Repeated measures models with multiple, correlated random effects

CAPS Methods Core

January 11, 2008

Steve Gregorich

Outline

Introduction to example data

Part 1

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

. Smoothing via linear mixed models

. More restricted spline models with correlated random effects

I will describe spline models of pre- and post-hysterectomy health-related quality of life (HRQOL) trajectories as well as the 'instantaneous' HRQOL change (or 'bump') attributable to the surgical intervention.

These models include correlated random intercept, trajectory, and 'bump' effects, which can address interesting research questions: e.g., are women's pre-surgical HRQOL trajectories associated with the 'instantaneous' HRQOL changes that are attributable the surgical intervention?

Overview

Part 2

. SEM approach to fitting growth models
    (latent growth curve models)

. Associative latent growth model (SEM)

These models allows for estimation of covariation between temporal changes (trajectories) in multiple dimensions. Such models are illustrated using longitudinal assessments of women's' self-reported sexual, physical, and mental health (e.g., are changes in sexual functioning associated with changes in mental health status?).

Introduction to example data

SOPHIA: Study of Pelvic Problems, Hysterectomy & Intervention Alternatives

Prospective observation study: N = 1493 women

<u>Eligibility criteria</u>

. 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

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

. interviewed every 6 months.

. aged 31-54 (mean = 42.5)

Introduction to example data

SOPHIA HRQOL (health-related quality of life) 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

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

Introduction to example data (cont.)

SOPHIA 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)

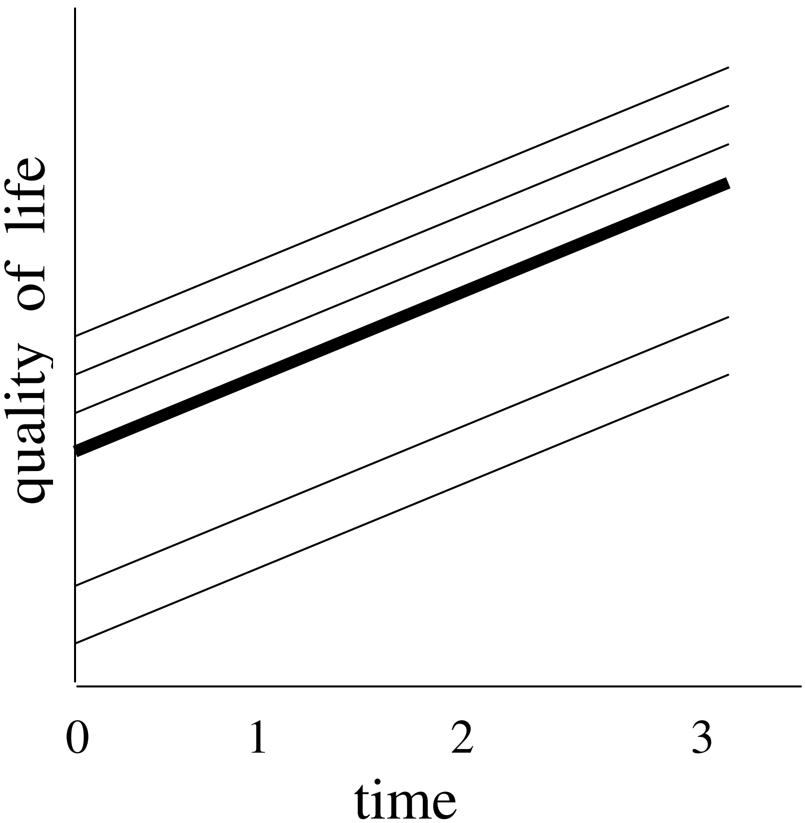
Pelvic Interference: 'To what extent have your pelvic problems, overall, interfered with your normal or regular sexual activity (with or without a partner)?' (α=.80)

For all variables, higher scores reflected higher levels of functioning

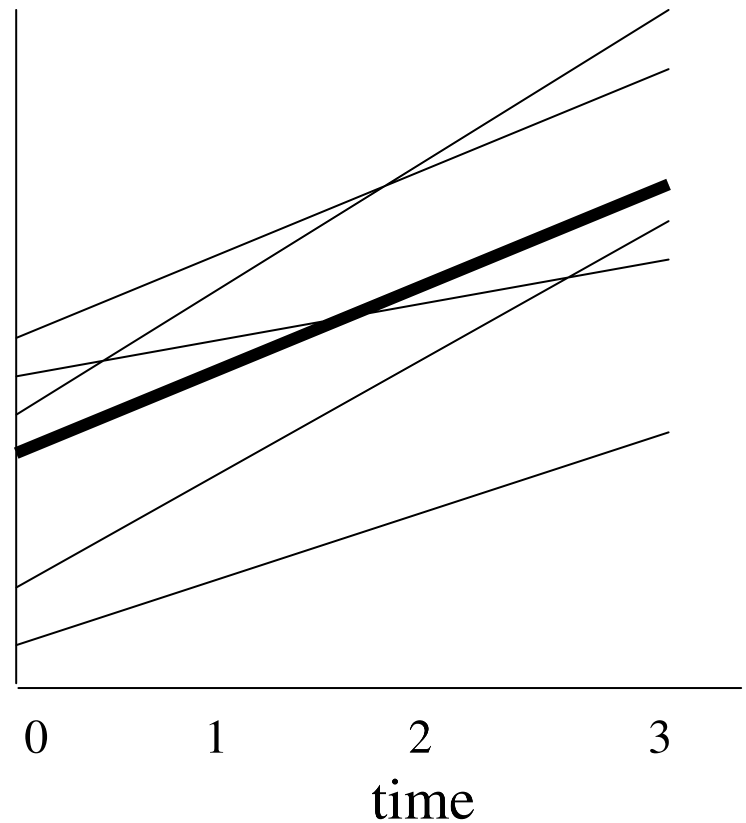
Part1: Review of linear random coefficient models for repeated measures

Cartoon examples of random coefficient models for repeated measures

Random intercepts



Random Slopes and Intercepts



. Allow for correlated random effects

Part1: 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

$$Y_{ij} = B_{0j} + B_{1j}Time_{ij} + e_{ij}$$

Level-2 or patient-level models

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

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

Combined Model

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

Part1: Linear growth curve model

Combined Model

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

where

$VAR(u_{0j}) = \tau_{00}$,

$VAR(u_{1j}) = \tau_{11}$,

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

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

Part1: 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

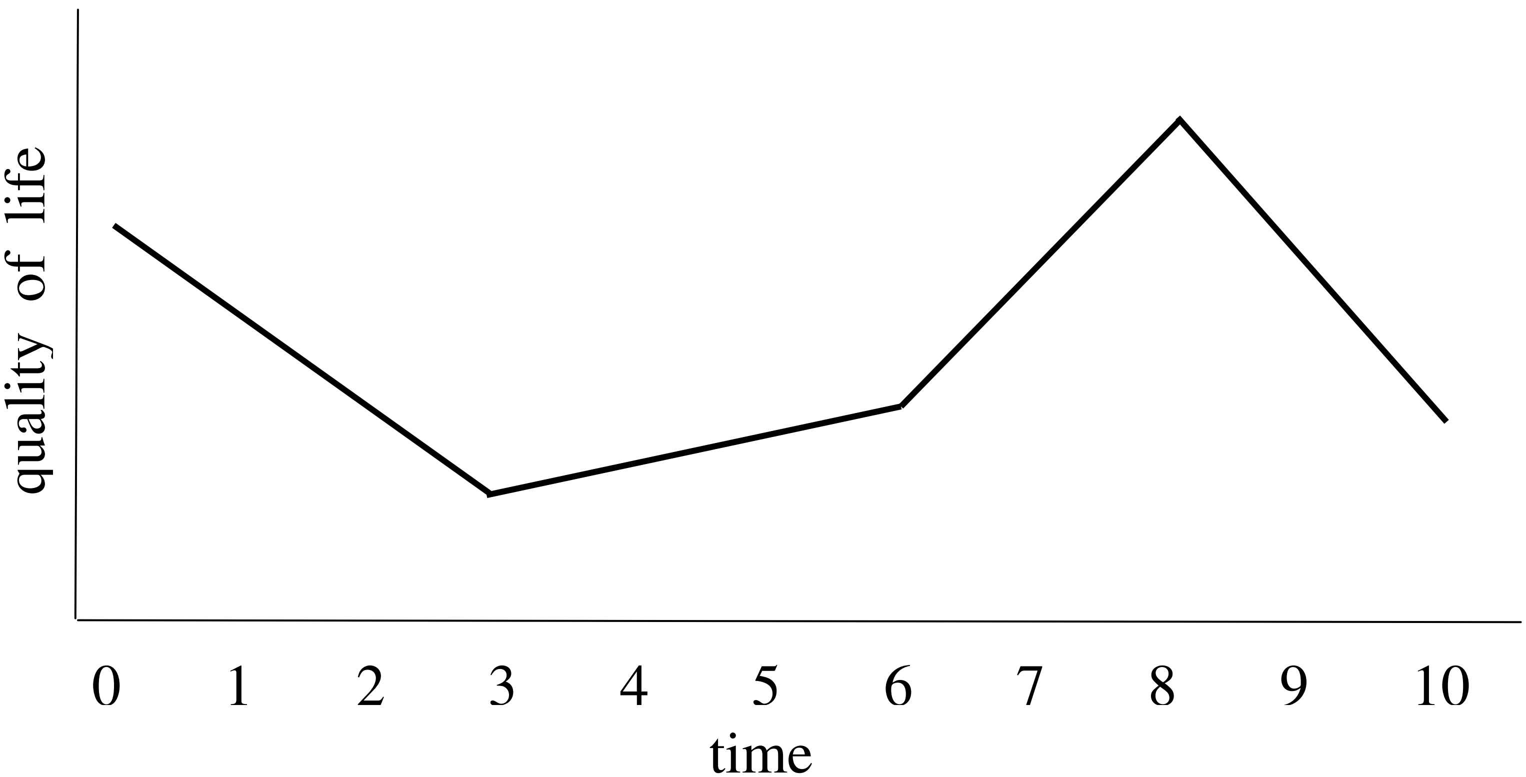
Such exploration can be accomplished with smoothing regression routines such as loess regression or thin-plate smoothing splines

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

Part1: Linear spline function



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

Part1: Linear spline function

. knots at times 3, 6, and 8

. The linear spline function is

$$f(t) = \beta_0 + \beta_1 t + \beta_2 (t-3)_+ + \beta_3 (t-6)_+ + \beta_4 (t-8)_+$$

where the linear spline basis function is

$$(u)_+ = u,\ u > 0$$

$$0,\ u \le 0$$

Part1: Linear spline function

. The linear spline function

$$f(X_1) = \beta_0 + \beta_1 X_1 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5$$

Example: with integer time values ranging from 0 through 10 and knots at times 3, 6, and 8 the values for X_1, X_3, X_4, and X_5 would equal

$$X_3 = (X_1 - 3)_+,$$

$$X_4 = (X_1 - 6)_+,$$

$$X_5 = (X_1 - 8)_+,$$

where

X_1 defines time of measurement

and

$$(u)_+ = u, u > 0$$
$$0, u \leq 0$$

(basis function)

X_1	X_3	X_4	X_5
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

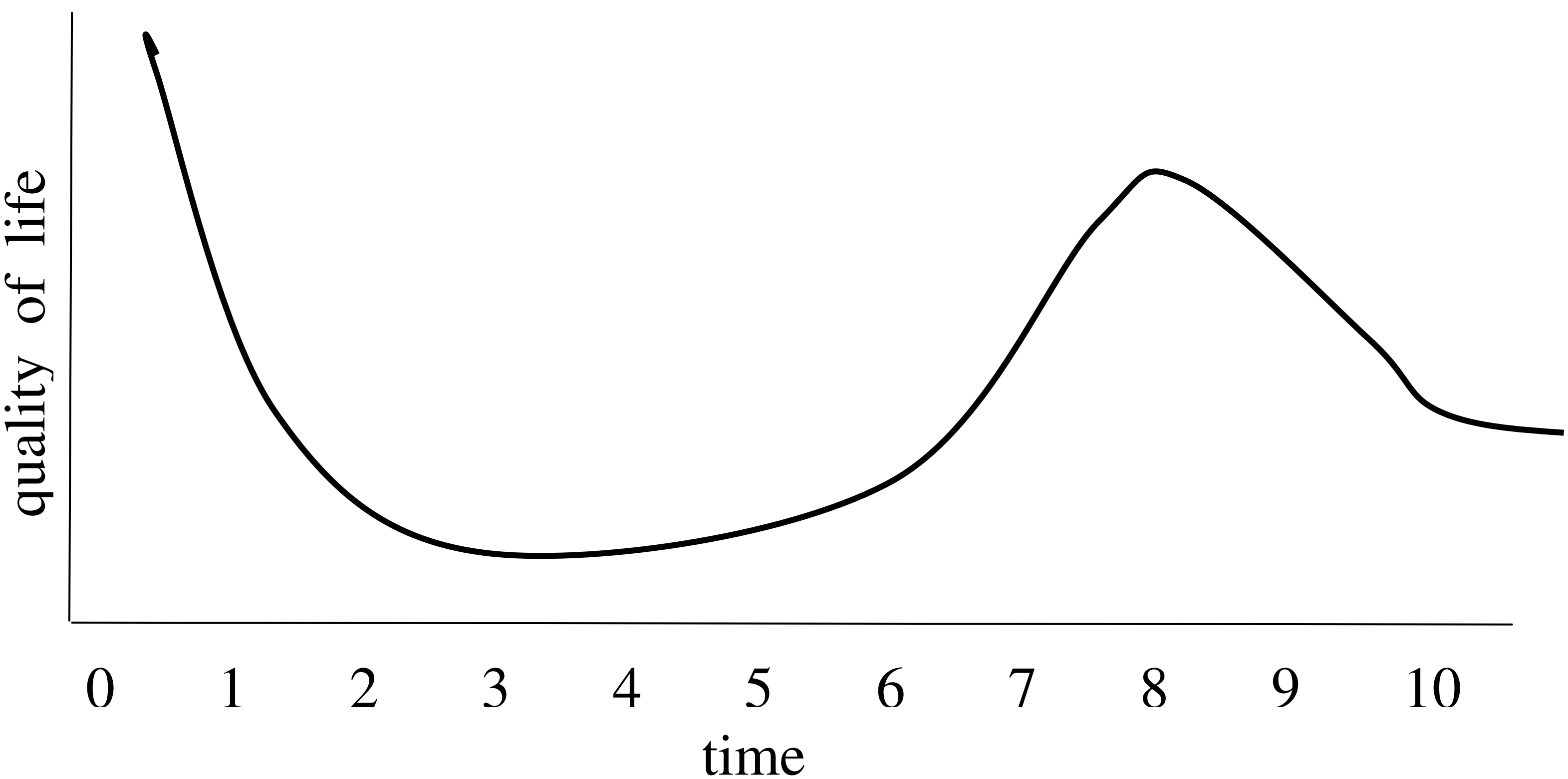
Part1: Quadratic spline function

Again, the linear spline function

$$f(X_1) = \beta_0 + \beta_1 X_1 \qquad + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5$$

The quadratic spline function

$$f(X_1) = \beta_0 + \beta_1 X_1 + \beta_2 X_1^2 + \beta_3 X_3^2 + \beta_4 X_4^2 + \beta_5 X_5^2$$



Part1: Smoothing via mixed models

Following the above example, calculate X_3, X_4, and X_5 (square those quantities to fit a quadratic regression spline model)

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

```
PROC MIXED DATA=mydata METHOD=REML;
  CLASS id ;
  MODEL qol = X1|X1 / SOULUTION OUTPRED=SMOOTH;
  RANDOM X3 X4 X5 / 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

The 'automatic' smoothing parameter equals $\sqrt{\sigma_\varepsilon / \sigma_u}$

Part1: 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

Part1: 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.

Part1: Example 1—Smoothing via mixed models

Effect of hysterectomy on perceived resolution of pelvic problems

14 knots along t_H chosen for spline model: years –7.0 through +6.0

Corresponding to each knot, I created a quadratic spline basis (sb01-sb14)

e.g., $sb14 = (t_H - 6)_+^2$

Outcome variable:

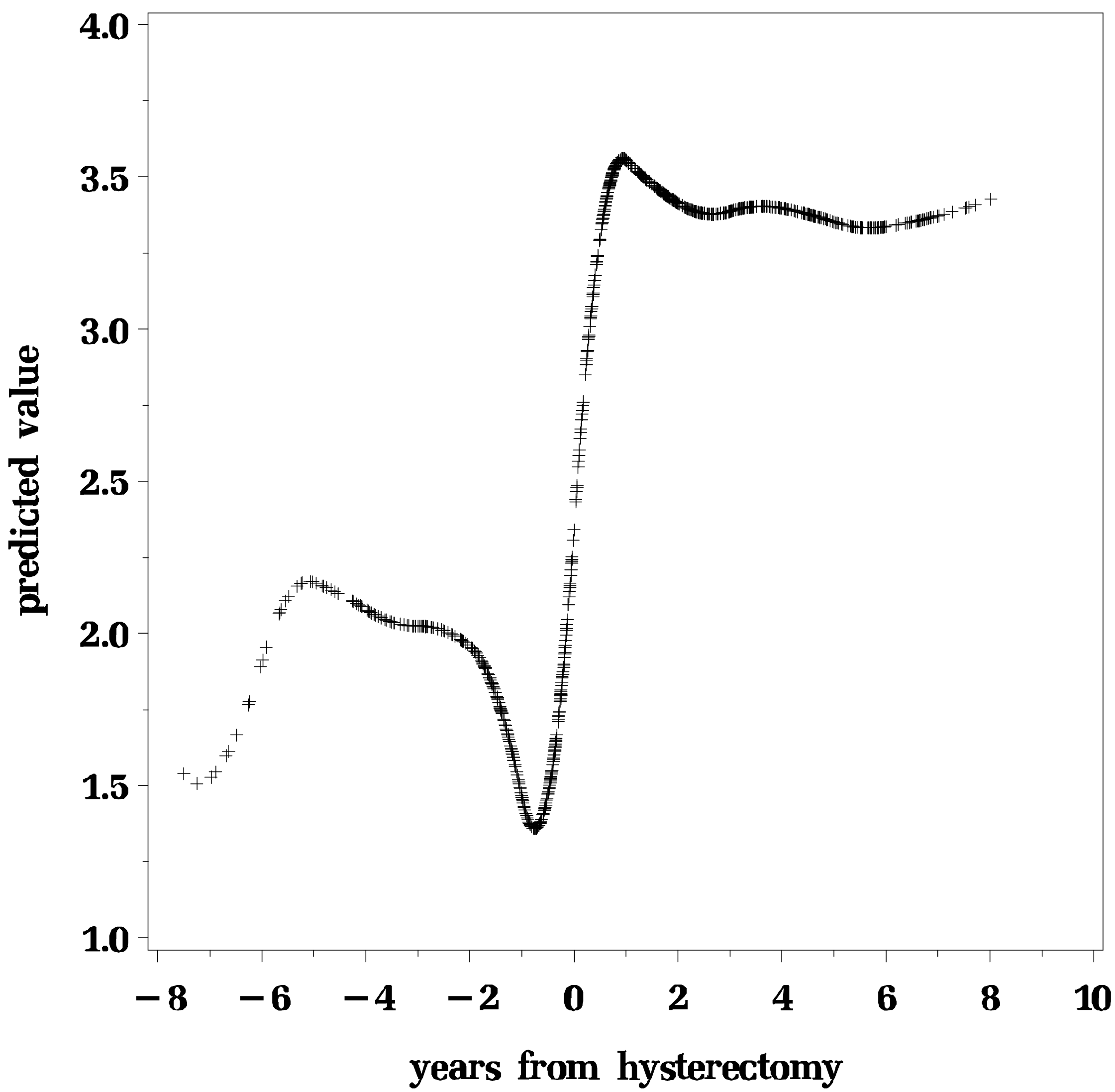
Perceived resolution of pelvic problems ('pps')

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

```
PROC MIXED DATA=mydata METHOD=REML;
 CLASS id;
 MODEL pps = tH|tH / S OUTPM=SMOOTH;
 RANDOM sb01-sb14 / TYPE=TOEP(1);
 RANDOM INTERCEPT / TYPE=UN SUBJECT= id;
 RUN;
```

HYST: smoothed model: pelvic problems resolved?

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



Part1: Example 1—Smoothing via mixed models

Effect of hysterectomy on perceived resolution of pelvic problems

Obviously, there was a large 'instantaneous' improvement at the time of the surgical intervention (hysterectomy)

Therefore, I fit a second smoothing model which

. included a binary indicator of hysterectomy status ('bump') (0=pre-hyst, 1=post-hyst),

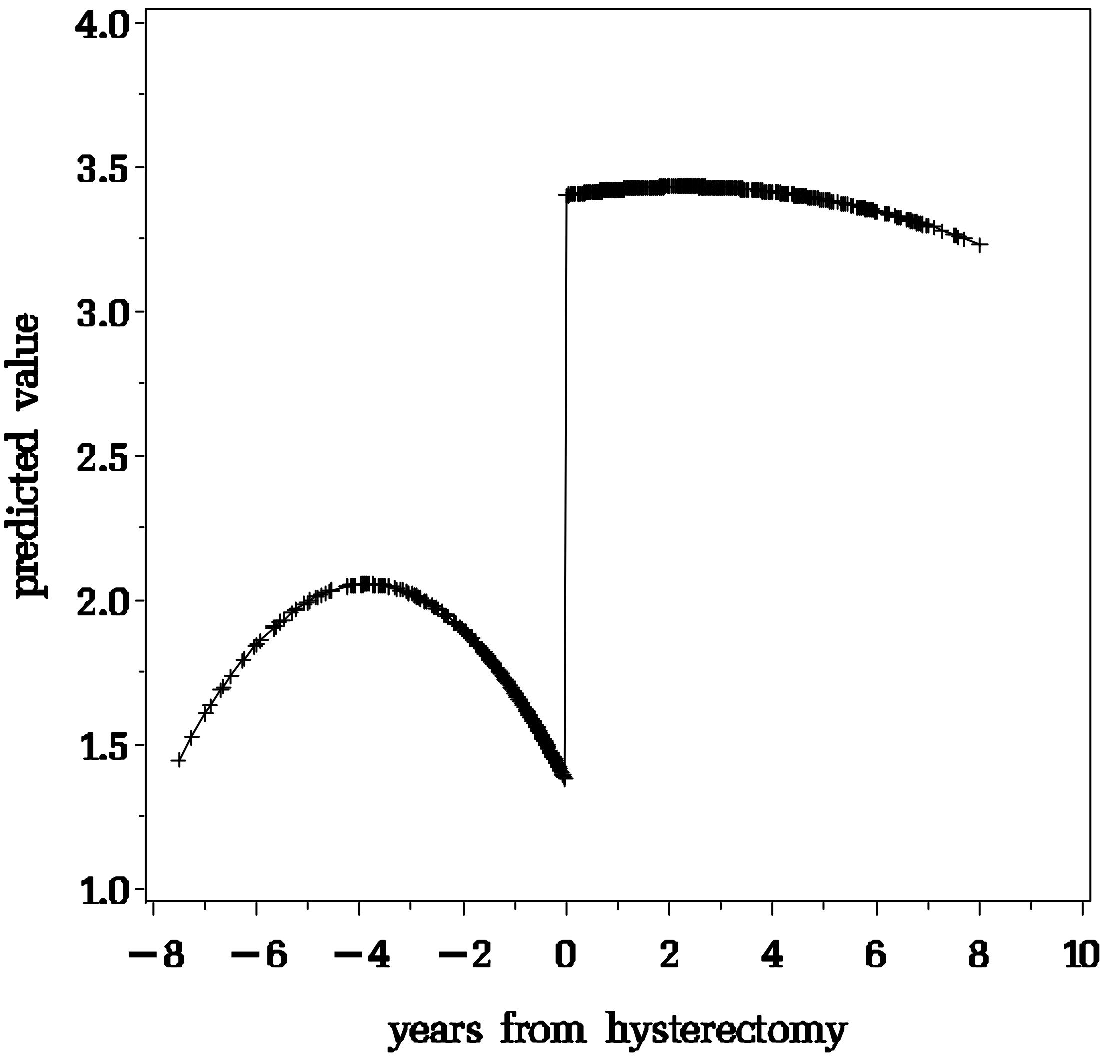
and

. smoothed the pre-hyst and post-hyst trajectories individually

```
PROC MIXED DATA=mydata METHOD=REML;
 CLASS id;
 MODEL pps = bump  tH(bump)  tH* tH(bump) / S OUTPM=SMOOTH;
 RANDOM sb01-sb14      / TYPE=TOEP(1);
 RANDOM INTERCEPT / TYPE=UN SUBJECT= id;
 RUN;
```

HYST: smoothed model: pelvic problems resolved?

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

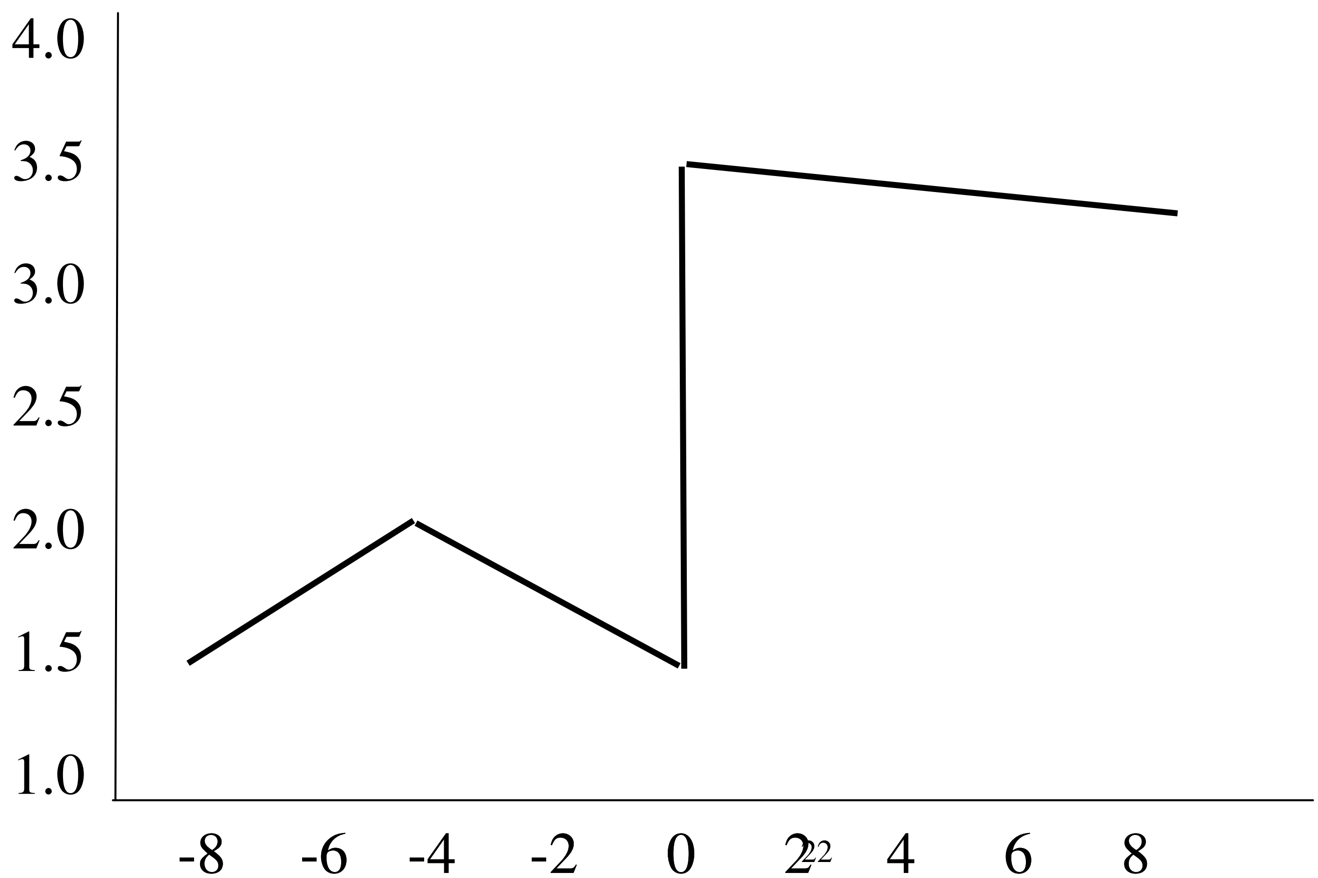


Part1: Example 1—

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

Effect of hysterectomy on perceived resolution of pelvic problems

The smoothed plot suggested that the overall trajectory might be approximated by a linear spline function with 2 knots (at times -4 and zero) and a 'bump' at the time of hysterectomy.



Part1: Example 1—

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

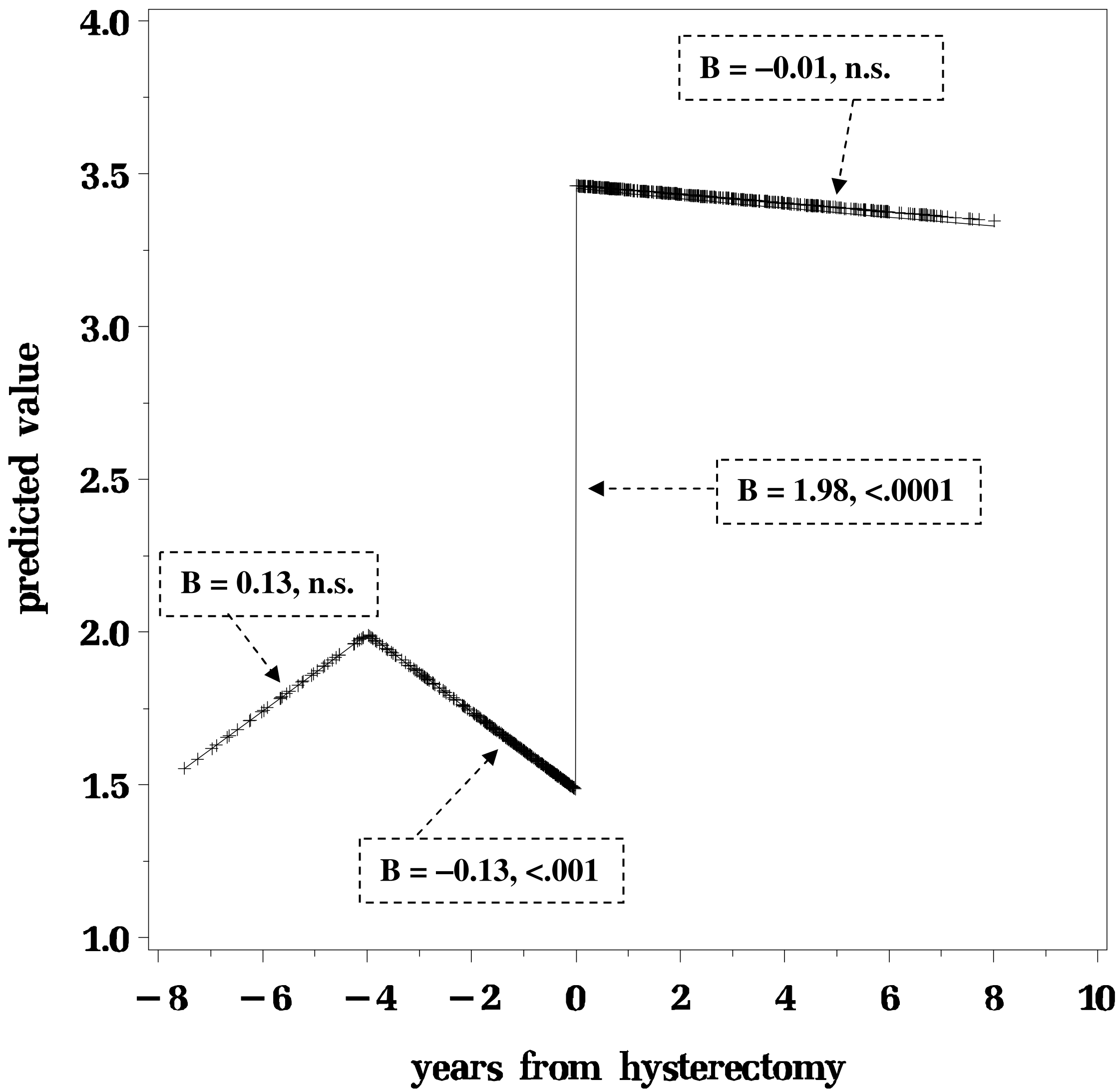
Effect of hysterectomy on perceived resolution of pelvic problems

```
PROC MIXED METHOD=REML DATA=mydata;
  CLASS id;
  MODEL pps =  bump  tH  sb01  sb02 /S  OUTPM=OUT;
   RANDOM INTERCEPT bump  / SUBJECT=id TYPE=UN;
RUN;
```

sb01 and sb02 are variables created from linear spline basis function, given knots at times -4 and 0 years

pelvic problems resolved?

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



Component	Est.
VAR(int)	0.25
VAR(bump)	0.59
COV(i,b)	-0.27
(r =	-0.69)
VAR(res)	0.3476

Part1: Example 2— Smoothing via mixed models
Effect of UPS on perceived resolution of pelvic problems

UPS is an acronym for uterine preserving surgery, e.g.,
myomectomy (to remove uterine fibroids)
uterine artery embolization (to shrink fibroids)
endometrial ablation (to control uterine bleeding)

Among those who did not have a hysterectomy, n=204 women had at least one UPS procedure during the study period

Occurrence and timing of UPS procedures was determined by self-report

The mean number of annual study observations was 6.10 (range 2-9) e.g., baseline plus 5.10 years of follow-up

I set each woman's final observed UPS date to year=0

The resulting 'time from UPS' variable, t_U, ranged from about −7.7 to +7.3 years.

Part1: Example 2— Smoothing via mixed models

Effect of UPS on perceived resolution of pelvic problems

14 knots along t_U chosen for spline model: years –7.0 through +6.0

Corresponding to each knot, I created a quadratic spline basis (sb01-sb14)

e.g., $sb14 = (t_U - 6)_+^2$

Outcome variable:

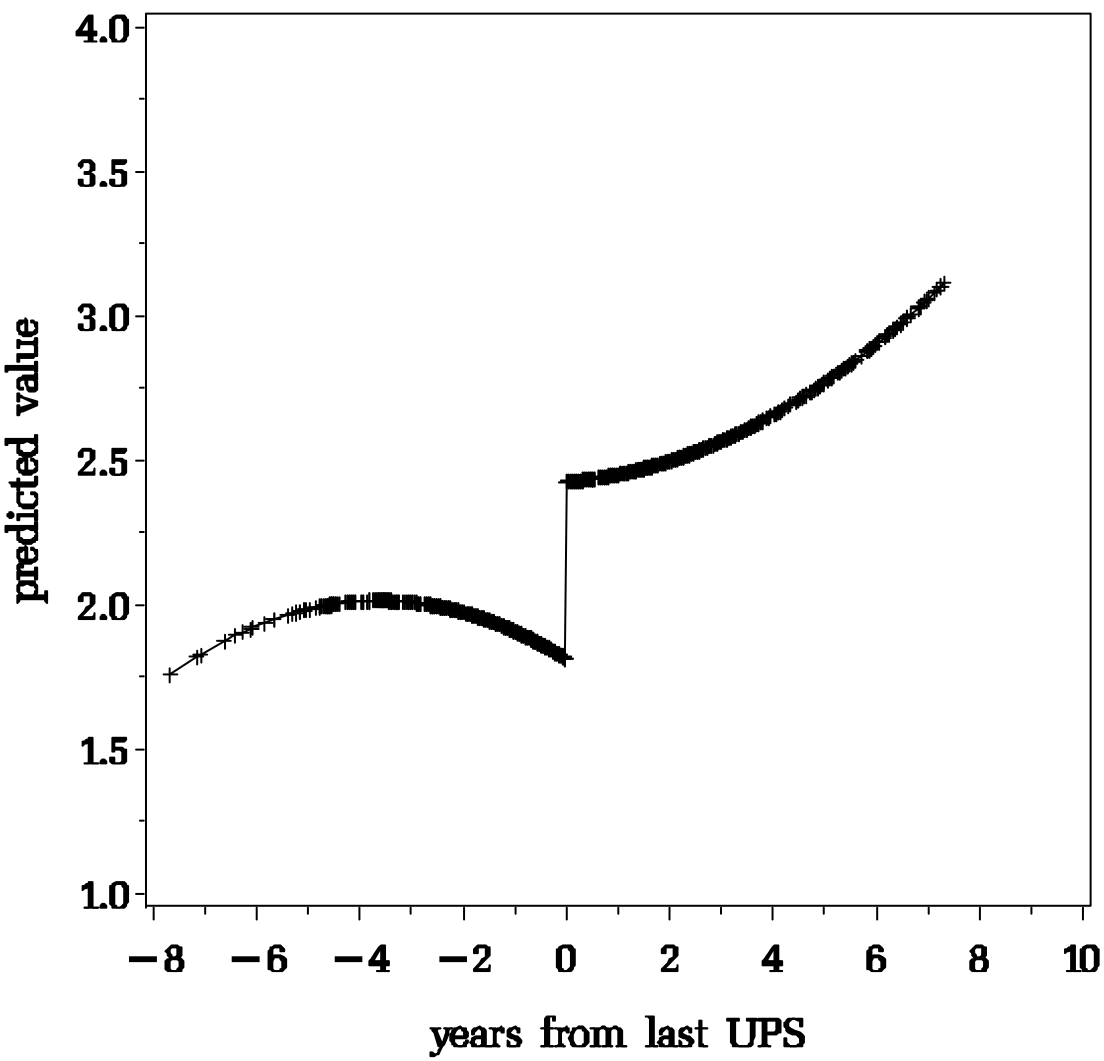
Perceived resolution of pelvic problems ('pps')

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

```
PROC MIXED DATA=mydata METHOD=REML;
 CLASS id;
 MODEL pps = tU|tU / S OUTPM=SMOOTH;
 RANDOM sb01-sb14      / TYPE=TOEP(1);
 RANDOM INTERCEPT / TYPE=UN SUBJECT= id;
 RUN;
```

UPS: smoothed model: pelvic problems resolved?

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

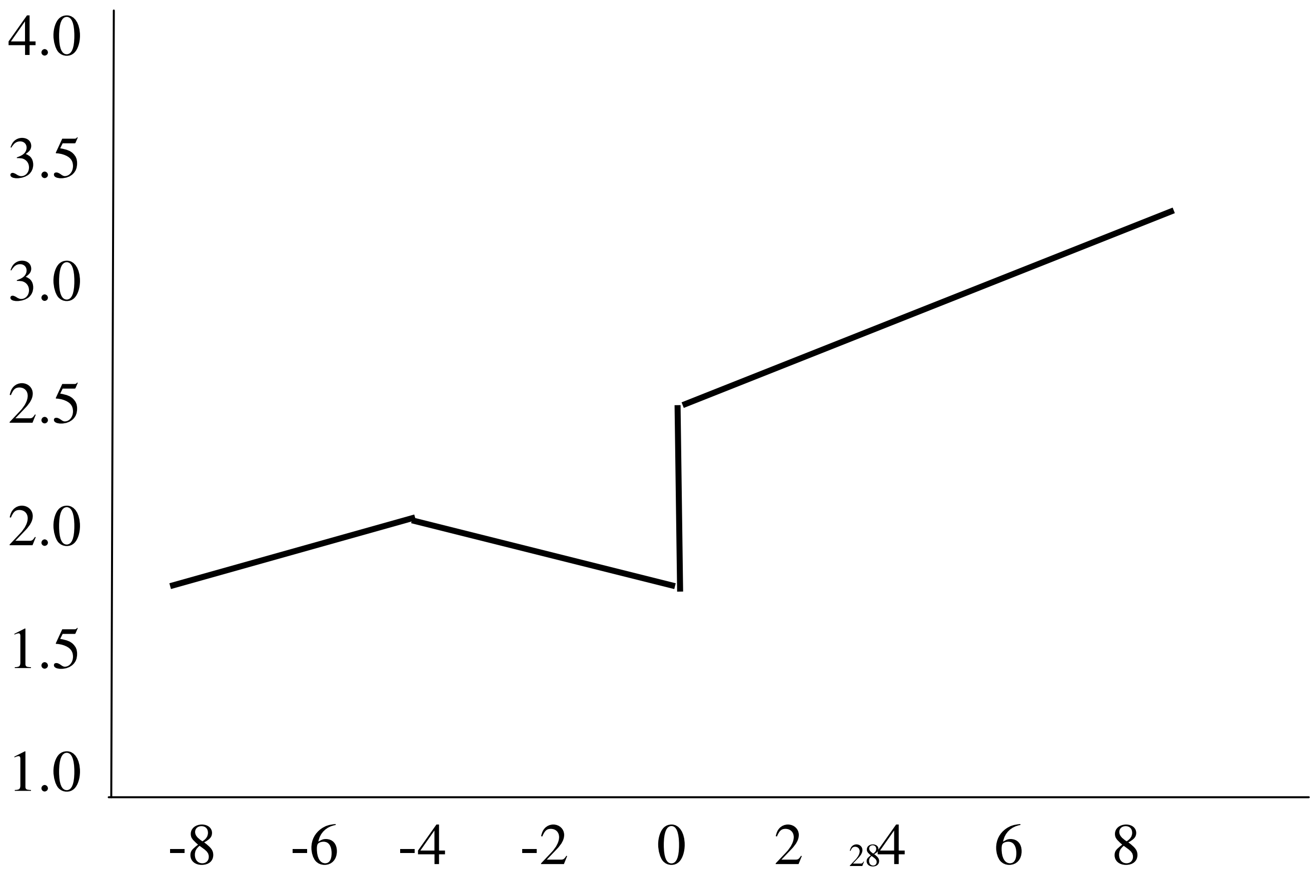


Part1: Example 2—

A two-knot linear spline with a 'bump' and random effects

Effect of UPS on perceived resolution of pelvic problems

The smooth plot suggested that the overall trajectory might be approximated by a linear spline function with 2 knots (at times -4 and zero) and a 'bump' at the time of hysterectomy



Part1: Example 2—

A two-knot linear spline with a 'bump' and random effects

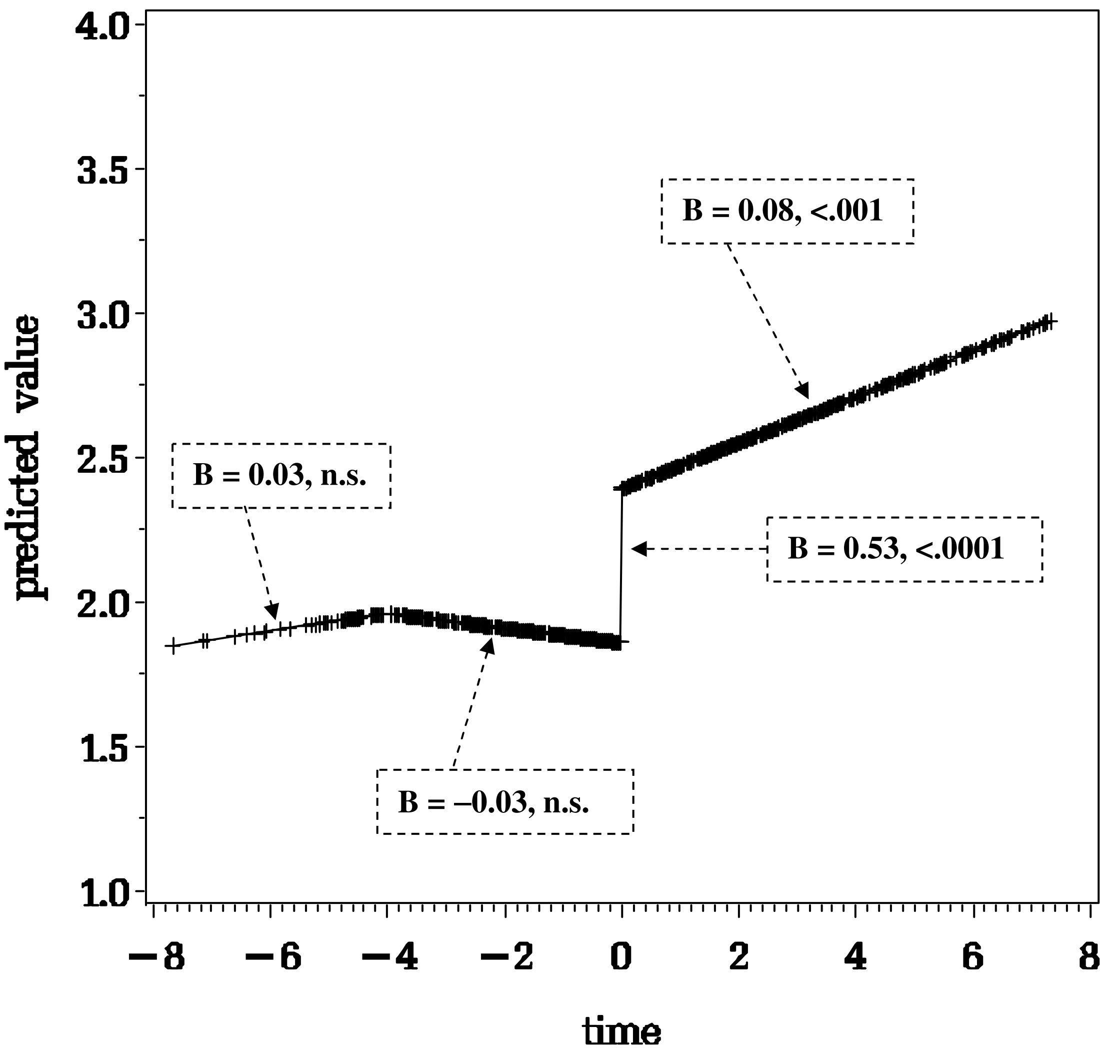
Effect of UPS on perceived resolution of pelvic problems

```
PROC MIXED METHOD=REML DATA=mydata;
  CLASS id;
  MODEL pps =  bump  tU  sb01  sb02 /S  OUTPM=OUT;
   RANDOM INTERCEPT bump  / SUBJECT=id TYPE=UN;
RUN;
```

sb01 and sb02 are variables created from the linear spline basis function, given knots at times -4 and 0 years

UPS: spline model: pelvic problems resolved?

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



Component	Est.
VAR(int)	0.51
VAR(bump)	0.60
VAR(seg2)	0.04
VAR(seg3)	0.06
COV(i,2)	-0.08
(r =	-0.51)
COV(2,3)	-0.04
(r =	-0.82)
VAR(res)	0.39

Part 1: Review

Spline functions w/in mixed models for smoothing 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 to model additional complexity in covariation among residuals can be used (e.g., PROC MIXED REPEATED statement).

In the two worked examples, variation in slopes (linear spline segments) was very low. Most between-patient variation was captured by the random intercept term.

Part 2: SEM approach to fitting growth curve models

Random intercepts and slopes are conceptualized as
inter-correlated latent variables

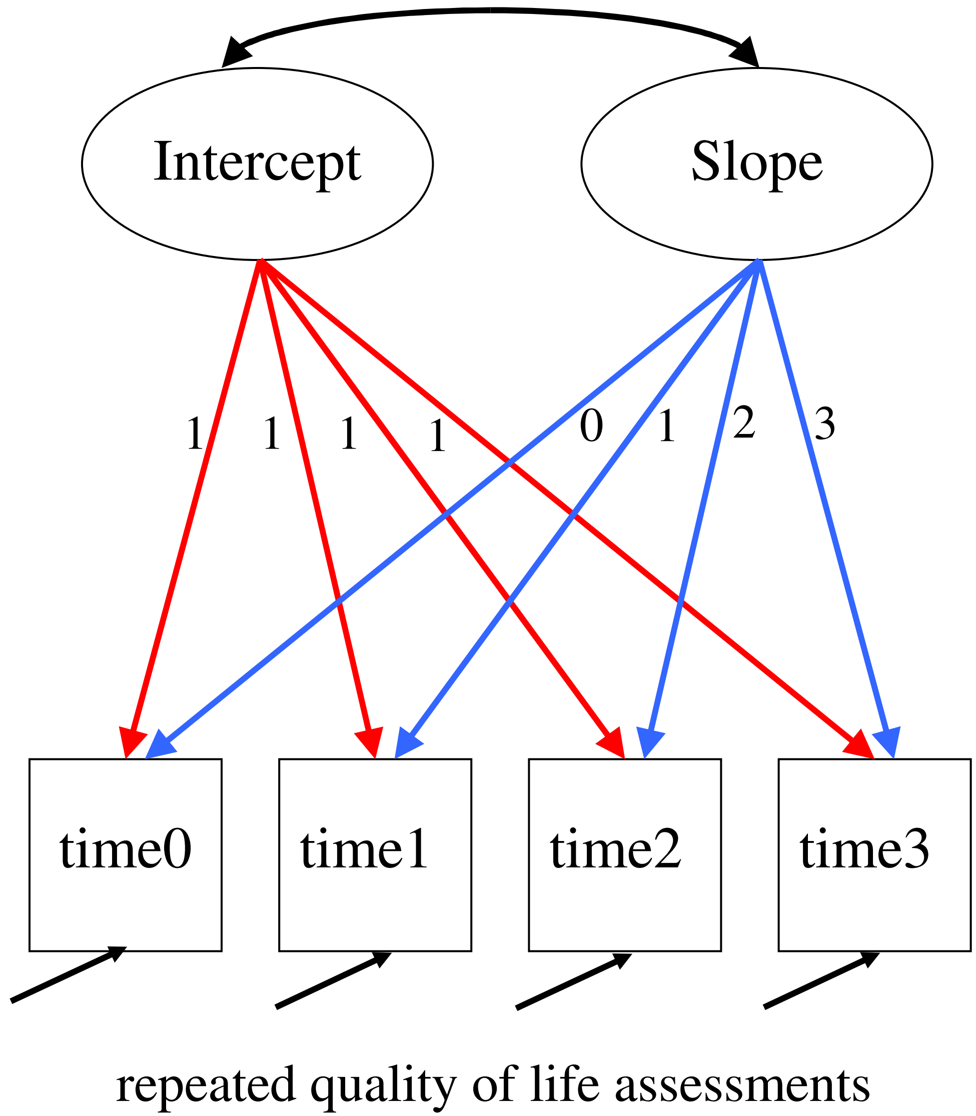
Most commonly,
. model fixed measurement occasions
. use observed covariance matrix and mean vector as input data

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

Part 2: SEM approach to fitting growth curve models

Representation of a linear growth curve model

a 'latent' growth curve (LGC) model



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 coefficient,

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

. $\mathbf{\Theta}$ holds residual variances of the xs

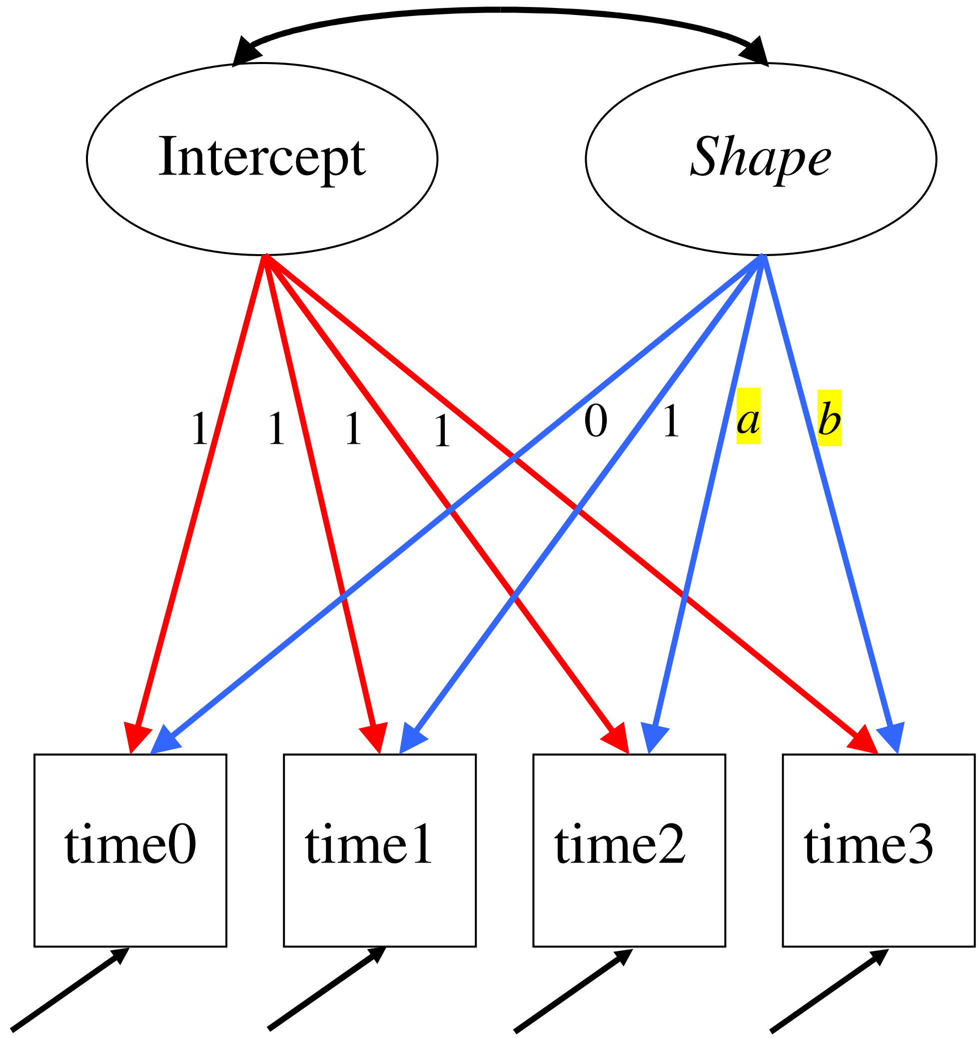
and

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

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

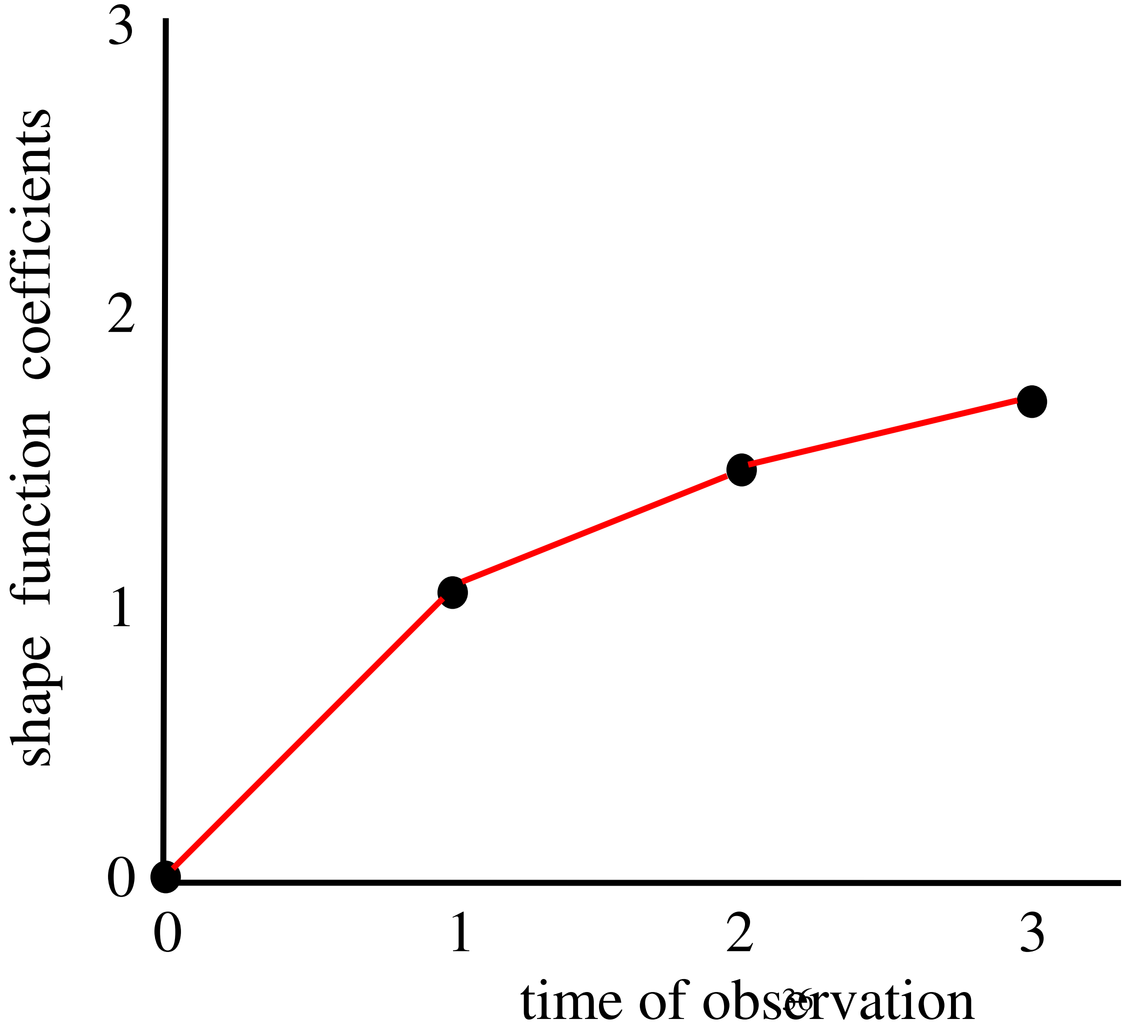
Allowing for non-linear trajectories



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

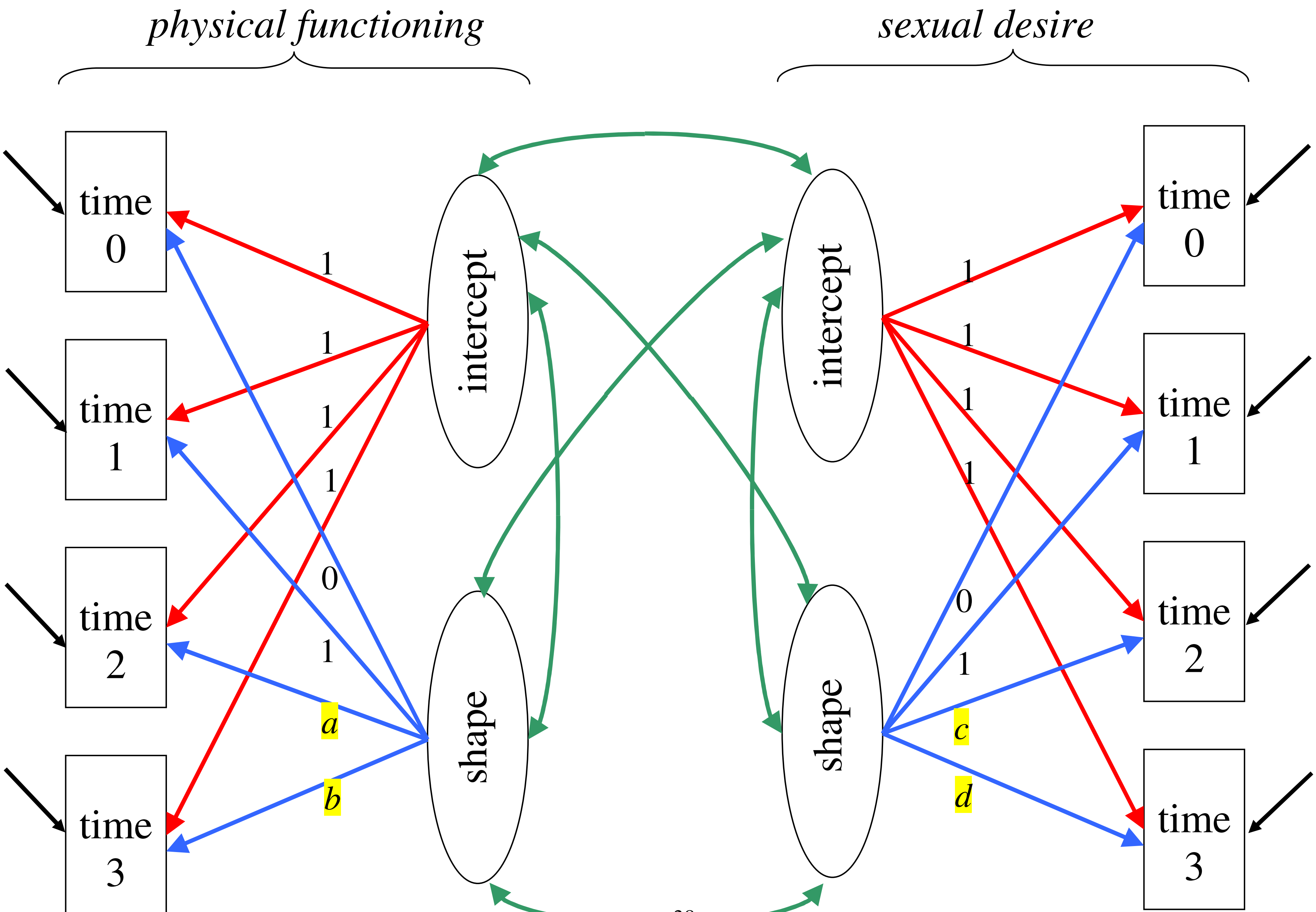


Part 2: SEM approach to fitting growth curve models
Associative latent growth curve models

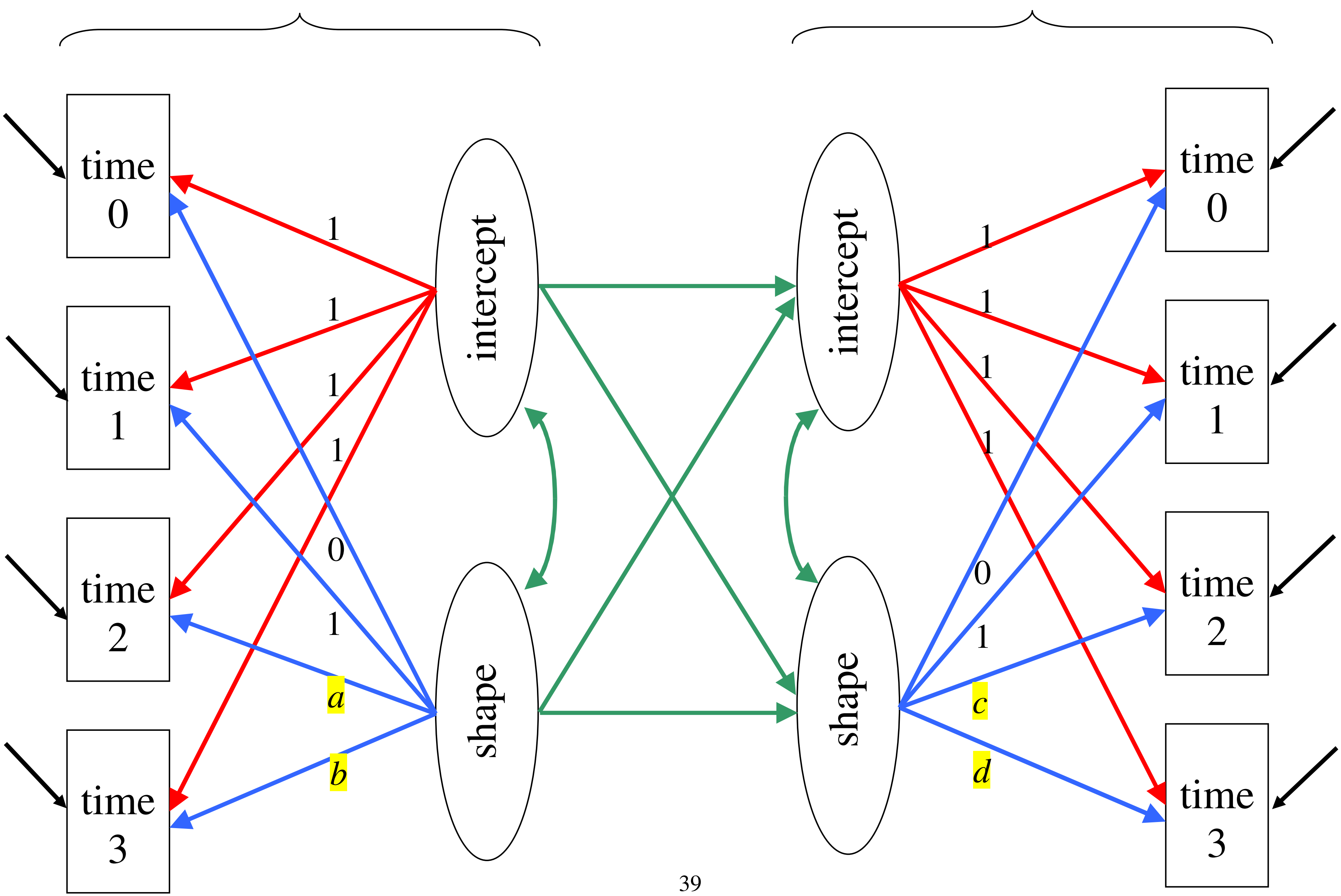
Simultaneously model growth curves for multiple outcomes

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.



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

<u>Related research questions</u>

. 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

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. Example application using the SOPHIA data

HRQOL (health-related quality of life) 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

PPS: perceived resolution of pelvic problems ('pelvic problems solved') (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
PPS	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