Repeated measures models with multiple, correlated random effects

April 8, 2008

Steve Gregorich

Foreword

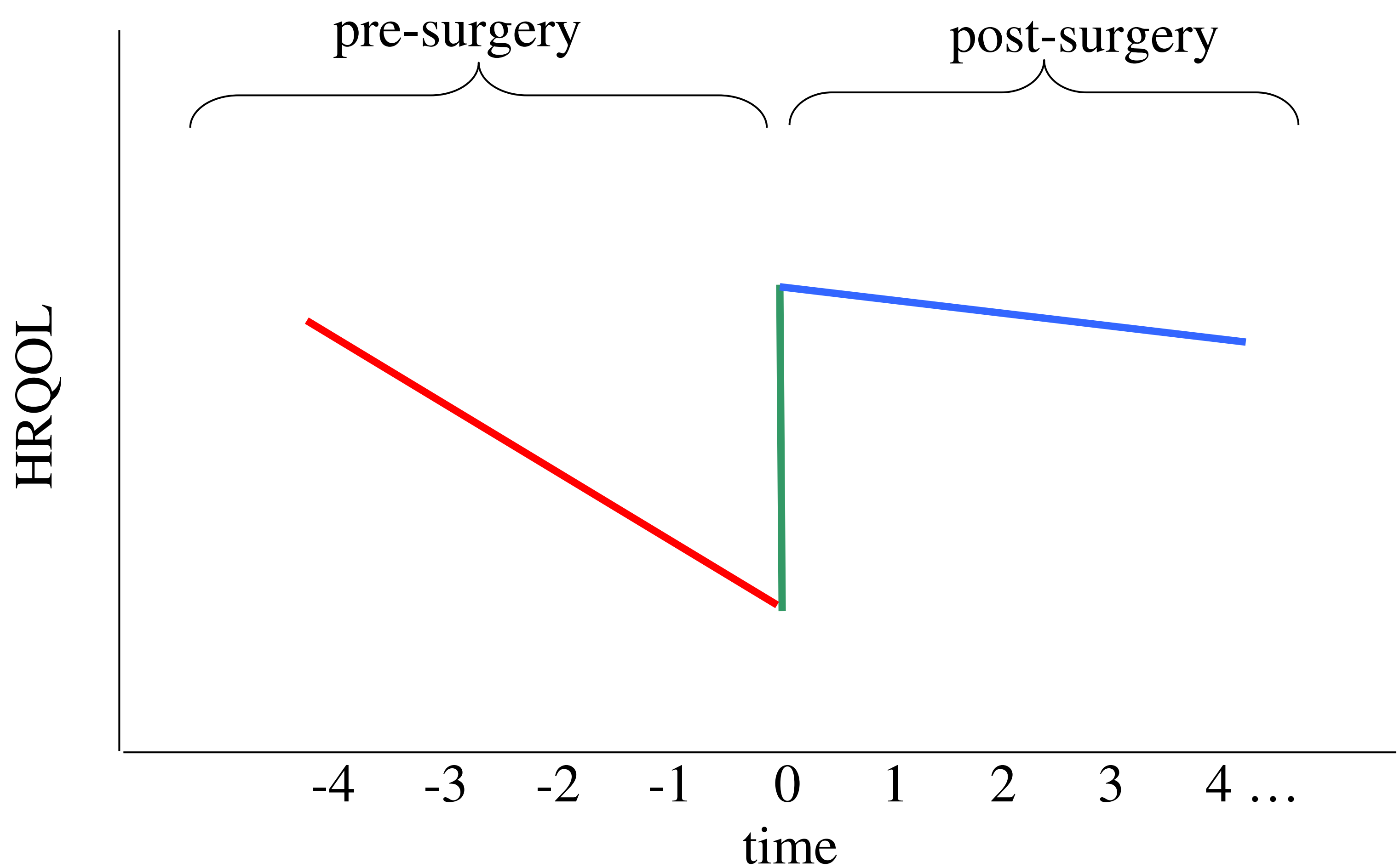
Application, nothing novel

More about the models than substance…
Mostly my goal is to describe analysis options

Some of the worked examples don't exploit the full strengths of the models that I describe

General research questions:

Part 1: pre- and post-surgical HRQOL trajectories



General research questions:

Part 1: pre- and post-surgical HRQOL trajectories

What are patients' HRQOL trajectories before & after a surgical intervention?

What is the average instantaneous 'bump' in HQROL resulting from surgery?

To what extent are patients' pre-surgical HRQOL trajectories associated with
the magnitude of their surgery-attributable HRQOL 'bumps'?
their post-surgical HRQOL trajectories?

To what extent are the magnitudes of patients' surgery-attributable
HRQOL 'bumps' associated with their post-surgical HRQOL trajectories?

General research questions:

Part 2: Associations between sexual functioning and HRQOL trajectories

Are women's sexual functioning trajectories associated with their HRQOL trajectories?

E.g., are changes across time on one dimension associated with changes across time on another dimension?

Study design

SOPHIA: Study of Pelvic Problems, Hysterectomy & Intervention Alternatives

Prospective observation study: N = 1493 women

Eligibility criteria

. pre menopausal

. sought care in the previous year for
pelvic pain,
abnormal uterine bleeding, and/or
fibroids

. no cancer of the reproductive tract

. never had a hysterectomy

. English or Spanish speaker

Study design

SOPHIA: Study of Pelvic Problems, Hysterectomy & Intervention Alternatives

<u>Basic design features</u>

. two cohorts: 1998/1999 (n=761) v 2003/2004 (n=732)

. interviewed every 6 months; maximum follow-up approximately 8 years

. aged 31-54 (mean = 42.5)

. I focus on data from annual interviews

The data: sub-samples and outcome

<u>Part 1: pre- and post-surgical HRQOL trajectories</u>

. Focus on n=168 women who had a hysterectomy during the study

. Observation times are centered around the time of surgical intervention

<u>HRQOL measure</u>

. *PRPP:* perceived resolution of pelvic problems
(1=not at all, 2=somewhat, 3=mostly, 4=completely)

The data: Part 2 sub-sample and outcomes

Part 2: associations between sexual functioning and HRQOL trajectories

. Focus on *n*=675 Cohort 2 women who consistently reported being sexual active across their baseline, year 1, and year 2 assessments.

. Estimate associations between intra-person trajectories of
HRQOL and
sexual functioning

HRQOL measures

PCS: physical functioning, role-related physical, bodily pain, health perception

MCS: role-related emotional, vitality, mental health, social function

Body Image: frequency of feeling feminine, good about one's body, physically unattractive, and sexually attractive

PRPP: perceived resolution of pelvic problems (1=not at all, 2=somewhat, 3=mostly, 4=completely)

The data (cont.): Part 2 sub-sample and outcomes

Part 2: associations between sexual functioning and HRQOL trajectories

Sexual functioning measures

SHOW-Q: a measure of sexual functioning (with or without a partner).

. Questions asked about the prior 4 weeks.
. 5-point response options

Satisfaction: 'How satisfied in general have you been with your ability to have and enjoy sex (with or without a partner)?' (α=.77)

Orgasm: 'When you had sexual activity, how much of the time did you experience orgasm?' (α=.84)

Desire: 'How often did you desire sex (with or without a partner)?' (α=.73)

Pelvic Interference: 'To what extent have your pelvic problems, overall, interfered with your normal or regular sexual activity (with or without a partner)?' (α=.80)

For all variables, higher scores reflected higher levels of functioning

Outline

Part 1: pre- and post-surgical HRQOL trajectories

1a. Linear random coefficient models for repeated measures (growth curve models)

1b. Smoothing growth data

- Basic spline models
- Smoothing splines via linear mixed models
- Example applications

1c. Linear spline models with correlated random effects

- Introduction to the model
- Example application

1d. Review

Outline of models considered: Part 2

Part 2: associations between sexual functioning and HRQOL trajectories

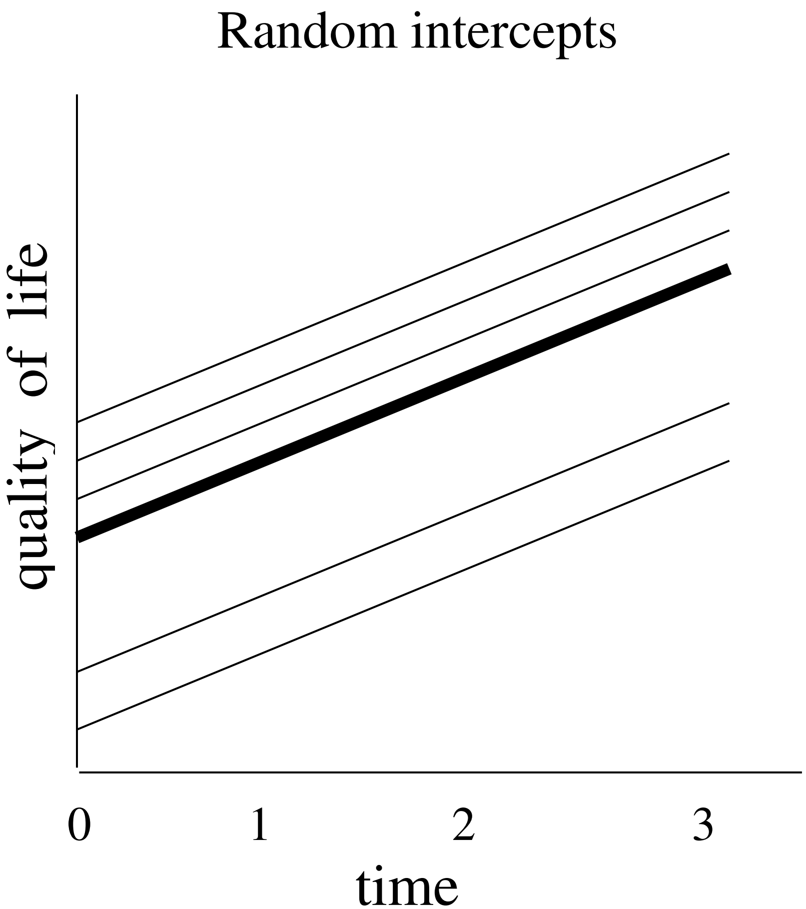
. Structural equation modeling framework for growth models

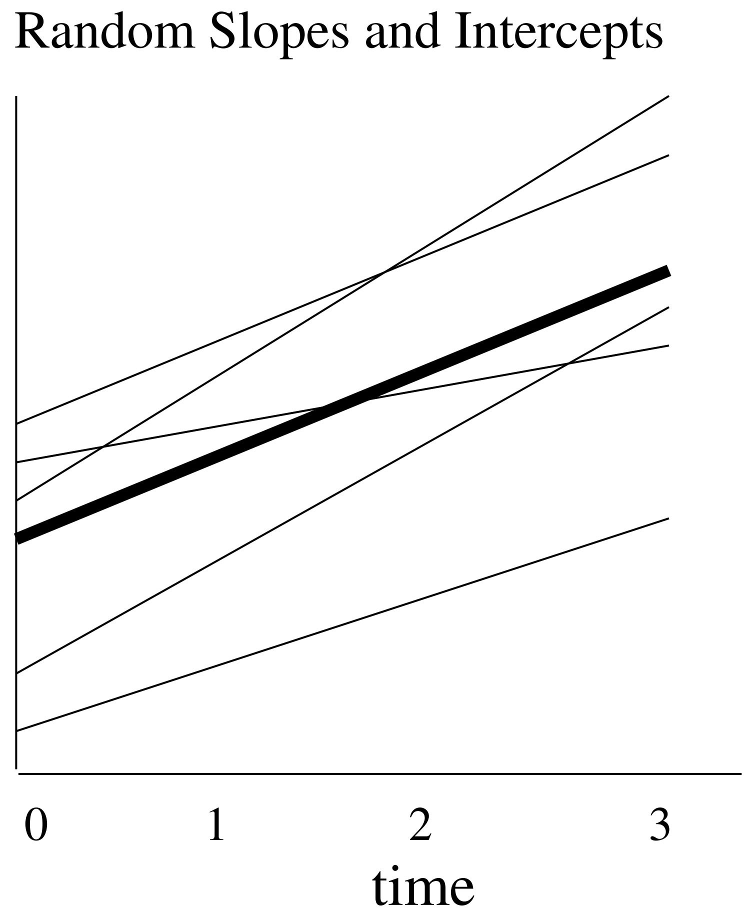
So-called, 'latent growth curve models', and

Associative latent growth model (SEM)

Part 1a: Linear random coefficient models for repeated measures

Cartoon examples of random coefficient models for repeated measures





. Allow for correlated random effects

Part 1a: Linear growth curve model

The random intercepts and slopes model for repeated measures is a linear growth curve model

Level-1 or time-level model

$Y_{ij} = B_{0j} + B_{1j}Time_{ij} + e_{ij}$, *i* indexes time and *j* indexes patients

Level-2 or patient-level models

$B_{0j} = \gamma_{00} + w_{0j}$, (mean outcome at time=0 for the *j*th patient)

$B_{1j} = \gamma_{10} + w_{1j}$, (effect of a one year increase for the *j*th patient)

Combined Model

$$Y_{ij} = \gamma_{00} + \gamma_{10}Time_{ij} + w_{0j} + w_{1j}Time_{ij} + e_{ij}$$

Part 1a: Linear growth curve model

Combined Model

$$Y_{ij} = \gamma_{00} + \gamma_{10}\text{Time}_{ij} + w_{0j} + w_{1j}\text{Time}_{ij} + e_{ij}$$

where

$\text{VAR}(u_{0j}) = \tau_{00}$,
$\text{VAR}(u_{1j}) = \tau_{11}$,
$\text{COV}(u_{0j}, u_{1j}) = \tau_{01}$, and

$$\text{VAR}\begin{bmatrix} u_{0j} \\ u_{1j} \end{bmatrix} = \begin{bmatrix} \tau_{00} & \tau_{01} \\ \tau_{01} & \tau_{11} \end{bmatrix}$$

Part 1b: Smoothing growth data

Linear growth/change across time is unlikely to obtain in many contexts
Need to allow for a non-linear trajectory

One approach is to explore trajectory shape in the data and
modify the linear model

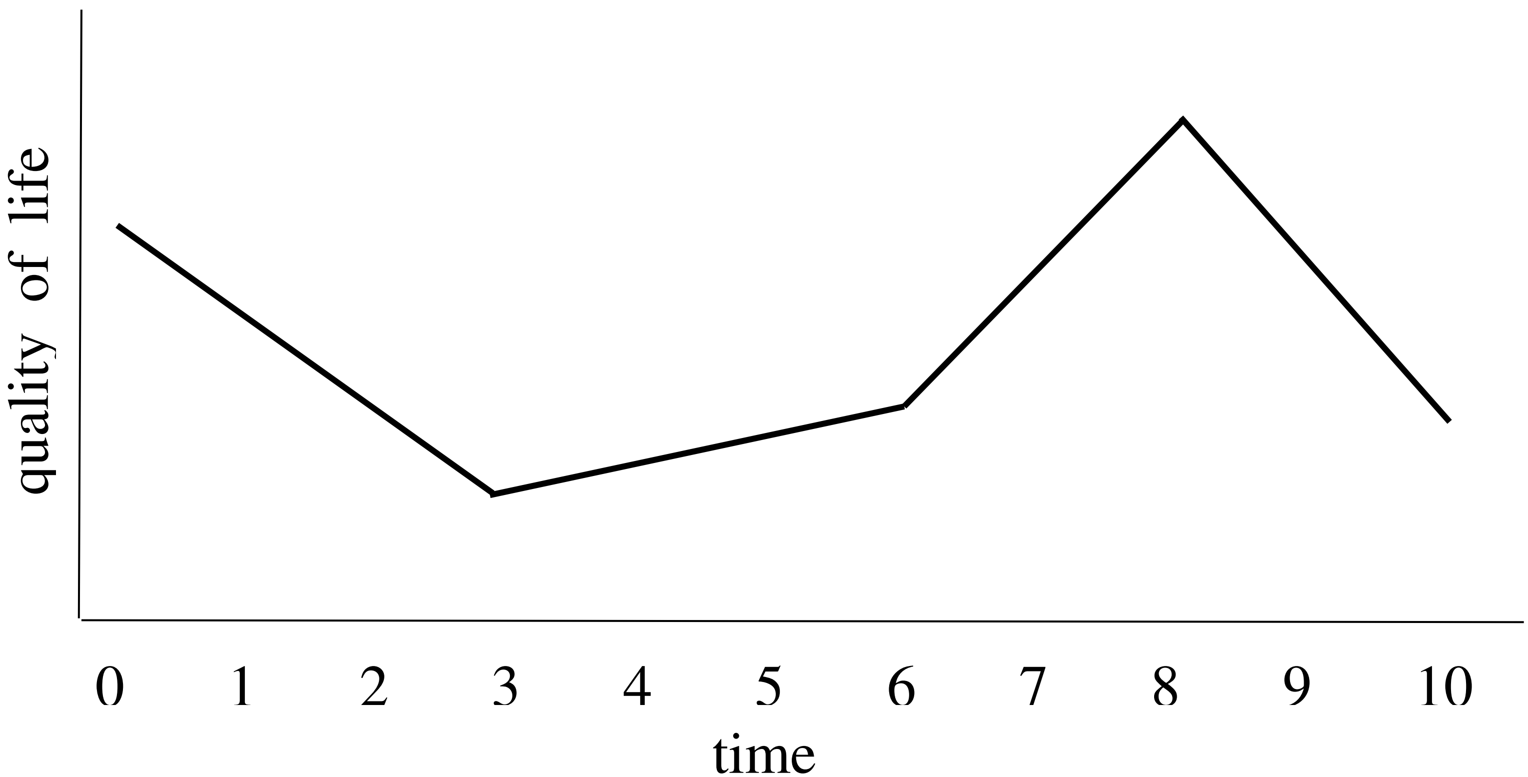
Such exploration can be accomplished with smoothing regression routines
such as loess regression

Typically, software for those routines does not explicitly accommodate
correlated responses (e.g., repeated measures)

Failure to consider non-independence of observations can result in
under-smoothing

To move forward, I will first introduce spline function models

Part 1b: Basic linear spline function example



The x-axis is divided into intervals—here at times 3, 6, and 8.

The interval endpoints at times 3, 6, and 8 are called 'knots'

Part 1b: Linear spline function example (continued)

. knots, κ_k, are at times 3, 6, and 8

. The linear spline function is

$$f(t) = \beta_0 + \beta_1 t + \sum_{k=1}^{K} u_k \left(t - \kappa_k\right)_+$$

where the linear spline basis function is

$$(t - \kappa_k)_+ = 0, \qquad t \le \kappa_k$$

$$t - \kappa_k, \ t > \kappa_k$$

Example: linear spline function with knots at times 3, 6, and 8

$$f(t) = \beta_0 + \beta_1 t + u_1 (t-3)_+ + u_2 (t-6)_+ + u_3 (t-8)_+$$

Part 1b: Linear spline function example

Integer time values ranging from 0 through 10 and knots at times 3, 6, and 8,

$$f(t) = \beta_0 + \beta_1 t + u_1(t-3)_+ + u_2(t-6)_+ + u_3(t-8)_+$$

The linear spline basis variables, $s_k = (t - \kappa_k)_+$, would equal

t	$s_3 = (t-3)_+$	$s_6 = (t-6)_+$	$s_8 = (t-8)_+$
0	**0**	**0**	**0**
1	**0**	**0**	**0**
2	**0**	**0**	**0**
3	**0**	**0**	**0**
4	**1**	**0**	**0**
5	**2**	**0**	**0**
6	**3**	**0**	**0**
7	**4**	**1**	**0**
8	**5**	**2**	**0**
9	**6**	**3**	**1**
10	**7**	**4**	**2**

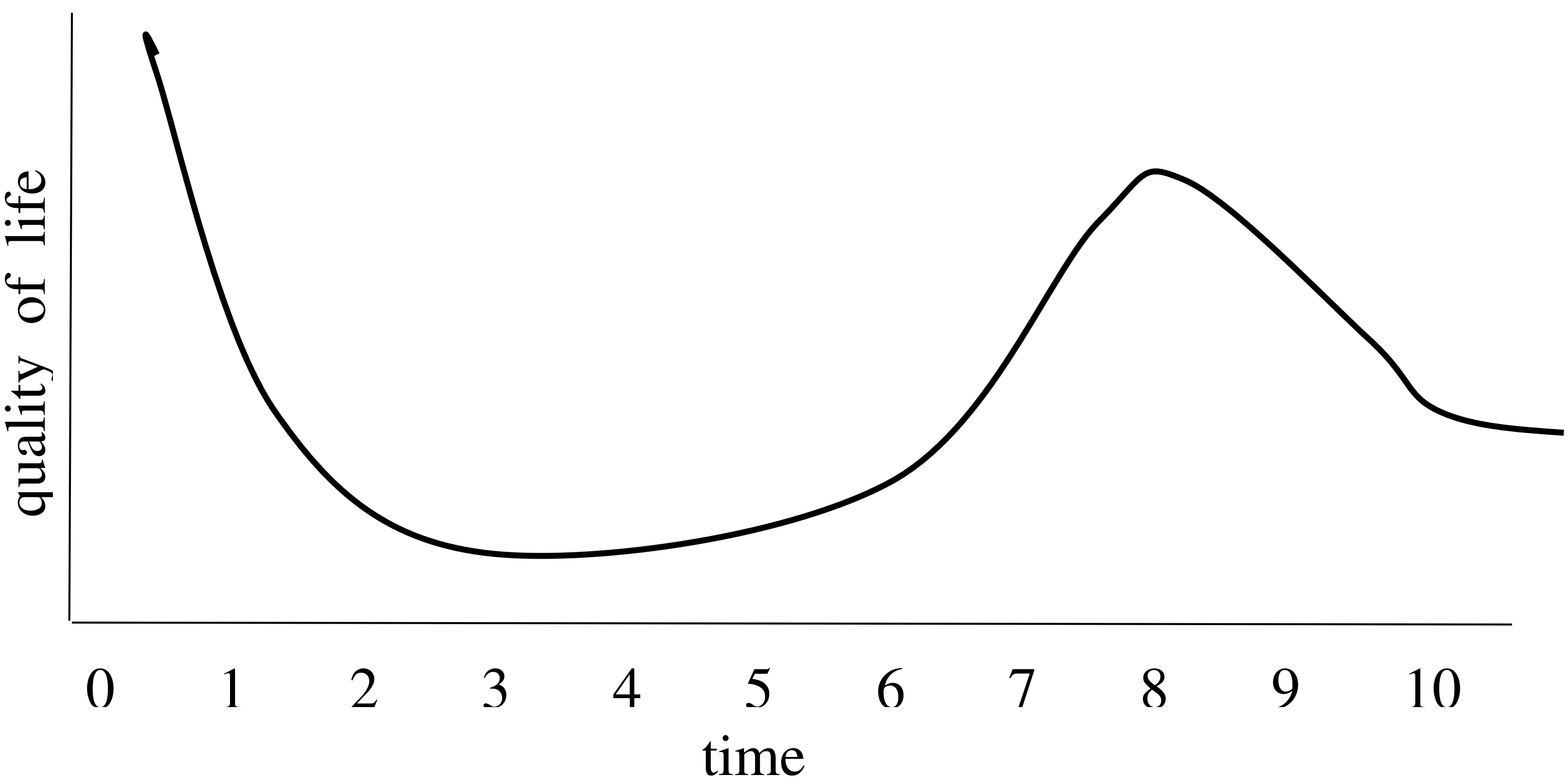
Part 1b: Quadratic spline function

Again, the linear spline function

$$f(t) = \beta_0 + \beta_1 t + \quad u_1(t-3)_+ + u_2(t-6)_+ + u_3(t-8)_+$$

The quadratic spline function

$$f(t) = \beta_0 + \beta_1 t + \beta_2 t^2 + u_1(t-3)_+^2 + u_2(t-6)_+^2 + u_3(t-8)_+^2$$



Part 1b: Smoothing spline regression via a linear mixed model

$$y = \mathbf{X\beta} + \mathbf{Zu} + \boldsymbol{\varepsilon}$$

where,
X is the fixed effects design matrix,
Z is the random effects design matrix, and

$$\begin{bmatrix} \mathbf{u} \\ \boldsymbol{\varepsilon} \end{bmatrix} \sim \mathrm{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_u^2 & 0 \\ 0 & \sigma_\varepsilon^2 \end{bmatrix} \right)$$

The **u** are estimated from a single distribution, and
the automatic smoothing parameter equals $\sigma_\varepsilon^2 / \sigma_u^2$

. Within the mixed models framework, you can model correlated responses with a random intercept or patterned residual covariance structure

Part 1b: Smoothing via mixed models

Following the above example, calculate the spline basis variables, s_k, as

$(t-3)_+$, $(t-6)_+$, and $(t-8)_+$

(square those quantities to fit a quadratic regression spline model)

Model the outcome as a function of the fixed effects of time and random effects of the spline basis variables

```
PROC MIXED DATA=mydata METHOD=REML;
 CLASS id ;
 MODEL hrqol = t  t*t / SOLUTION OUTPRED=SMOOTH;
 RANDOM s1 s2 s3 / TYPE=TOEP(1);
 RANDOM INTERCEPT / TYPE=UN SUBJECT=id;
 RUN;
```

where

TYPE=TOEP(1) estimates a single variance component that is shared by all random effects

Part 1b: Example 1--Intro
Effect of hysterectomy on perceived resolution of pelvic problems

Related research questions

. Among women with non-cancerous uterine conditions, to what extent does a hysterectomy affect their perceptions that their pelvic problems have been solved?

What are patients' trajectories of their perceived pelvic problem resolution in the years before and after hysterectomy?

Do patients' pre-surgical health perceptions affect the level of perceived problem resolution that is attributable to the surgical intervention?

Do patients' pre-surgical health perceptions affect whether post-surgical improvements in perceived health are maintained in the years after surgery?

The goal is to fit a reasonable growth model that can address these questions

Part 1b: Example 1—Smoothing via mixed models
Effect of hysterectomy on perceived resolution of pelvic problems

n=168 women had a hysterectomy during the study period
time of hysterectomy was determined by self-report and chart review

The mean number of annual study observations was 5.45 (range 2 - 9)
e.g., baseline plus 4.45 years of follow-up

Because women who enrolled in the study were heterogeneous wrt, age, symptom duration, and symptom severity, time-on-study was not a very meaningful metric.

Therefore, I set each woman's hysterectomy date to year=0

The resulting 'time from hysterectomy' variable, t_H, ranged from about −7.5 to +8.0 years.

However, the extreme time points were sparsely represented.
Therefore, I only modeled data observed on the interval $-3 \leq t_H \leq 6$

Part 1b: Example 1—Smoothing via mixed models

Effect of hysterectomy on perceived resolution of pelvic problems

. 8 knots chosen along $t_{H,}$ at times –2, –1, 0, 1, 2, 3, 4, and 5

. Corresponding to each knot, I created a quadratic spline basis variable (s_1 through s_8), e.g., $s_8 = (t_H - 5)_+^2$

Outcome variable:

Perceived resolution of pelvic problems (PRPP)

(1=not at all, 2=somewhat, 3=mostly, 4=completely)

. Fit (a) a fixed effect quadratic spline model and

(b) a random effect smoothing spline model

Both models included a binary indicator representing hysterectomy status where b=0 for pre-hysterectomy observations and b=1, post-hysterectomy

e.g., the fixed effect quadratic spline model,

$$f(t_H) = \beta_0 + \beta_1 t_H + \beta_2 t_H^2 + \beta_2 b + \sum_{k=1}^{K} u_k \left(t_H - \kappa_k\right)_+^2$$

Part 1b: Example 1—Smoothing via mixed models

Effect of hysterectomy on perceived resolution of pelvic problems

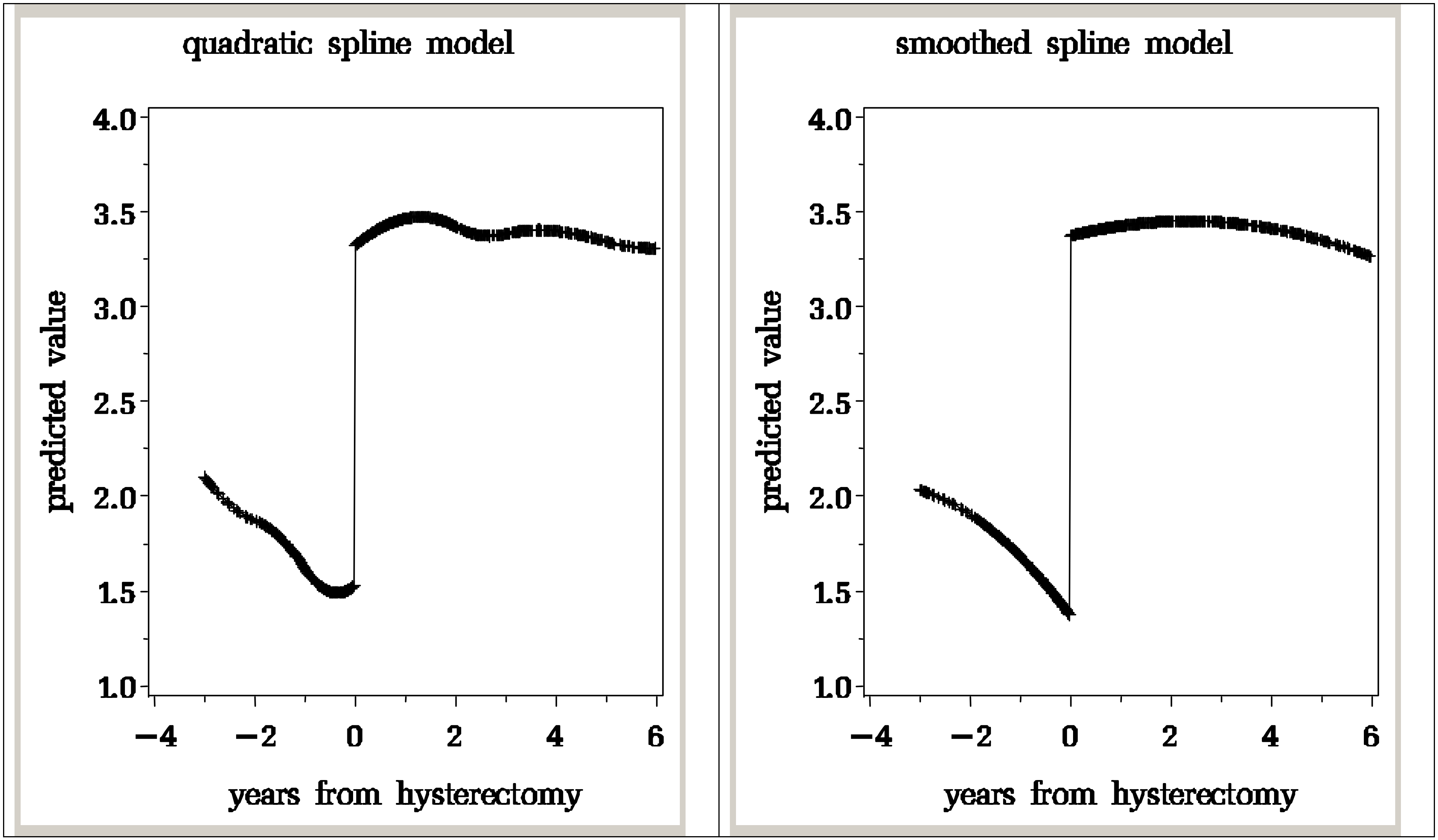
These models smoothed the pre-hyst and post-hyst trajectories individually,

e.g., for the smoothing spline regression model

```
PROC MIXED DATA=mydata METHOD=REML;
  CLASS id  b;
  MODEL prpp = b  tH(b)  tH* tH(b) / S OUTPM=SMOOTH;
  RANDOM s1 – s8      / TYPE=TOEP(1);
  RANDOM INTERCEPT / TYPE=UN SUBJECT= id;
  RUN;
```

Part 1b. Smoothed models

Perceived resolution of pelvic problems: pre- and post-hysterectomy
(1=not at all; 2=-somewhat; 3=mostly; 4=completely)

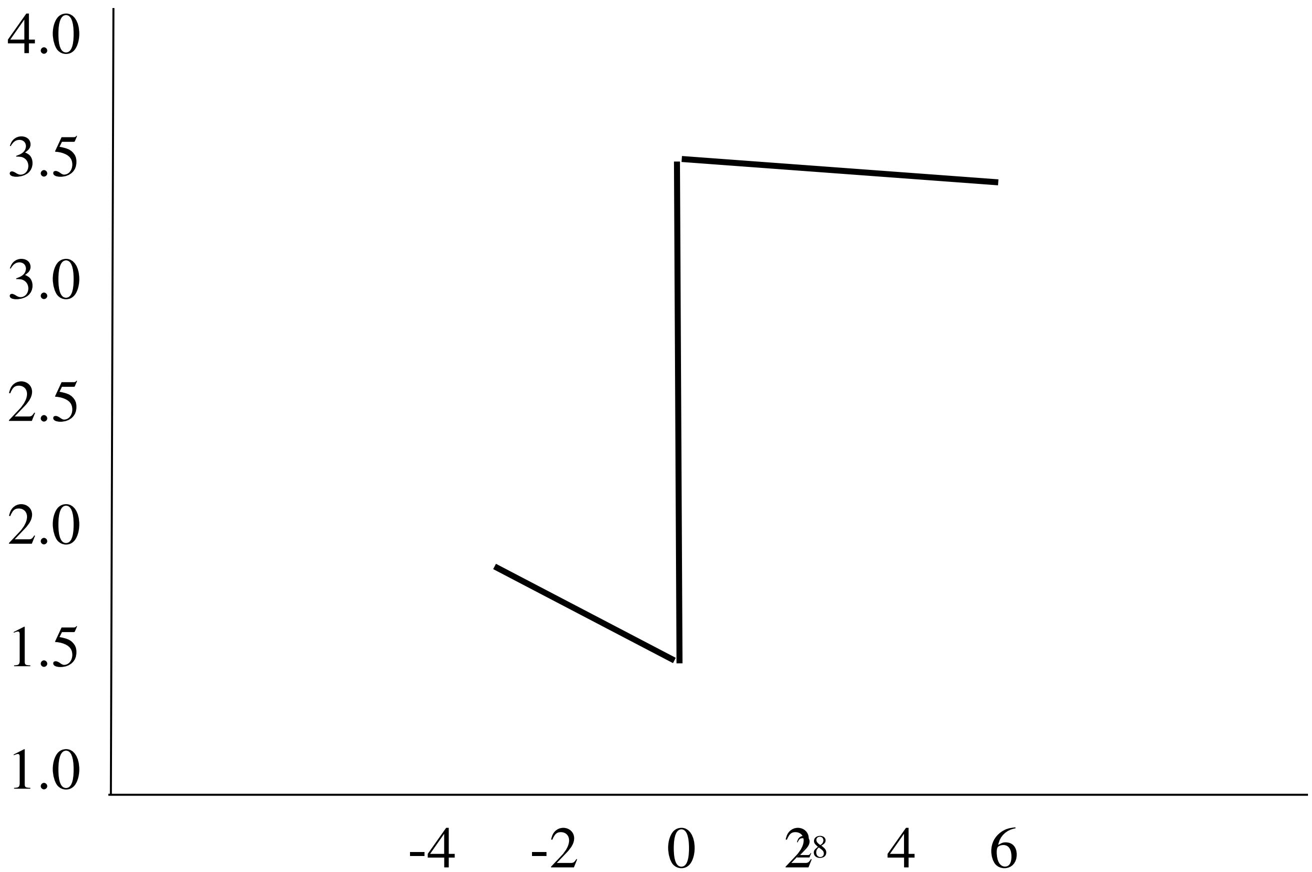


Part 1c: Example application

A one-knot linear spline model with a 'bump' & random effects

Effect of hysterectomy on perceived resolution of pelvic problems

The smoothed plots suggested that the overall trajectory might be approximated by a linear spline function with one knot as well as a 'bump' at the time of hysterectomy (time=0).



Part 1c: Example application

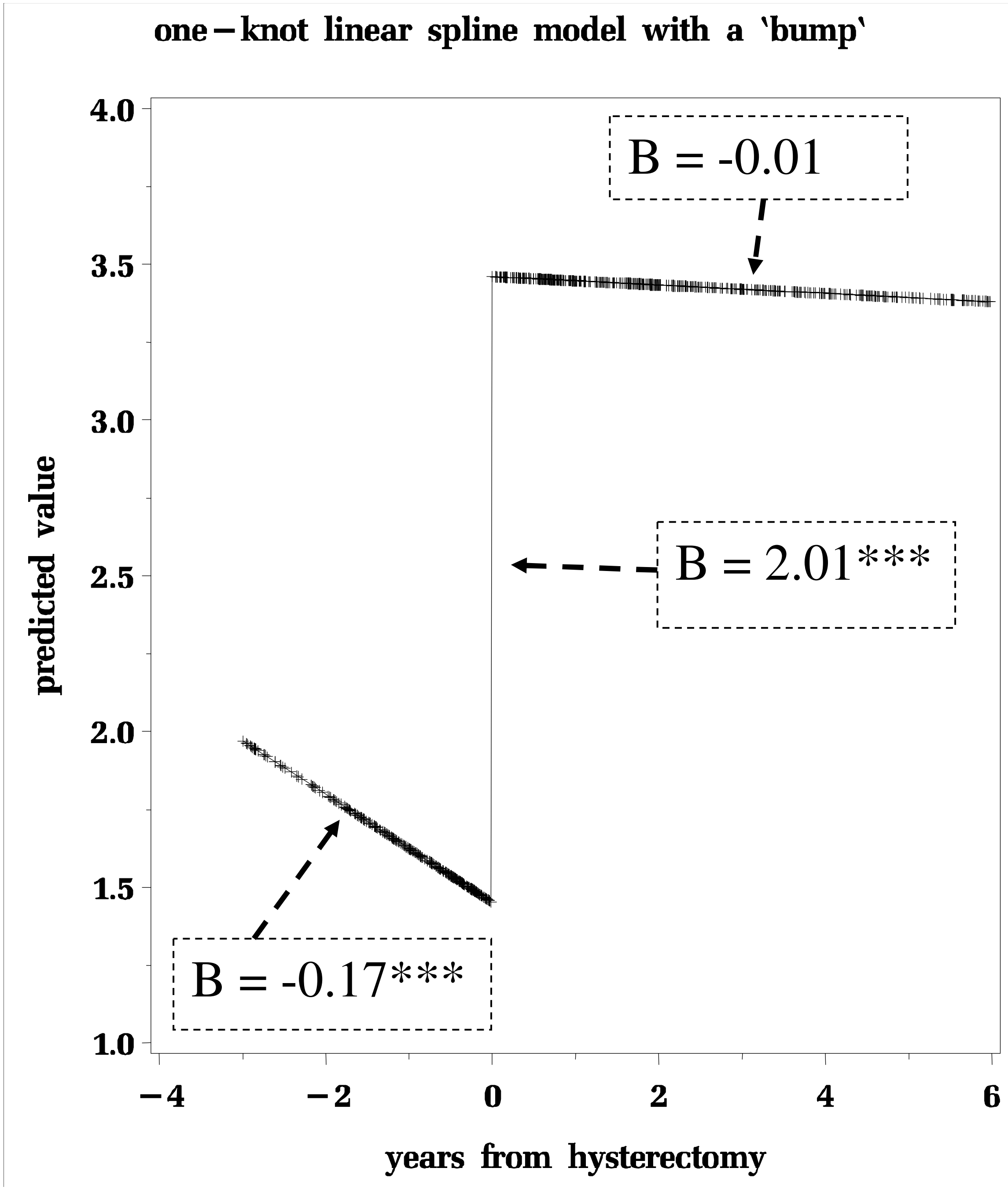
A one-knot linear spline model with a 'bump' & random effects

Effect of hysterectomy on perceived resolution of pelvic problems

```
PROC MIXED METHOD=REML DATA=mydata;
 CLASS id _b_;
 MODEL prpp =  b  tH(_b_) /S  OUTPM=OUT;
  RANDOM INTERCEPT b  / SUBJECT=id TYPE=UN;
RUN;
```

where, _b_ is identical to b, but is defined as a CLASS variable.

* I fit random effects for pre- and post-surgery trajectories, and allowed them to covary with all respondent-level random effects. Those effects were non-significant and were dropped from the model.



Component	Est.
VAR(int)	0.25
VAR(bump)	0.59
COV(i,b)	-0.26
	(r = -0.69)
VAR(res)	0.35

---

post- v pre-HYST slopes, $\Delta = 0.16^{**}$

---

intercept = 1.45

Part 1: Review

Smoothing splines w/in mixed models framework can smooth longitudinal 'growth' data and accommodate correlated responses within individuals

The smoothed data can guide selection of a simpler linear spline models

Such models may provide adequate approximations and
will be more accessible to some audiences.

When observation times are fixed, a repeated measures approach can
model additional complexity in covariation among residuals

In the worked example, variation in slopes (linear spline segments)
was very low. Most between-patient variation was captured
by the random intercept and 'bump' terms.

Part 2: Associations between sexual functioning and HRQOL outcomes

<u>SEM approach to fitting growth curve models</u>

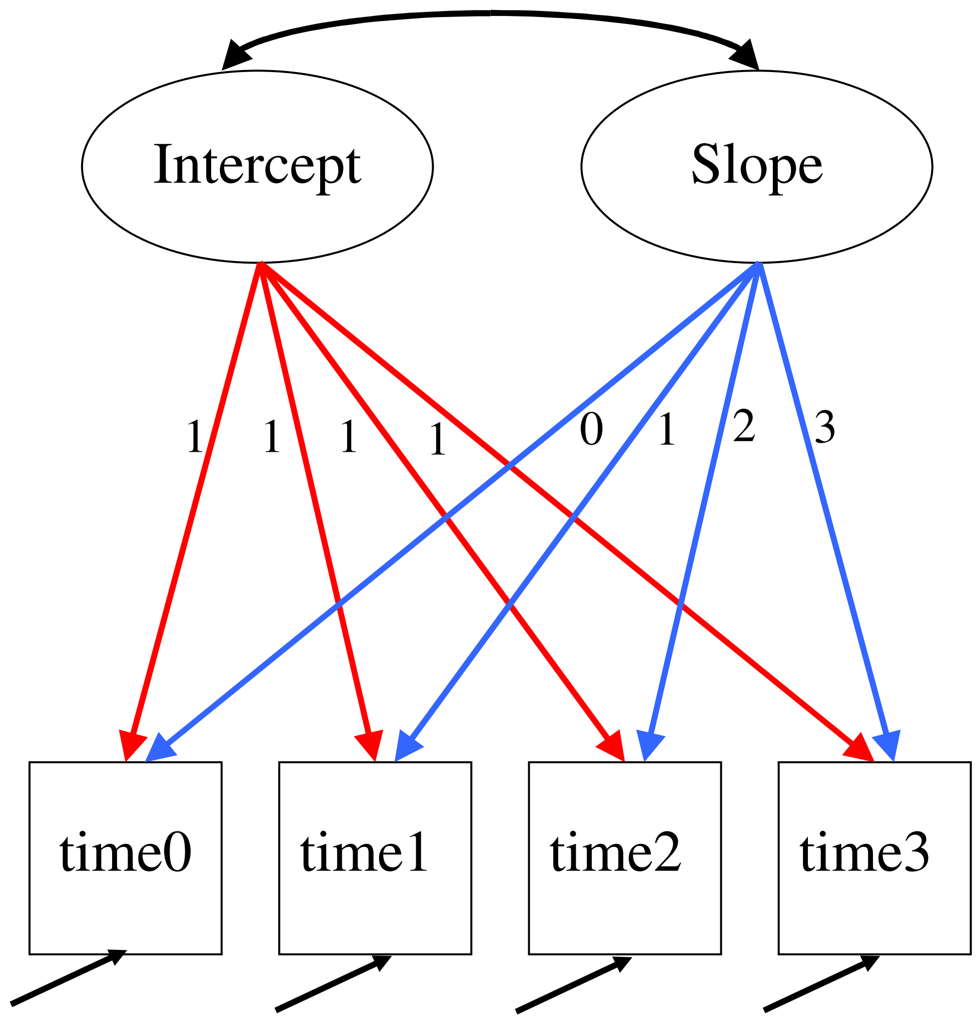
. Latent growth curve models
Random intercepts and slopes are conceptualized as inter-correlated latent variables

. Most commonly,
model fixed measurement occasions
use observed covariance matrix and mean vector as input data

Extensions, via FIML, to arbitrary measurement occasions and raw data input

Part 2: SEM approach to fitting growth curve models

A linear 'latent' growth curve (LGC) model



repeated quality of life assessments

Part 2: SEM approach to fitting growth curve models

$$\text{COV}(\mathbf{X}) = \mathbf{\Lambda}_x \mathbf{\Phi} \mathbf{\Lambda}_x' + \mathbf{\Theta},$$

where

. $\mathbf{\Lambda}_x$ holds the constant and slope coefficients,

. $\mathbf{\Phi}$ holds the covariances among random intercepts and slopes, and

. $\mathbf{\Theta}$ holds residual variances of the *x*s

and

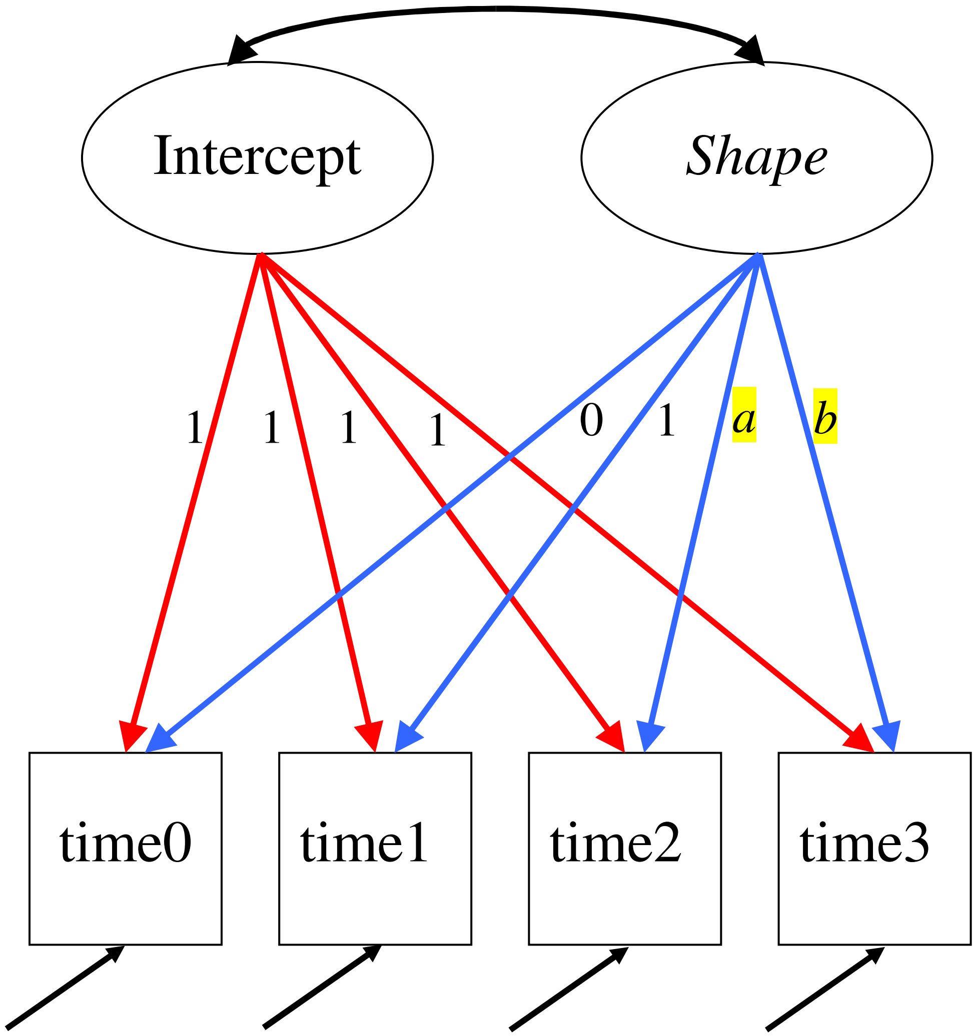
$$\text{MEAN}(\mathbf{X}) = \mathbf{\Lambda}_x \mathbf{K}$$

where $\mathbf{K}$ holds the mean intercept and slope values

Part 2: SEM approach to fitting growth curve models

LGC with optimal shape function estimates

Allowing for non-linear trajectories

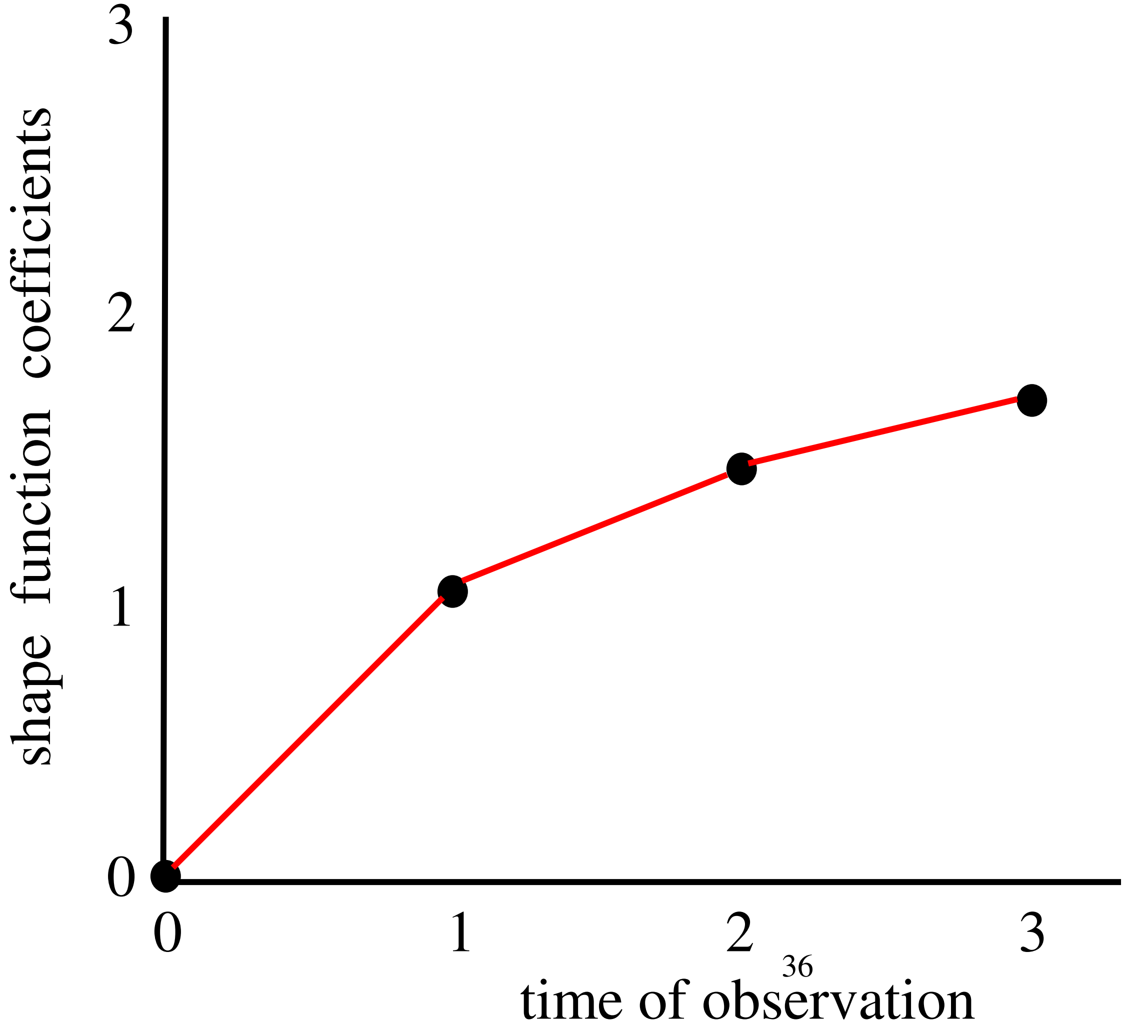


repeated quality of life assessments

Part 2: SEM approach to fitting growth curve models

LGC with optimal shape function estimates

Plot the actual time of observation against the shape function coefficients to reveal the trajectory shape



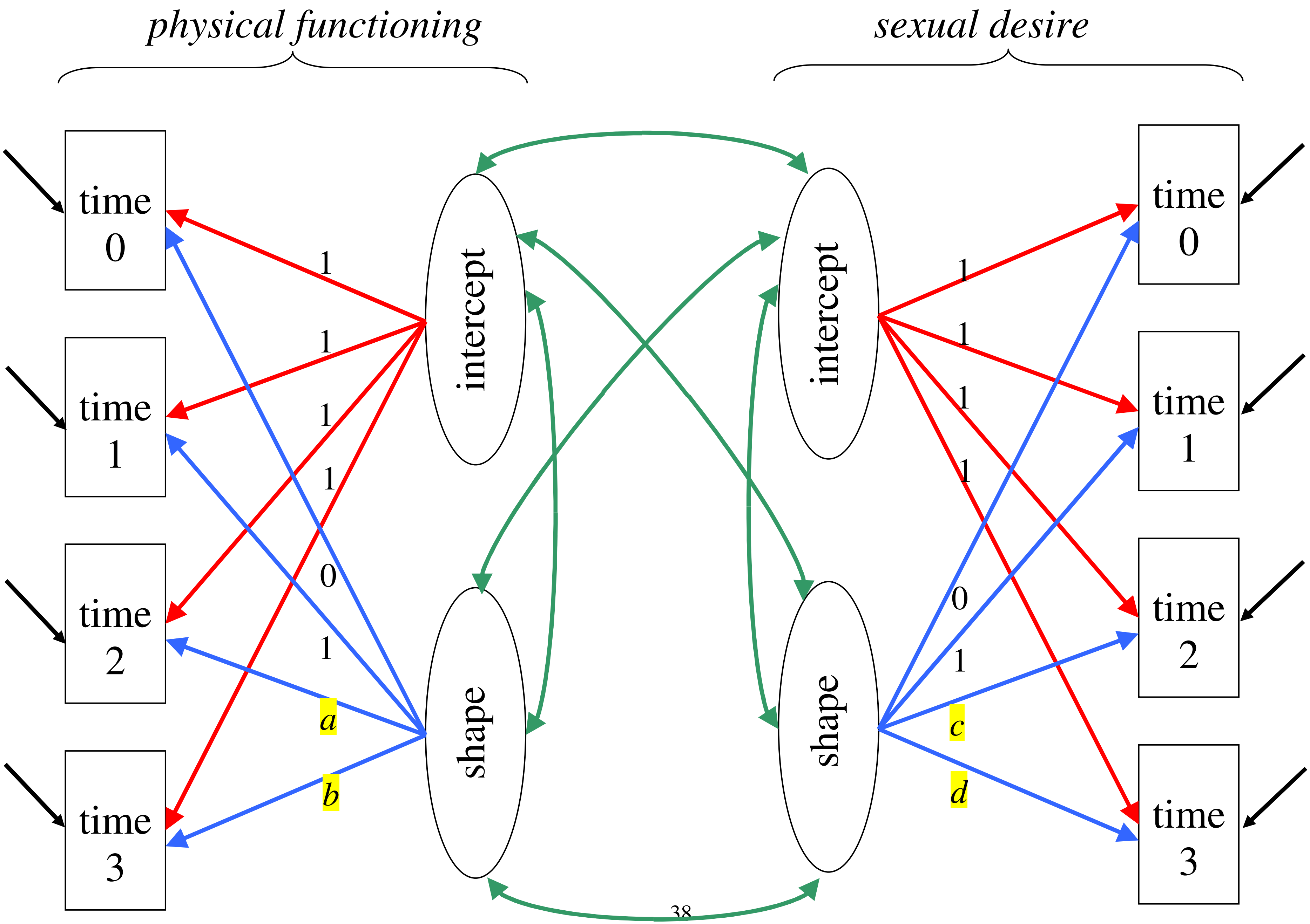
Part 2: SEM approach to fitting growth curve models

Associative latent growth curve models

. Simultaneously model growth curves for multiple outcomes

. Estimate inter-outcome covariation of intercepts and trajectories

Part 2: *Associative* latent growth model w/ optimal shape function est.



Part 2: *Associative* latent growth model with optimal shape function est.

Example application using the SOPHIA data

. *n*=675 SOPHIA cohort II women who were consistently sexually active across the baseline, year 1, and year 2 assessments

Related research questions

. Among women with non-cancerous uterine conditions, to what extent do *changes* in sexual functioning correlate with *changes* in other measures of HRQOL?

Part 2: *Associative* latent growth model with optimal shape function est.

Example application using the SOPHIA data

. Sexual functioning measures

Satisfaction: 'How satisfied in general have you been with your ability to have and enjoy sex (with or without a partner)?' (α=.77)

Orgasm: 'When you had sexual activity, how much of the time did you experience orgasm?' (α=.84)

Desire: 'How often did you desire sex (with or without a partner)?' (α=.73)

Pelvic Interference: 'To what extent have your pelvic problems, overall, interfered with your normal or regular sexual activity (with or without a partner)?' (α=.80)

Part 2: *Associative* latent growth model with optimal shape function est.

<u>Example application using the SOPHIA data</u>

. HRQOL measures

PCS: physical functioning, role-related physical, bodily pain, health perception

MCS: role-related emotional, vitality, mental health, social function

Body Image: frequency of feeling feminine, good about one's body, physically unattractive, and sexually attractive

PRPP: perceived resolution of pelvic problems
(1=not at all, 2=somewhat, 3=mostly, 4=completely)

Part 2: *Associative* latent growth model with optimal shape function est. Example application using the SOPHIA data

Correlations between sexual functioning and quality of life *trajectories*

	Satisfaction	Orgasm	Desire	Pelvic Interf.
Orgasm	0.693			
Desire	0.857	0.843		
Pelvic interf.	0.822	0.628	0.607	
PCS	0.058	0.063	0.121	0.581
MCS	0.448	0.290	0.196	0.154
Body Image	0.597	0.598	0.407	0.201
PRPP	0.223	0.037	0.028	0.732

Part 2: LGC versus growth curves fit w/in the mixed models framework

. LGC can estimate optimal growth coefficients for fixed
measurement occasions

. LGC can estimate associative growth models
Associative growth models may be possible within the mixed models framework, by specifying a multivariate outcomes models (?)

. Both can accommodate non-uniform measurement occasions

The End