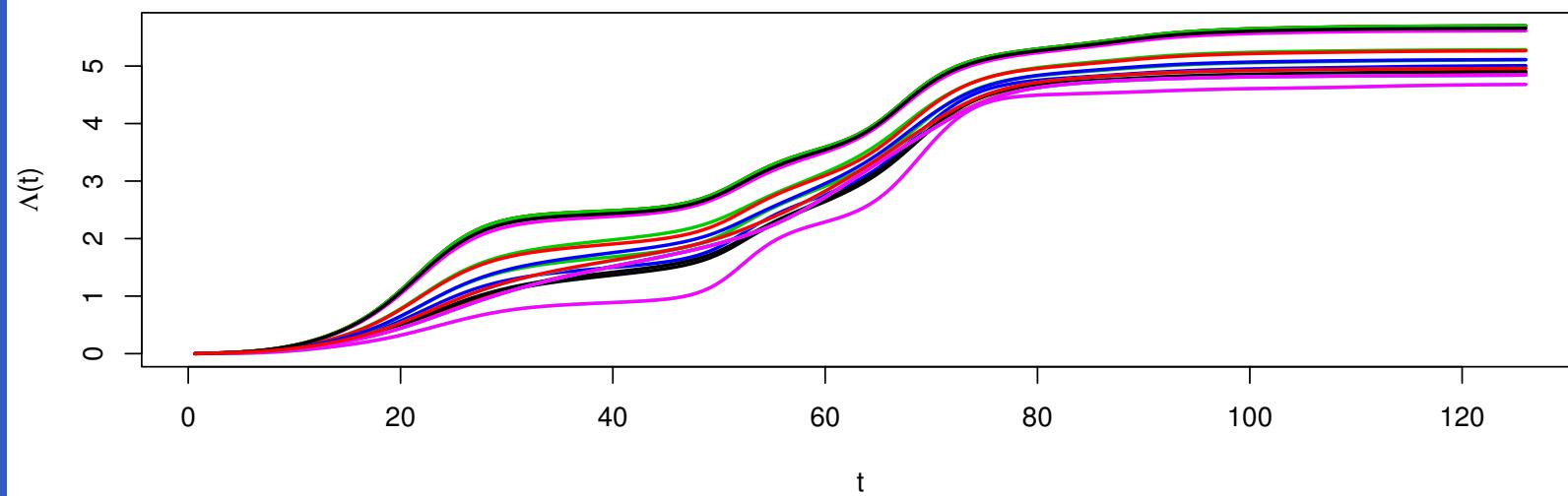
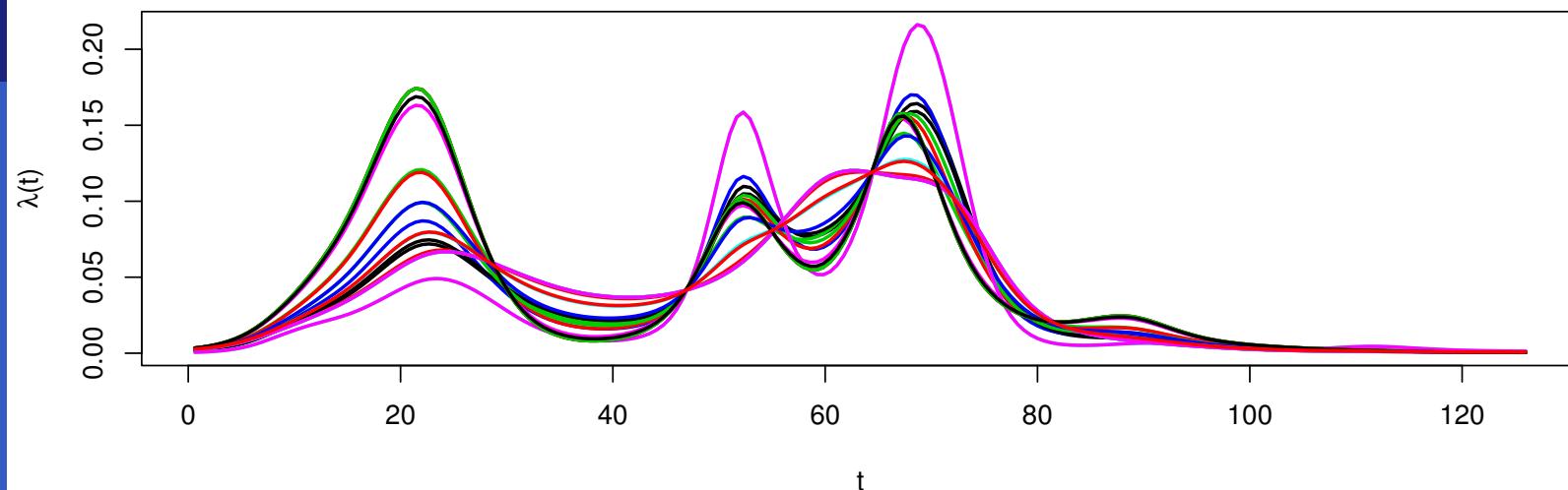
Trt – Group Level

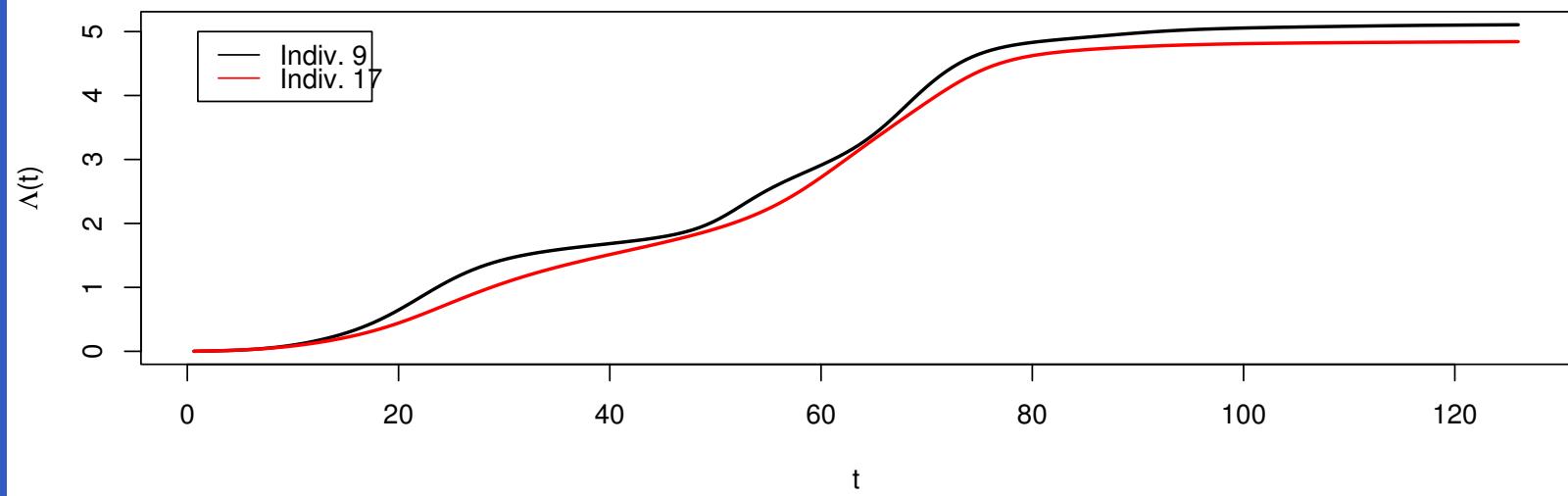
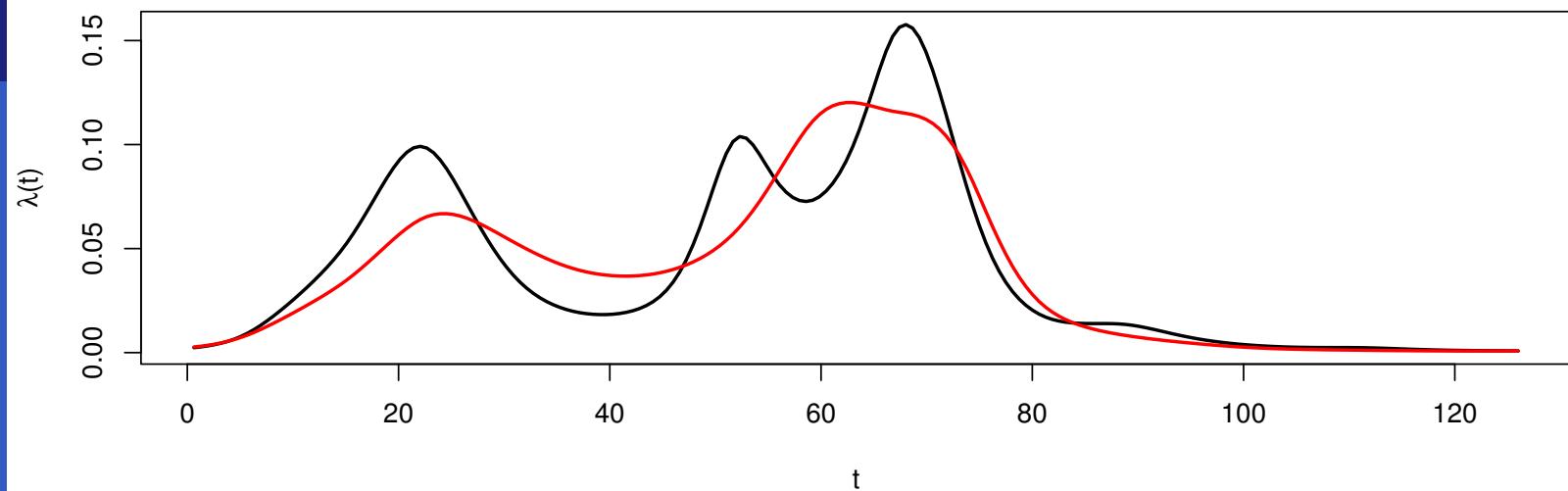


Baseline – Overall



Baseline – Individual Level





Estimates

	Group 1	Group 2	Group 3
Trt	3.87 (0.40)	3.74 (0.41)	1.99 (0.51)
Disp	0.49 (0.03)	0.93 (0.40)	0.69 (0.17)
Prob	0.21 (0.12)	0.41 (0.15)	0.38 (0.14)

Simulation

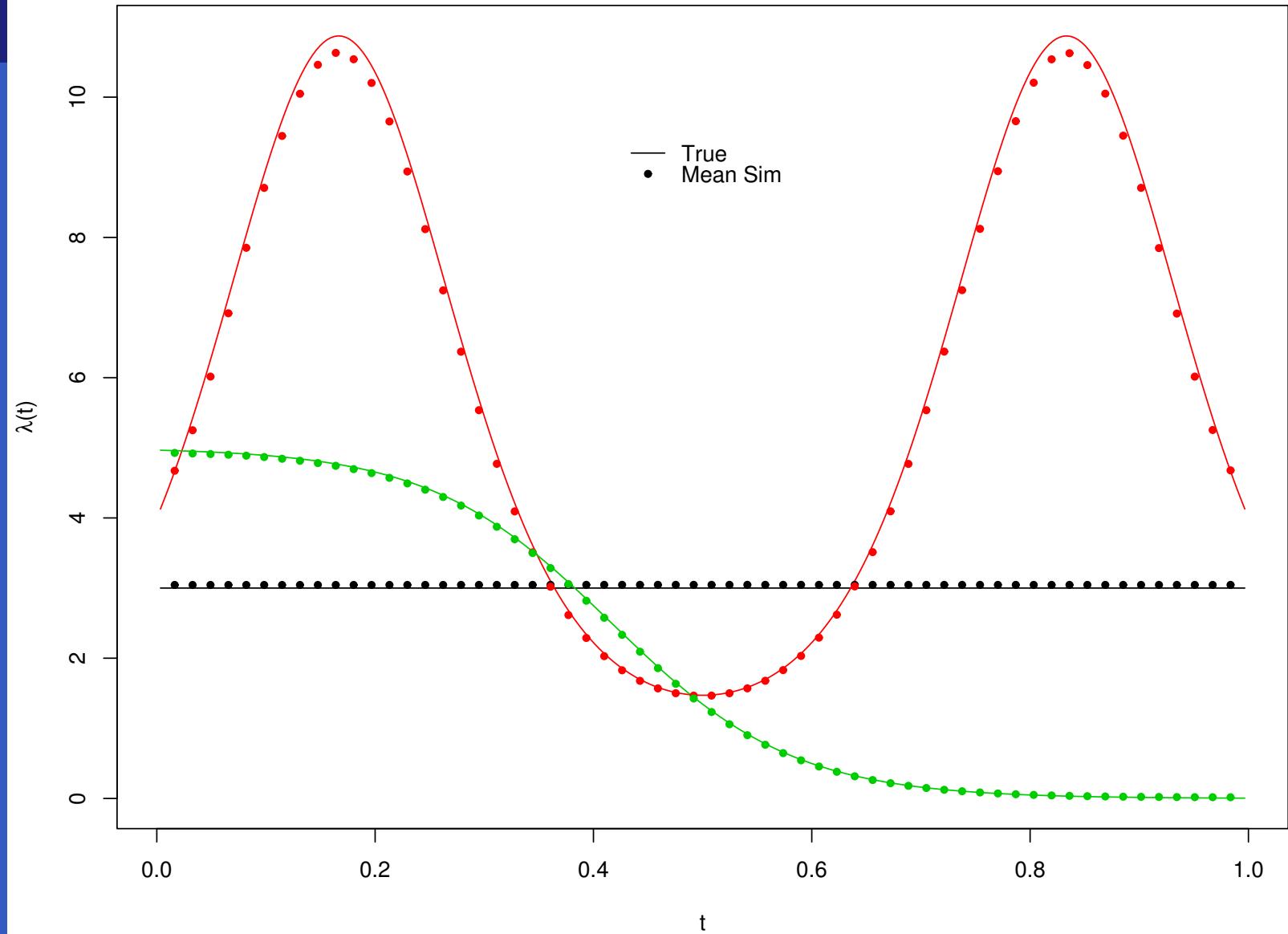
- Generate 3 component finite mixture of MNHPP models

Simulation

- Generate 3 component finite mixture of MNHPP models
- Forms:
 - Flat (homogeneous)
 - Good treatment
 - Bimodal

Simulation

- Generate 3 component finite mixture of MNHPP models
- 80 individuals, 30 follow-up
- $\mathbf{p} = [0.1, 0.4, 0.5]^T$
- $\boldsymbol{\tau} = [0.6, 0.4, 0.3]^T$

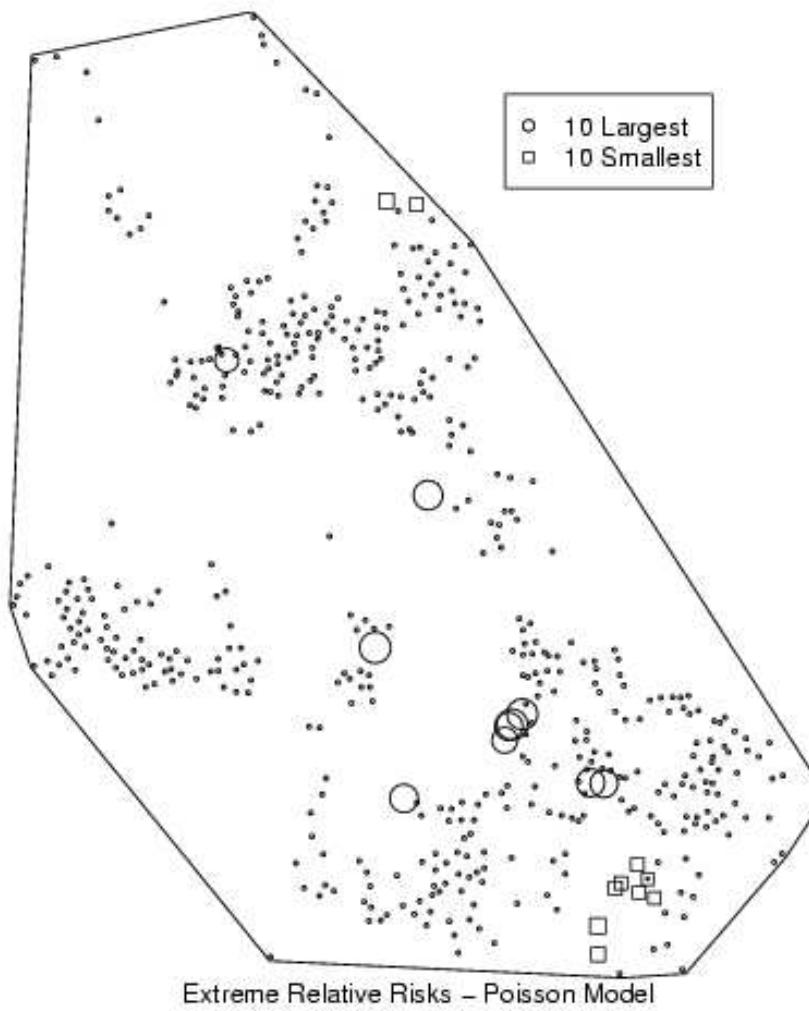


Simulation Results

		True	Mean	Sim. SE	Rob. SE	95% CP
Group 1	p_1	0.100	0.094	0.032	0.037	0.904
	$\lambda_1(t)$	—	—	—	—	~0.890
	τ_1	0.600	0.502	0.236	0.149	0.467
Group 2	p_2	0.400	0.407	0.056	0.061	0.945
	$\lambda_2(t)$	—	—	—	—	~0.937
	τ_2	0.400	0.437	0.108	0.185	0.938
Group 3	p_3	0.500	0.499	0.056	0.057	0.946
	$\lambda_3(t)$	—	—	—	—	~0.935
	τ_3	0.300	0.301	0.069	0.082	0.907

Future

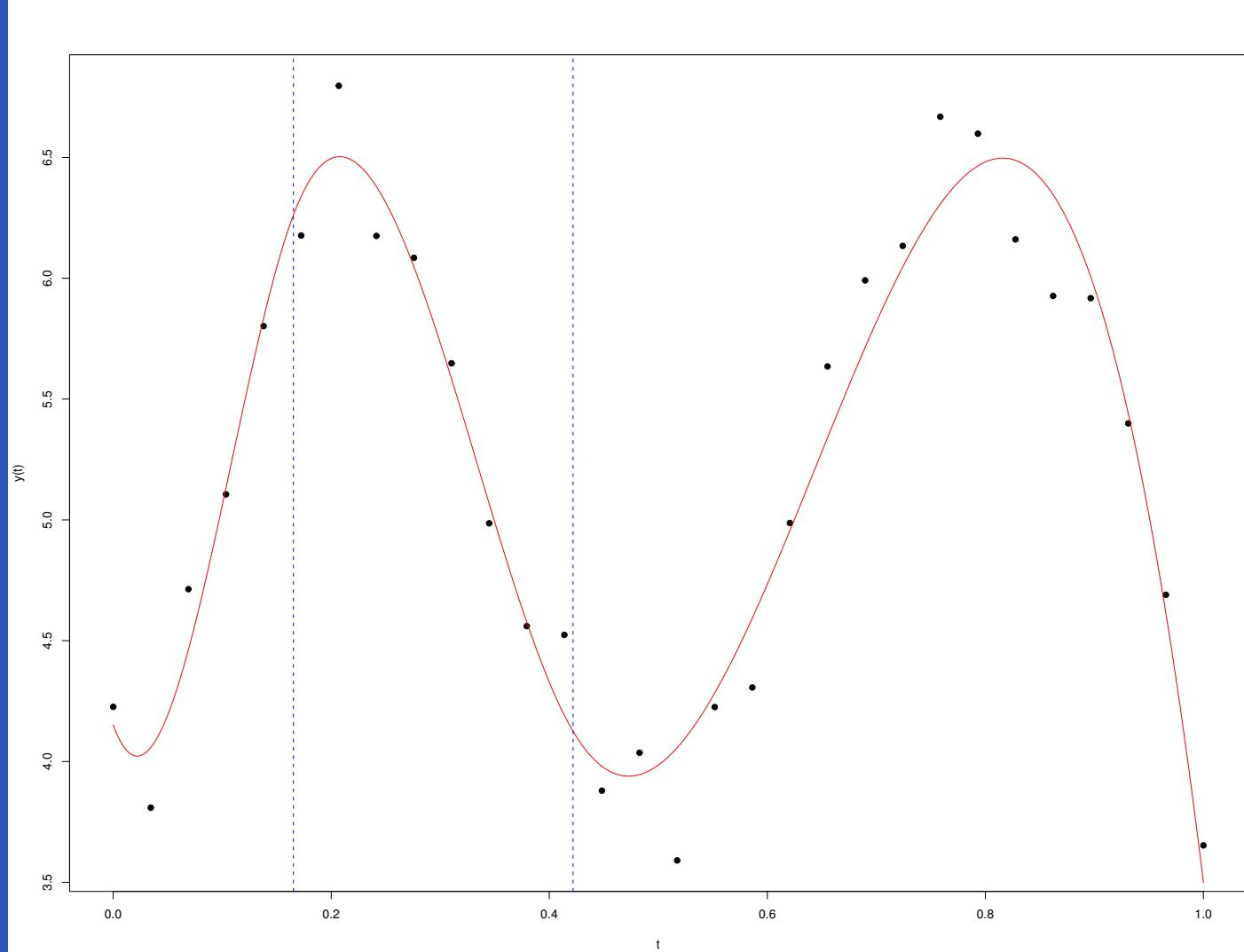
- Spatial Random Effects
- Efficiency lost by using Panel data?
- Estimating Equation Bootstrap
- Better, faster algorithm

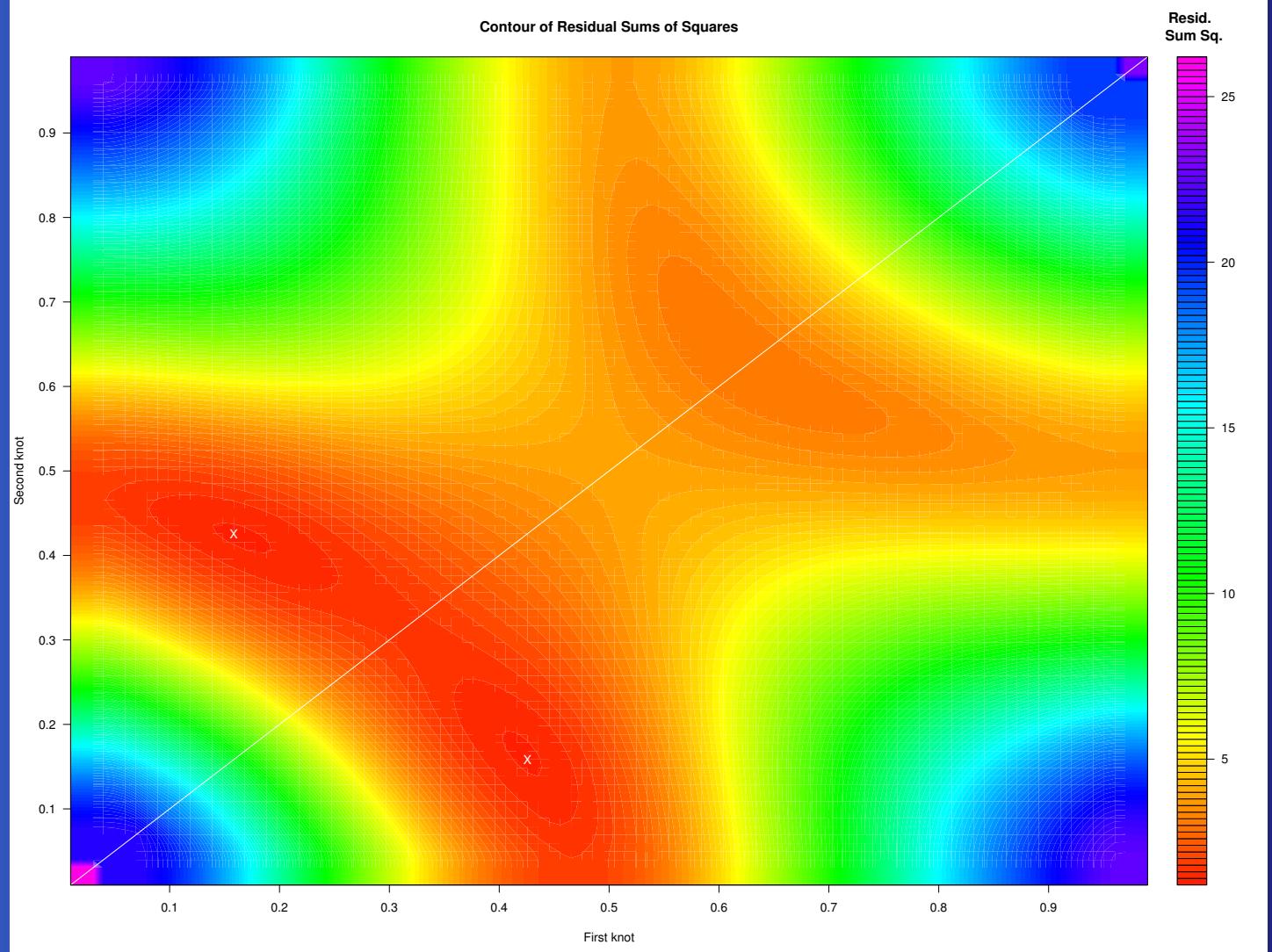


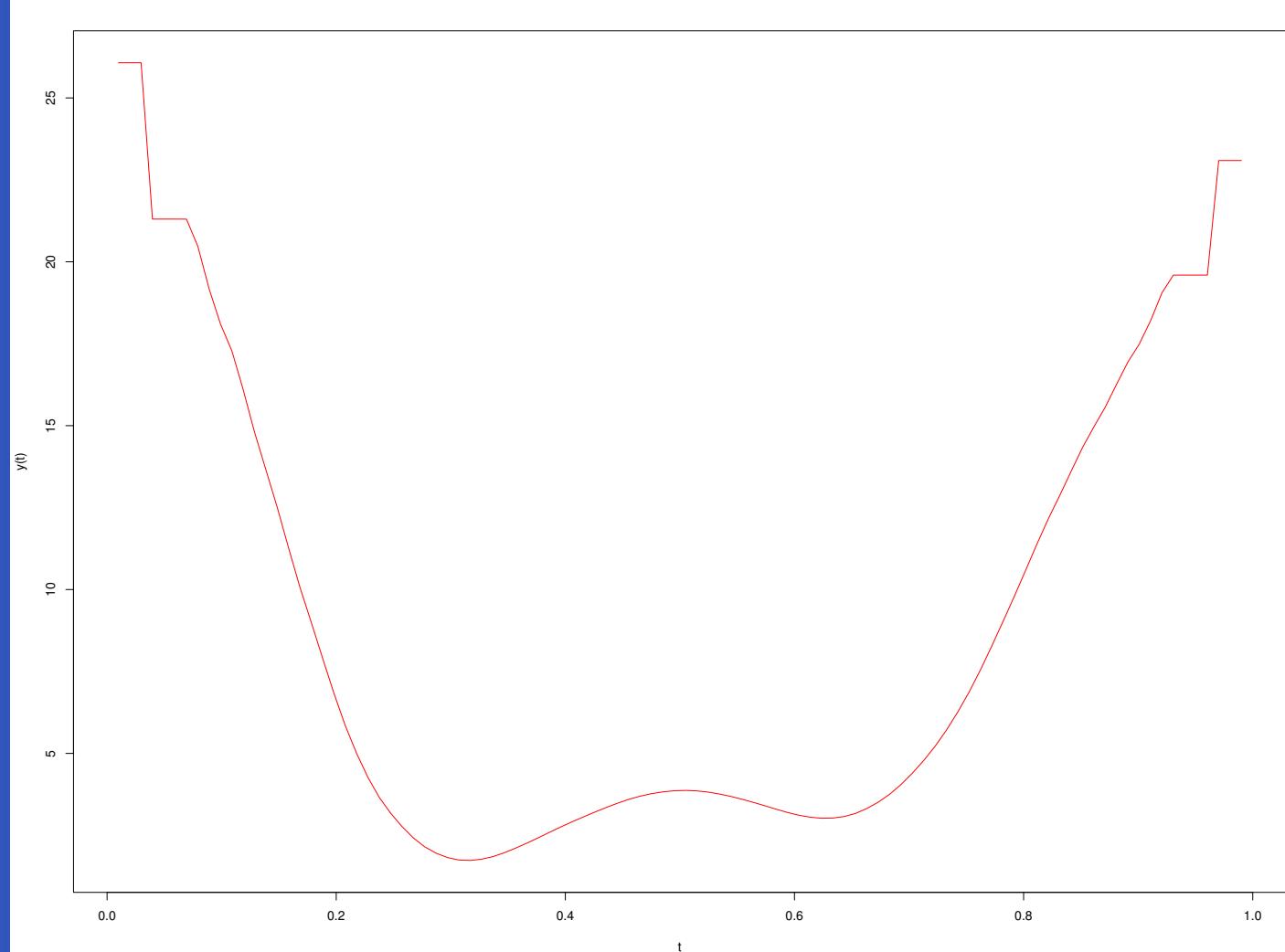
Future

- Spatial Random Effects
- Efficiency lost by using Panel data?
- Estimating Equation Bootstrap
- Better, faster algorithm

Illustration







Finding knots ξ ($\delta = 0$):

$$\max_{\boldsymbol{\xi}} \sum_{i=1}^M \tilde{\mathbf{z}}_i^T \widetilde{\mathbf{W}}_i(\boldsymbol{\xi}) \tilde{\mathbf{z}}_i$$

Finding knots ξ ($\delta = 0$):

$$\max_{\boldsymbol{\xi}} \sum_{i=1}^M \tilde{\mathbf{z}}_i^T \widetilde{\mathbf{W}}_i(\boldsymbol{\xi}) \tilde{\mathbf{z}}_i$$

where

$$\tilde{\mathbf{z}}_i = \left\{ \log[n_{ij}/(T_{ij} - T_{i(j-1)})] - \mathbf{x}_i^T \tilde{\boldsymbol{\beta}} - E[\nu_i | \tilde{\mu}_{i+}] \right\}_{e_i \times 1}$$

$$\widetilde{\mathbf{W}}_i = \text{diag}\{\tilde{\mu}_i\} \mathbf{B}_i \left(\mathbf{B}_i^T \text{diag}\{\tilde{\boldsymbol{\mu}}_i\} \mathbf{B}_i \right)^{-} \mathbf{B}_i^T \text{diag}\{\tilde{\boldsymbol{\mu}}_i\}$$