PSY792F SEM

Week 7 – Latent Growth Curve Modeling

Mark A. Prince, PhD, MS

When to use Latent Growth Curve Modeling (LGCM)

- When you want to model repeated measures of a variable
  - Typically, change over time
  - And you have a relatively small number of time points (e.g. $3 < t < 20$)
    - Not a strict rule, but less than 3 and you just have pre-post and only a linear relationship can fit, more than 20 and you should consider multilevel modeling
- When you want to see if initial position (i.e., intercept) and/or change (i.e., slope) in a construct predicts or is predicted by other variables
  - Intercepts and slopes can be IVs or DVs in your model
- When you have a theory about how a construct should change (e.g., increase linearly) or when you want to explore how a construct changes
  - LGCM is designed to test a priori assumptions
- Typically a series of comparison models are run to find the best fit
  - E.g., linear, quadratic, log, no growth

Why use LGCM?

- More flexible than repeated measures (M)ANOVA – by a lot!
  - Can look at many possible patterns of change
  - Can include covariates of intercept, slope, or both
  - Can use intercepts and slopes as mediators or moderators
  - Can model separate growth patterns among groups (i.e., multi-group analysis)
  - Can look at parallel growth among multiple constructs simultaneously
  - Can have time metrics that are uneven
  - Can put the intercept anywhere
  - Can model discontinuous change
  - Can handle missing data

Theoretical Overview

- LGCM has the same assumptions as regression and SEM
- LGCM uses some mathematical ‘tricks’ to give special meaning to latent variables
  - Assigning all 1’s to the factor loadings makes a latent variable an intercept
  - Assigning a series of numbers (e.g., 0, 1, 2, 3) creates a slope
    - Slopes can take any shape
      - Examples include
        - 0, 1, 2, 3 – linear
        - 0, 1, 4, 9 – quadratic
        - 0, .69, 1.10, 1.39 – log
        - 1, 1, 1, 1 – no growth
        - -3, -2, -1, 0 – linear with intercept at the end

Some pictures of linear models
from Preacher, 2010

$$\Lambda_A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \quad \Lambda_B = \begin{bmatrix} 1 & -2 \\ 1 & 0 \\ 1 & 2 \\ 1 & 4 \end{bmatrix} \quad \Lambda_C = \begin{bmatrix} 1 & -3 \\ 1 & -2 \\ 1 & -1 \\ 1 & 0 \end{bmatrix} \quad \Lambda_D = \begin{bmatrix} 1 & 10 \\ 1 & 20 \\ 1 & 30 \\ 1 & 40 \end{bmatrix},$$

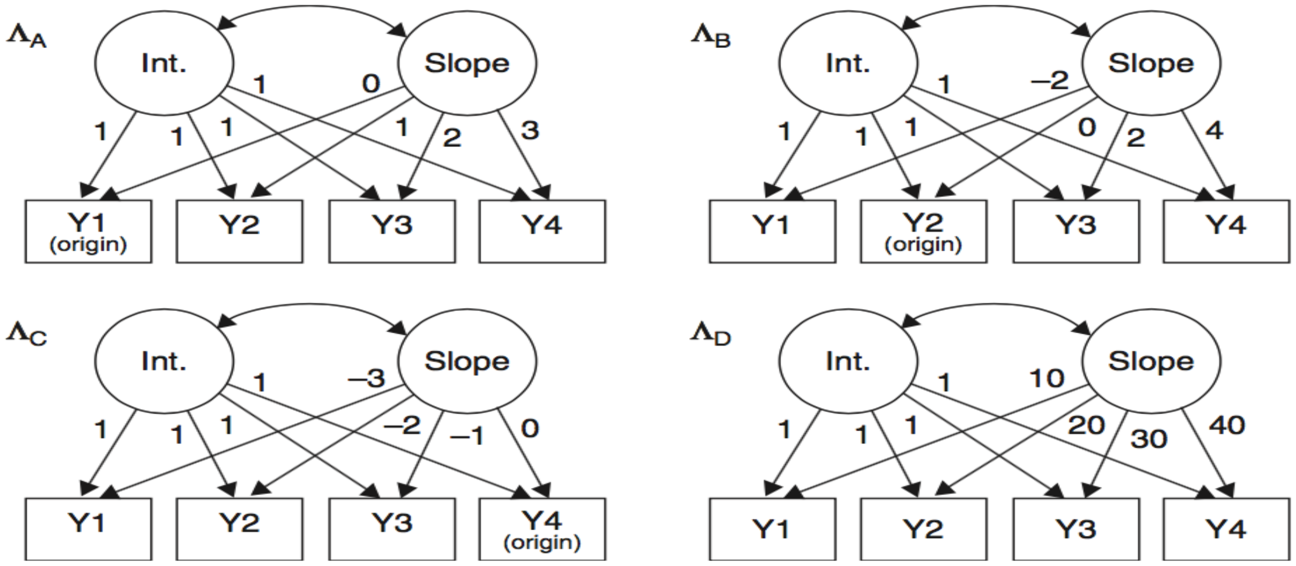


Figure 14.1 How the Loadings in Λ_A, Λ_B, Λ_C, and Λ_D Might Be Represented in Path Diagrams.

Full path model with all paths and symbols

The triangle represents a constant, which makes paths alpha 1 and 2 carry the value of the means of the intercept and slope

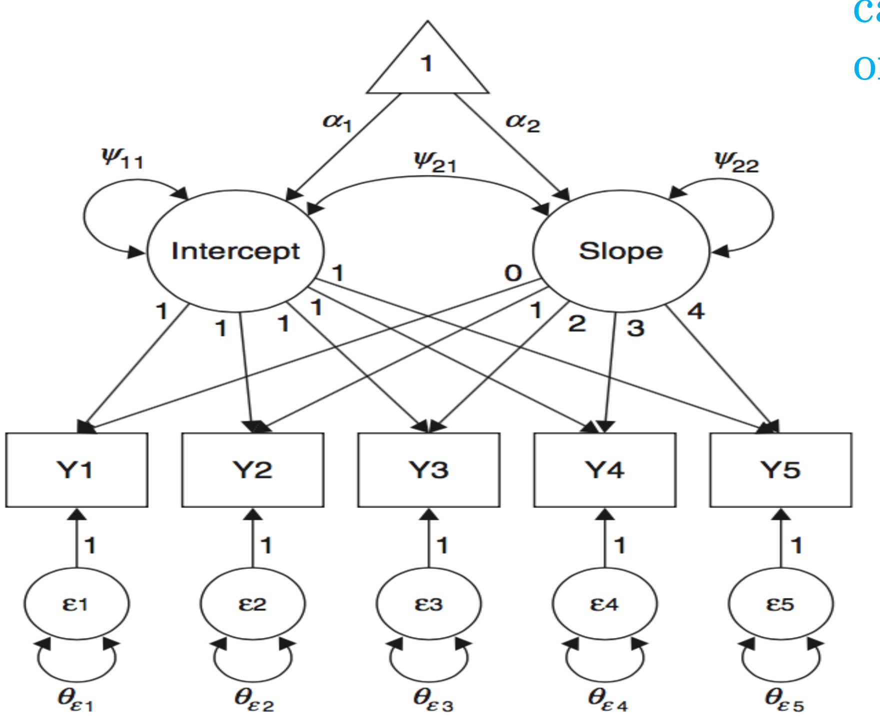
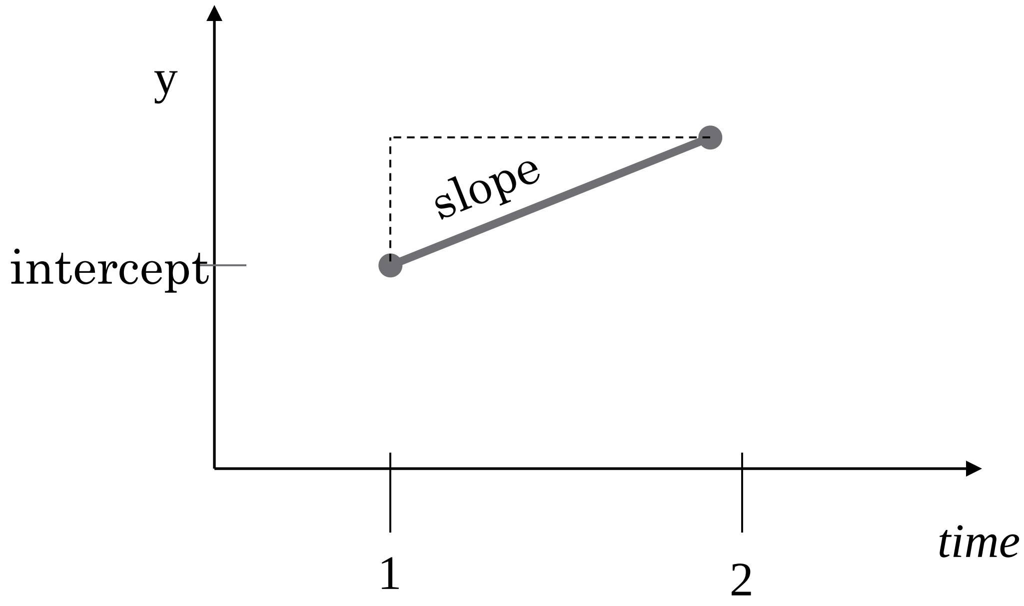


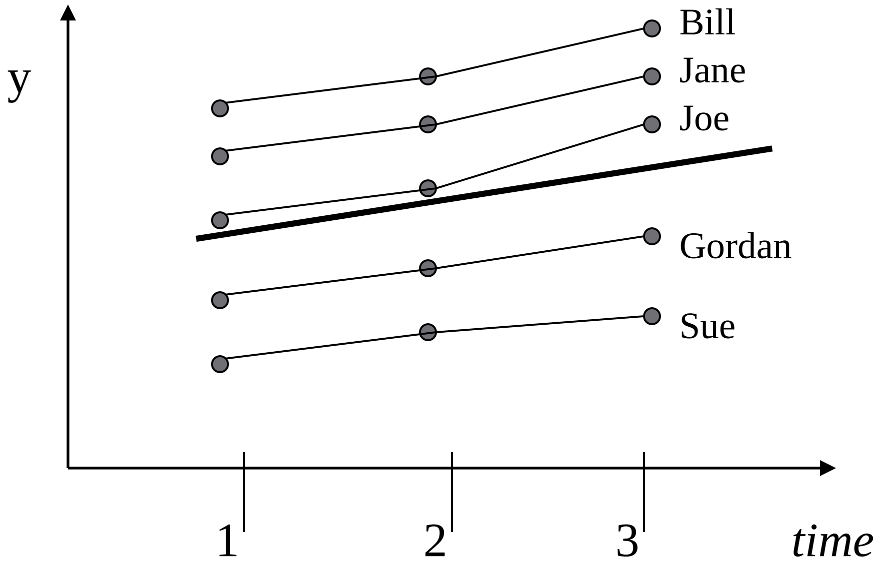
Figure 14.2 Path Diagram of a Linear Latent Growth Curve Model with Random Intercepts, Random Slopes, and Unconstrained Residual Variances.

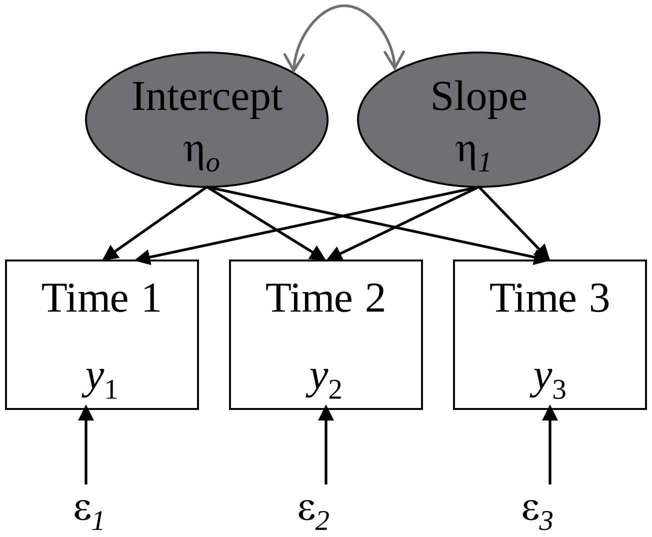
Growth with Two Measurement Occasions

- Straight line fits perfect
- 0 degrees of freedom



Latent Growth Curve = longitudinal continuous latent variable





ex. Depression scores measured over time can be summarized by taking the average initial level and slope for all individual trajectories

Parameters of Growth Model

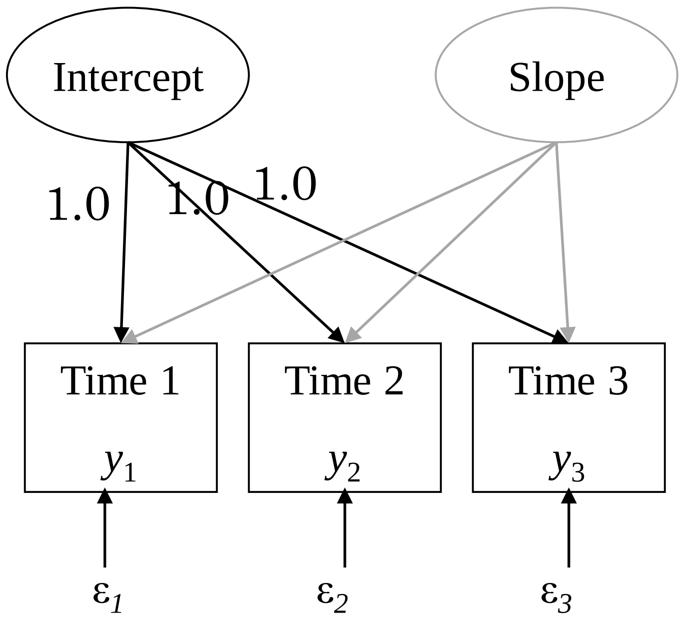
η0i (initial status, level, baseline, intercept): Systematic variation in the outcome variable at the time point where the time score is zero

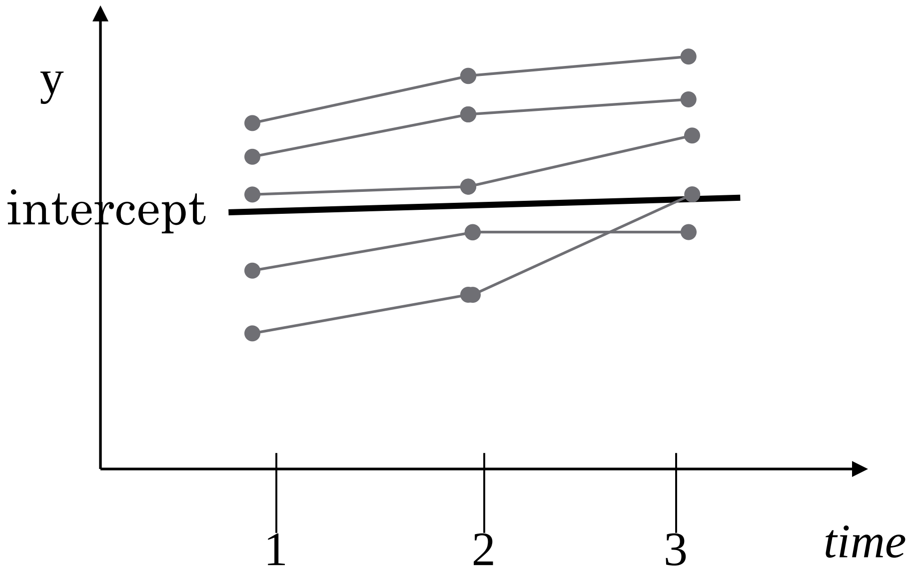
Mean = average of the outcome over individuals at the time point where the time score is zero; when the first time score is zero, it is the intercept of the average growth curve, also called initial status

Variance = variance of the outcome over individuals at the time point where the time score is zero, excluding the residual variance

Intercept (i.e., initial status)

- Unstandardized regression coefficients for intercept predicting all three time scores is set at a constant value
  - 1 is convention, but other values would produce the same result.
- All values are fixed at the same value, thus the intercept represents the constant level of “Y”, <u>if there were no growth</u>





Parameters of Growth Model

η_{1i} (growth rate, trend, slope):

Systematic rate of change (increase or decrease) in the outcome for a time score increase of one unit.

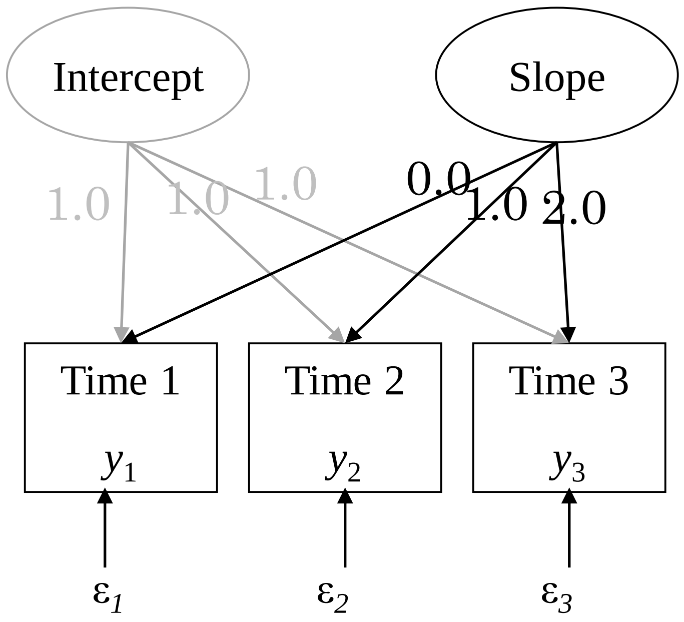
Mean = average rate of change over individuals

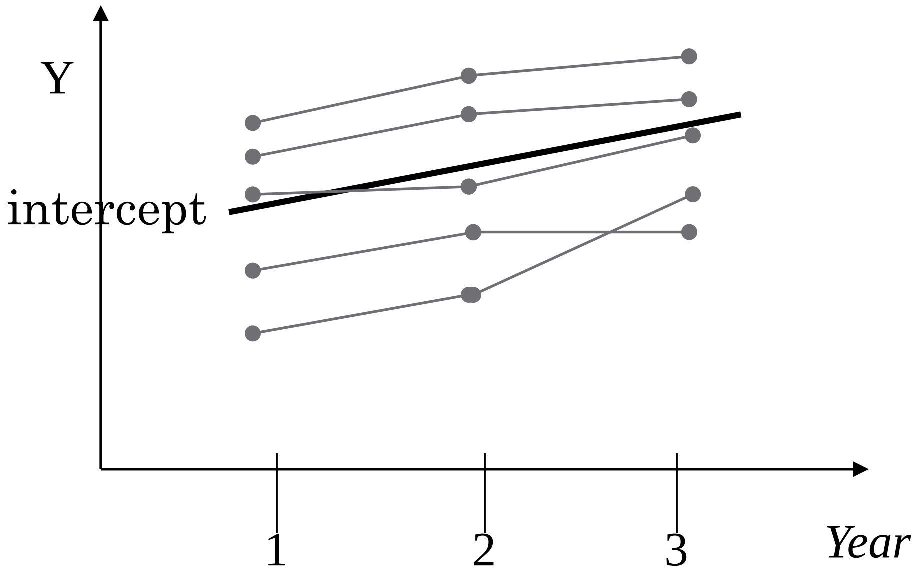
Variance = variance in rate of change over individuals

Covariance with intercept = relationship between individual intercept and slope values

Slope (i.e., rate of change)

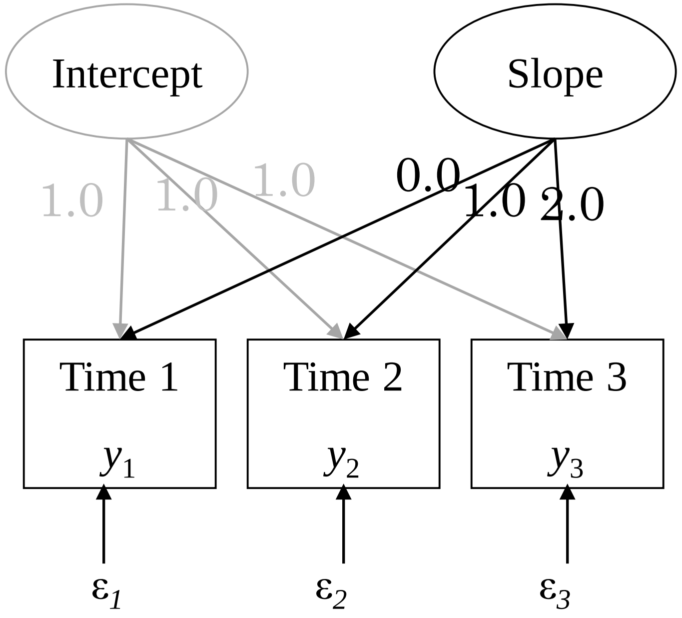
- Unstandardized regression coefficients for slope predicting all three time scores are set to reflect ordering of time
  - Time point 1 represents 0 growth
  - Time point 2 represents 1 year later
  - Time point 3 represents 2 years later

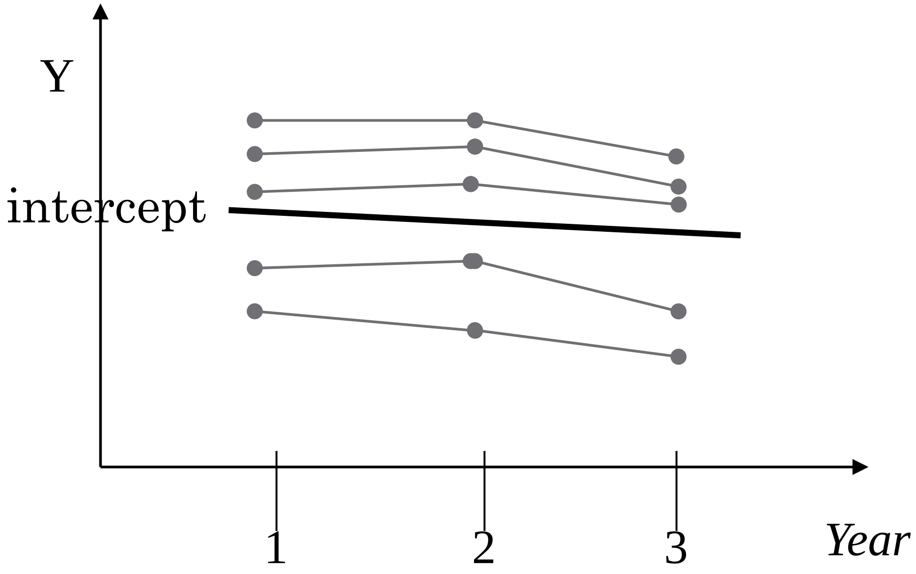




Slope (i.e., rate of change)

- Unstandardized regression coefficients for slope predicting all three time scores are set to reflect ordering of time
  - Time point 1 represents 0 growth
  - Time point 2 represents 1 year later
  - Time point 3 represents 2 years later





Time Scale

- Choice of how to code and scale time is important!
- The intercept η_{0i} is the expected value of y at the origin of time (i.e., when t equals 0).
- The “origin of time” (where slope coefficient = 0) has a direct and predictable effect on the:
  - Mean of the intercept
  - Variance of the intercept
  - Covariance between the intercept and slope
- The origin of time does not impact the estimate of the slope, because the slope is the expected change in y for a 1-unit change in time

Common Methods for Coding Time

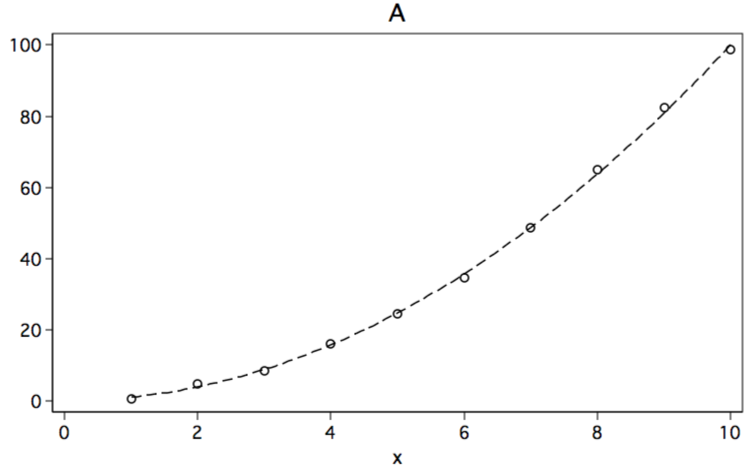
- Code time to produce parameter estimates that are more easily and readily interpretable
- Code time to address a substantive question.
  - Interested in individual differences at the beginning of the assessed growth process, then make origin of time at the initial assessment (e.g., 0, 1, 2, 3)
  - Interested in the end of the assessed growth process, then make origin of time at the last assessment (e.g., 3, 2, 1, 0)
  - Interested in alternative functional forms (e.g., power function), then code the first two coefficients as 0, 1, and allow the rest to be estimated and the mean slope represents a curve which can be essentially any function

Nonlinear Growth

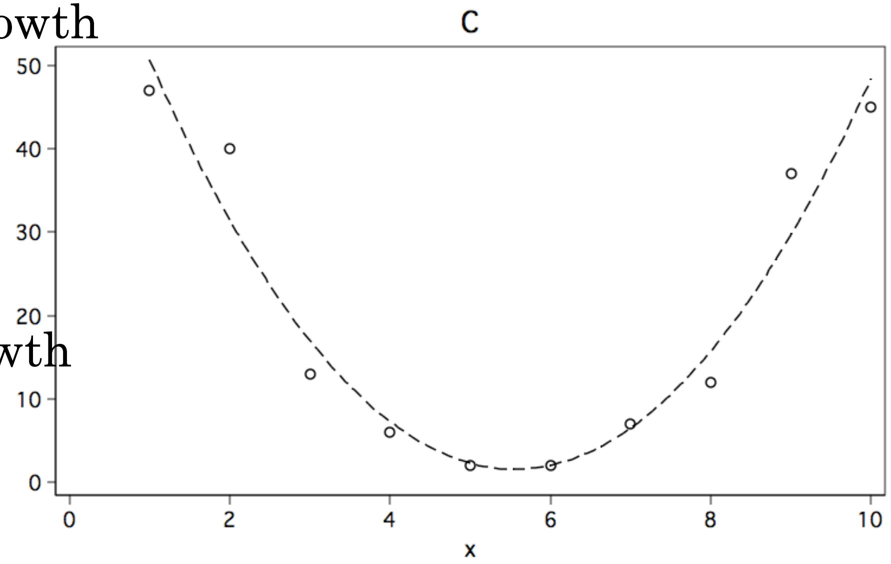
- If estimating linear and quadratic (or higher order polynomials), provide specification by squaring/cubing/etc time scores
  - Linear = 0, 1, 2, 3; quadratic = 0, 1, 4, 9; cubic = 0, 1, 8, 27
  - Linear = 0, 2, 4; quadratic = 0, 4, 16; cubic = 0, 8, 64
- Recommend > 3 timepoints for quadratic+ model
- Can also specify the first two time scores and allow the remainder to be estimated by the model

Quadratic Growth Curves – Positive (smile)

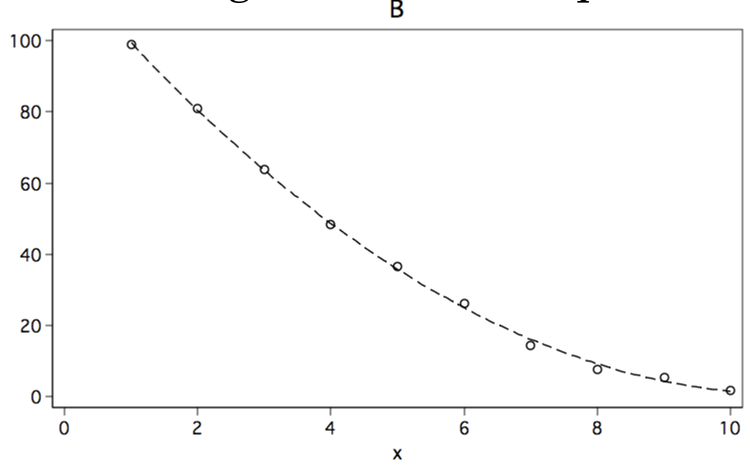
Significant positive linear + quadratic growth



Non-significant negative linear + quadratic growth

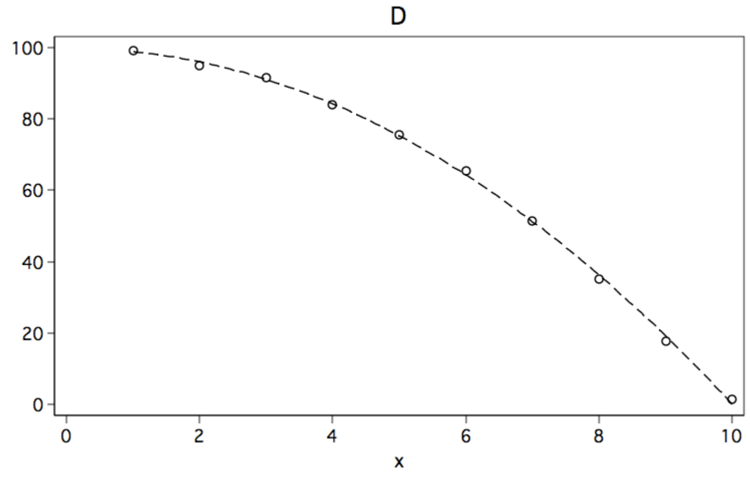


Significant negative linear + quadratic growth

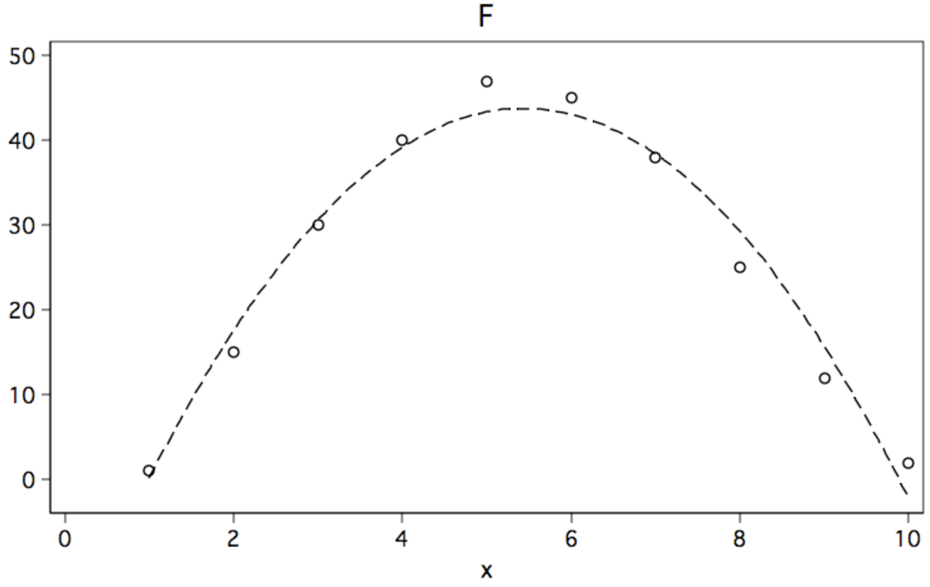


Quadratic Growth Curves – Negative (frown)

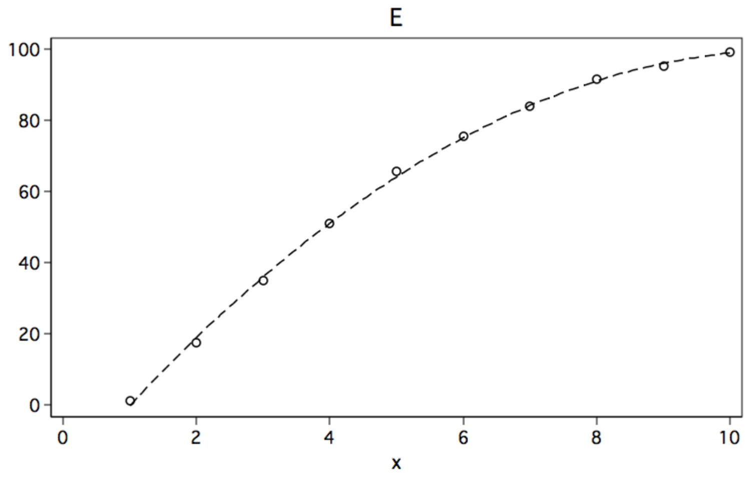
Significant negative linear negative quadratic growth



Non-significant negative linear negative quadratic



Significant positive linear negative quadratic



Nonlinear Growth

- Linear and quadratic or polynomial effects are very highly correlated if time is coded as starting at 0, consider “centering time” (subtract mean of time)
  - e.g., linear coding of 0 1 2 3 4 becomes -2 -1 0 1 2
  - Quadratic coding of 0 1 4 9 16 becomes -4 -1 0 1 4

Effect Size in LGCM

Feingold (2009)

- Mean values of intercepts and slopes are good effects themselves
- Compare slopes between groups (e.g., treatment vs. control)
- $d_{\text{GMA-CHANGE}} = \boldsymbol{\beta}_{11}/\text{sqrt}(\boldsymbol{\tau})$
  - Used to calculate power
- $d_{\text{GMA-RAW}} = \boldsymbol{\beta}_{11}(\text{time})/\text{SD}_{\text{RAW}}$
  - Used to calculate magnitude of the effect
- $\boldsymbol{\beta}_{11}$ - the difference between the groups in mean growth rates (mean change for group 1 – mean change for group 2)
- $\boldsymbol{\tau}$ - estimate of within group variability of the 'true' score of the slopes
  - $\boldsymbol{\tau}$ is calculated as $\text{SD}_{\text{CHANGE}}$ = standard deviation of the change scores
- Time – use the end time point to obtain the difference between model-estimated means of the two groups at the end of the study (adjusted for baseline differences)
- SD_{RAW} – pretest or baseline SD

Notes on interpretation

- Means
  - i – the average value for the intercept
  - s – the average change over time
  - i with s – the correlation between the intercept and slope
    - + higher starting point associated with steeper slope
    - - higher starting point associated with flatter slope
  - Predictors of i tell us if a variable is associated with where someone starts in the growth process
  - Predictors of s tell us if a variable is associated with someone's rate of change
  - Freely estimated time scores
    - Compare to linear, quadratic, etc. expectations
    - If you get, 0, 1, 2.5, 3.7 – that means the growth is happening faster than linear
    - If you get, 0, 1, 1.2, 1.4 – that means that the growth is happening slower than linear

Mplus code

- New symbol "|" used to indicate intercept and slope specifications
  - The | symbol, located above the enter key is referred to as 'bar' or 'pipe'
  - It is used in conjunction with i, s, and q commands
  - Also LGCM uses @ to indicate fixed time metrics or * for a starting point for time metrics

- Examples

- i s | x1@0 x2@1 x3@2 x4@3 x5@4;
  - This will give you an intercept and a linear slope

- i s q | x1@0 x2@1 x3@2 x4@3 x5@4;
  - This will give you an intercept, a linear slope, and a quadratic slope

- i s | x1@0 x2@1 x3 x4 x5;
  - Will estimate the time metric for times 3 through 5 – this will tell you how close to linear the trend is

- i s | x1@0 x2@1 x3*2 x4*3 x5*4;
  - This will estimate the time metric using the indicated values as a starting point

Simulated data code example

Mplus users guide example
TITLE: this is an example of a linear growth model for a continuous outcome
DATA:
FILE IS ex6.2.dat;
VARIABLE: NAMES ARE y11-y14;
MODEL: i s | y11@0 y12@1 y13@2 y14@3;

To free the growth trajectory use:
MODEL: i s | y11@0 y12@1 y13 y14;

To provide starting values use:
MODEL: i s | y11@0 y12@1 y13*2 y14*3;

- let's compare outputs

Real data code examples

- Two studies
  - Rural – examined three alcohol interventions for people in rural communities
    - Needed to use dummy coded variables
    - Needed to use censored outcomes
    - **model write up is from this paper, which is currently under review**

Extensions

- Covariates
- Multiple groups
- Second order latent growth
- Parallel process

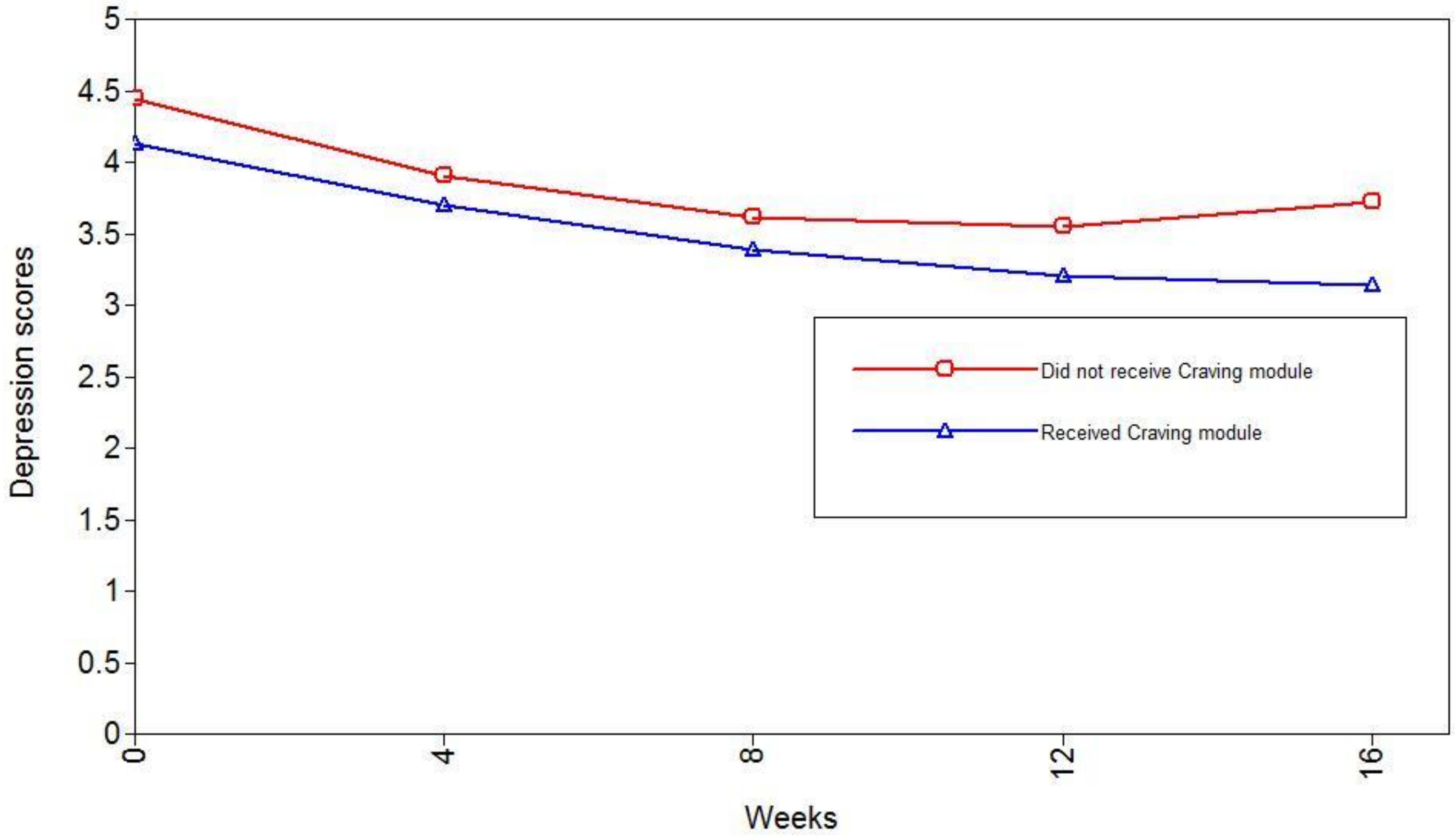
Adding Covariates

```
MODEL:
i s q | depres0@0 depres4@1 depres8@2 depres12@3 depres16@4;


i s q ON gender age adstot
```

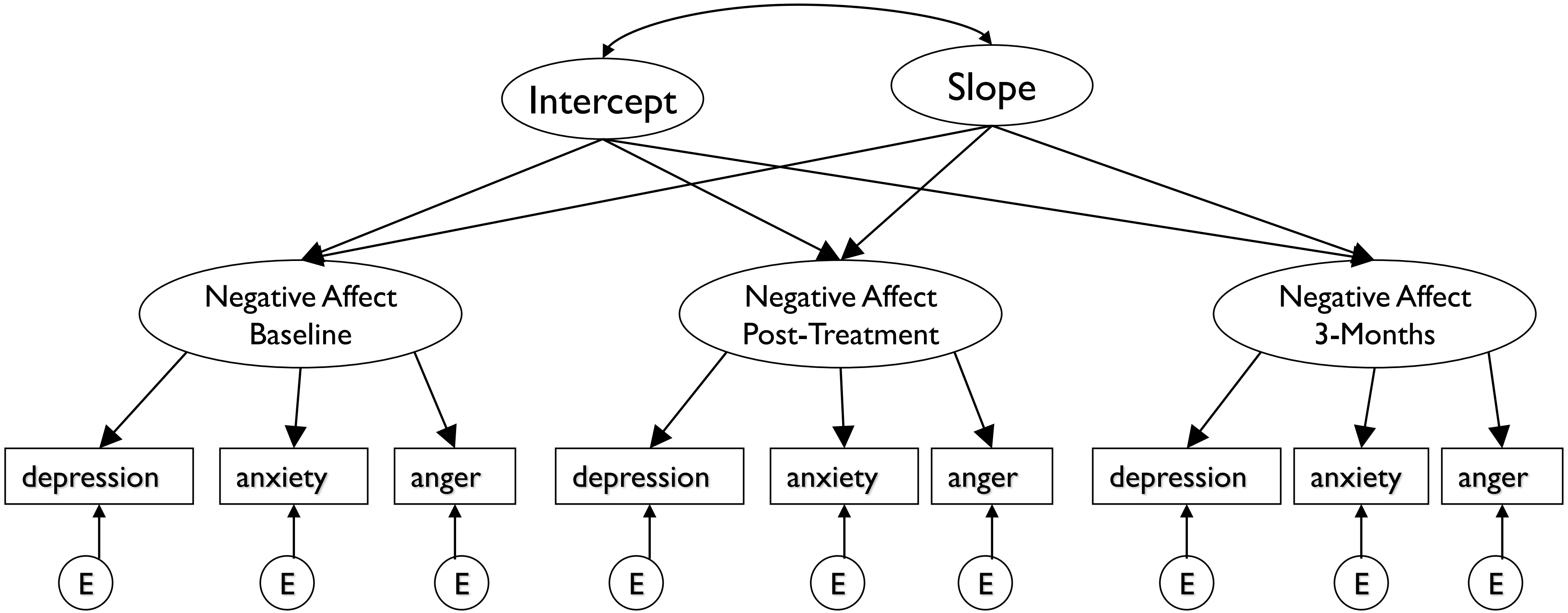
Multiple Groups Growth Modeling

```
GROUPING IS cravmod (0=No 1=Yes);



MODEL:

i s q | depres0@0 depres4@1 depres8@2 depres12@3 depres16@4;
```

Multiple-Groups Growth Modeling



Second Order Latent Growth Model



Parallel Process Latent Growth Model