HLM versus SEM Perspectives on Growth Curve Modeling

Hsueh-Sheng Wu

CFDR Workshop Series

August 3, 2015

BGSU
Center for Family and Demographic Research

Outline

- What is Growth Curve Modeling (GCM)
- Advantages of GCM
- Disadvantages of GCM
- Graphs of trajectories
- Inter-person difference in trajectories
- Key concepts of GCM
- SEM and HLM approaches to GCM
- Sample SEM and HLM Syntax in Stata
- Respective strengths of SEM and HLM approaches
- Conclusions

What is Growth Curve Modeling (GCM)

- Growth curve modeling is a technique to describe and explain an individual's change over time.

- Main Research Questions:
  - What are the patterns of change for individuals over time
  - What factors are related with the patterns of change over time

- Data requirement:
  - Panel data
  - The more waves of data you have, the more complex models you can estimate. For example, with three-wave panel data, you can test a linear growth curve model only, but with four-wave panel data, you can test both linear and curvilinear growth curve models.

Advantages of GCM

- Examine constructs measured at several time points simultaneously, not just at one point in time.
- Make clear distinction between two concepts: (1) intra-individual change and (2) inter-person differences. The first concept refers to how the outcome variable varies for the same individuals over time. The second concept hints on why people differ in their changes over time. This distinction is often blurred when cross-sectional data or traditional statistical methods are used.
- Provide a more comprehensive framework to understand change over time: (1) whether change occurs over time, (2) what is the functional form between change and time, (3) what parameters are used to describe the functional form, and (4) what factors are associated with these parameters.
- Incorporate time-invariant and time-varying variables in the analyses
- Include respondents even when respondents had missing data on some of the time points

Disadvantage of GCM

- GCM can only be used if the data meet the following criteria:
  - at least 3 waves of panel data
  - Outcome variables should be measured the same way across waves
  - Data set need to have a time variable
- No theories dictate the functional form between outcome and time. Thus, researchers often need to explore and decide the best empirical functional form for the outcome variable first.
- GCM can be computationally intensive for complex models.

Graphs of Trajectories

- Researchers usually use the term, trajectory, to describe the patterns of changer for individuals over time.
- Different trajectories describe different functional forms between time and the construct of interest.
- The simplest trajectory is a linear trajectory, defined by two parameters, intercept and slope. More complex trajectories are defined by intercept, slope, and additional parameters.

Four Sample Trajectories and Their Parameters

- A linear trajectory

  $Y = 4 + 0.05 * Age + \sigma$

- A piecewise linear trajectory

  $Y = 4 + 0.05 * Age + \sigma$ if X <50

  $Y = 4 + 0.05 * Age + 0.05 * Age + \sigma$ if X >=50

- A curvilinear trajectory

  $Y = 4 + 0.05 * Age + 0.03 * \mathrm{Age}^2 + \sigma$

- A cubic trajectory

  $Y = 4 + 0.05 * Age + 0.03 * \mathrm{Age}^2 + (-0.0001) * \mathrm{Age}^3 + \sigma$

Quadratic

Intercept

Slope

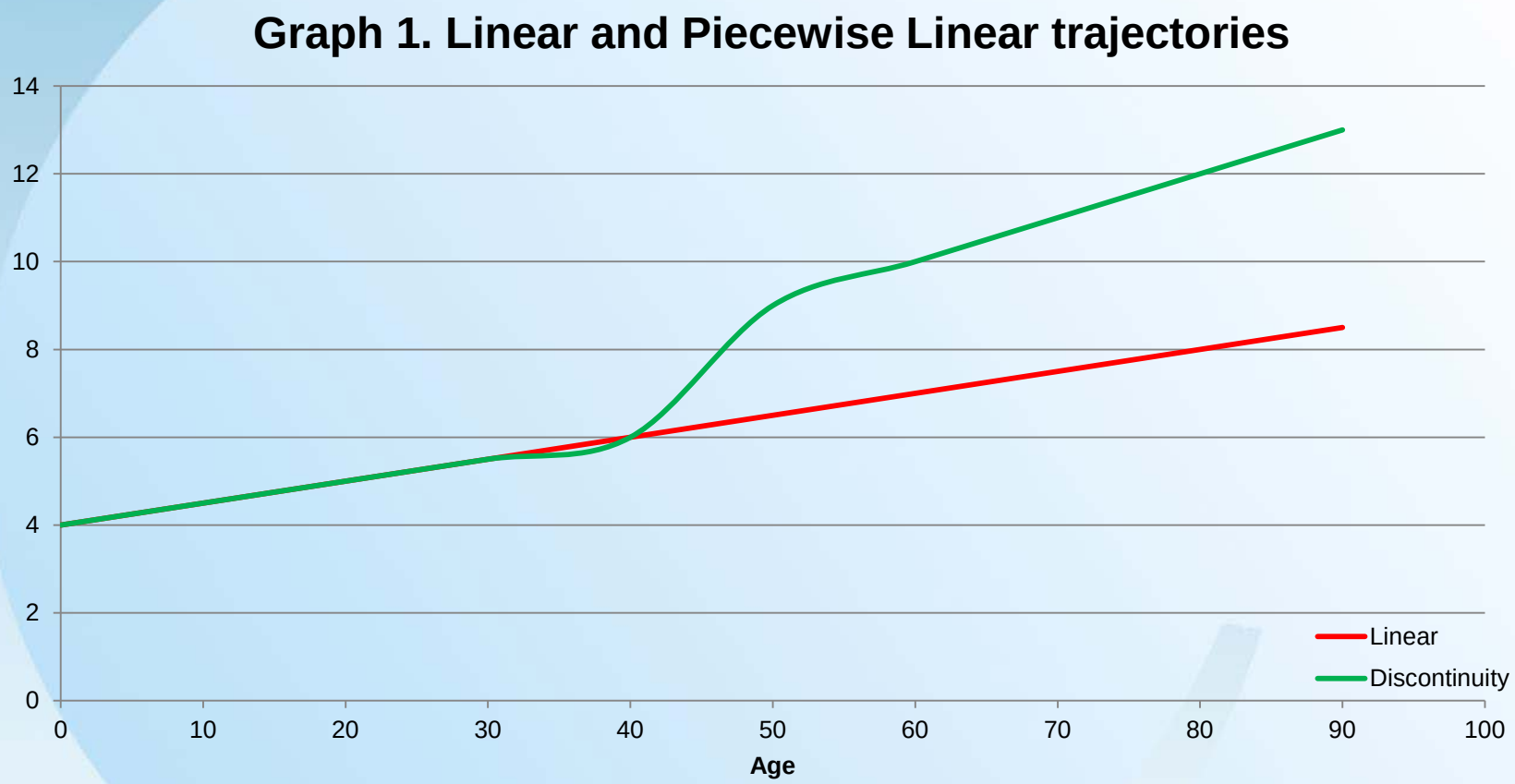
Cubic

Center for Family and Demographic Research

Table 1. Predicted Values for Sample Trajectories

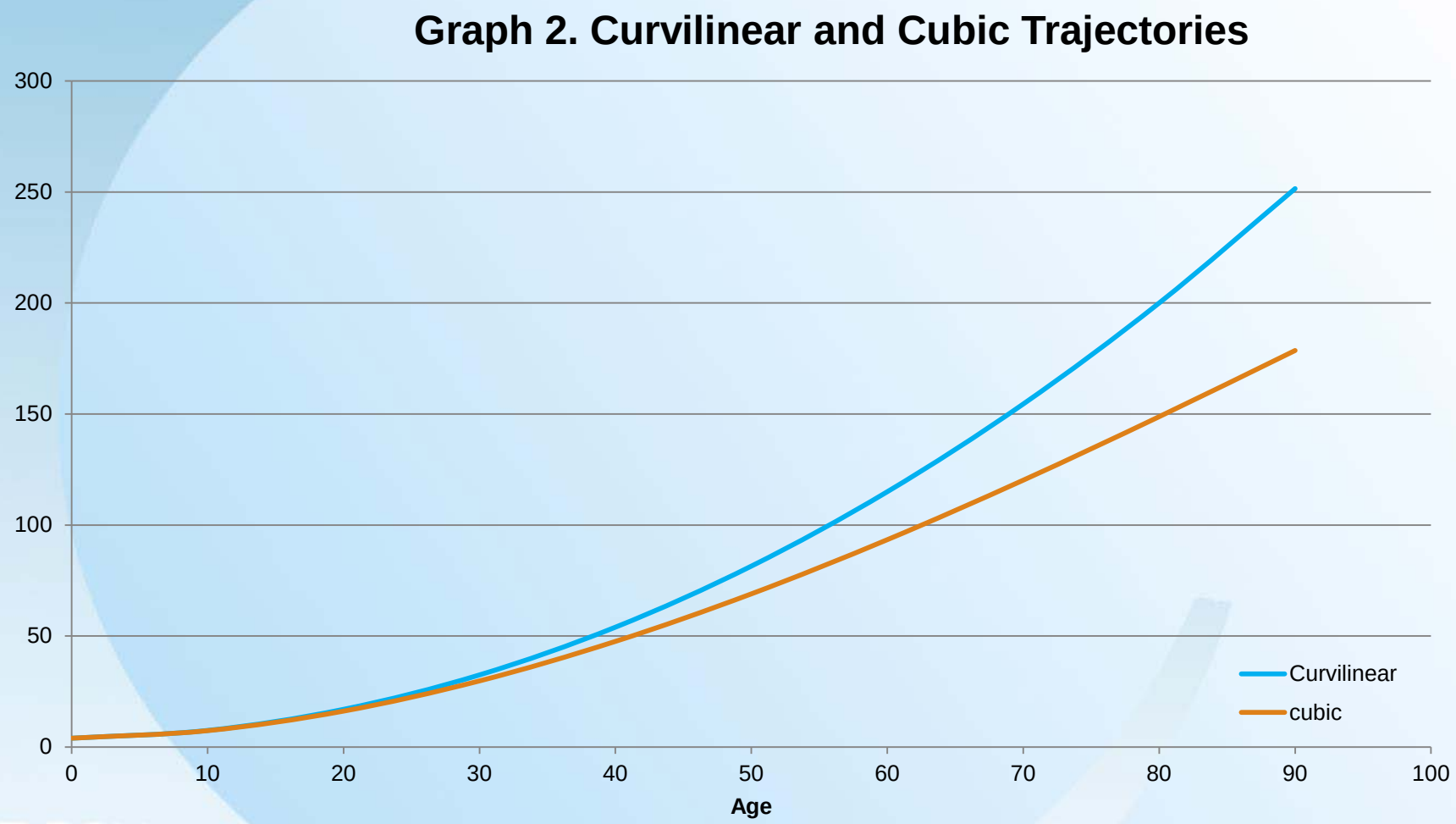
Table 1. Predicted Values for Four Different Trajectories				
Age	Linear	Discontinuity	Curvilinear	Vubic
0	4	4	4	4
10	4.5	4.5	7.5	7.4
20	5	5	17	16.2
30	5.5	5.5	32.5	29.8
40	6	6	54	47.6
50	6.5	9	81.5	69
60	7	10	115	93.4
70	7.5	11	154.5	120.2
80	8	12	200	148.8
90	8.5	13	251.5	178.6

Linear and Piecewise Linear Trajectories



BGSU Center for Family and Demographic Research

Curvilinear and Cubic Trajectories



BGSU Center for Family and Demographic Research

Example of Inter-person Differences in Linear Trajectories

Table 2. Graphs of Inter-person Differences in Trajectories

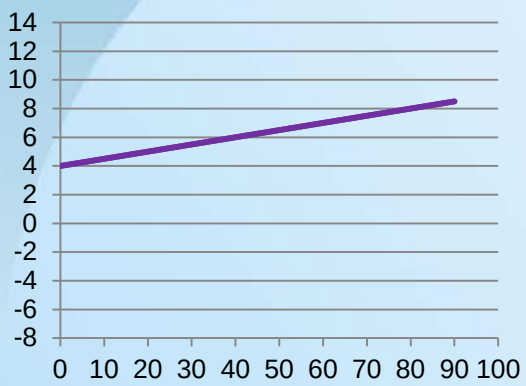
	Woman		Man	
	intercept	slope	intercept	slope
Same Intercept				
Same slope	4	0.05	4	0.05
Different slope	4	0.05	4	0.1
Different Intercep				
Same slope	4	0.05	3	0.05
Different slope	4	0.05	3	0.1
Different slope	4	0.05	3	-0.1

BGSU Center for Family and Demographic Research

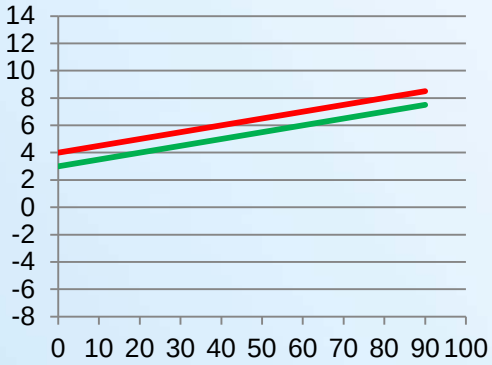
Table 2. Graphs of Inter-person Differences in Trajectories				
	Woman		Man	
	intercept	slope	intercept	slope
Same Intercept				
(1) Same Slpe	4	0.05	4	0.05
(2) Different Slpe	4	0.05	4	0.1
Different Intercep				
(3) Same Slpe	4	0.05	3	0.05
(4) Different Slpe	4	0.05	3	0.1
(5) Different Slpe	4	0.05	3	-0.1

Graphs of Inter-person Differences in Trajectories

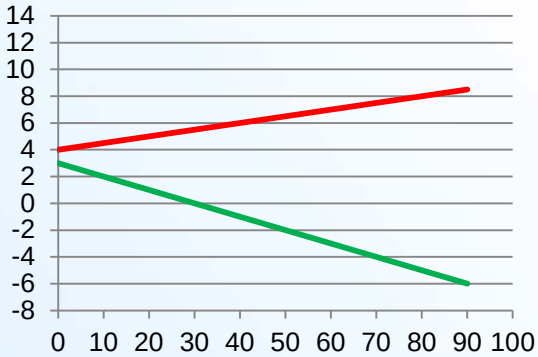
(1) Same Intercept and slope



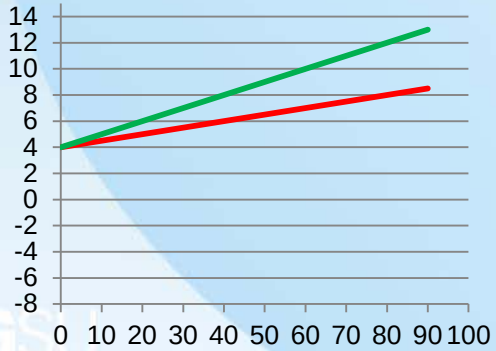
(3) Different Intercept and same slope



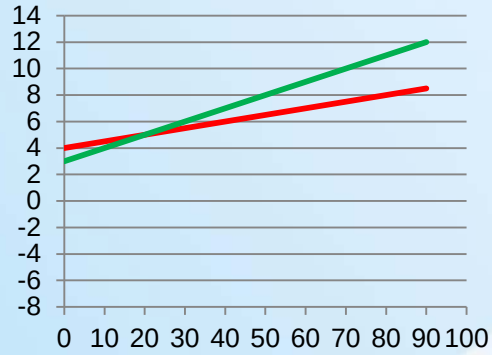
(5) Different Intercept and slope

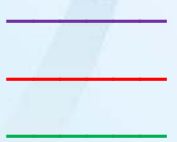


(2) Same Intercept and different slope



(4)Different Intercept and slope





Women and men
Women
Men

BGSU

Center for Family and Demographic Research

Key Concepts of Growth Curve Modeling

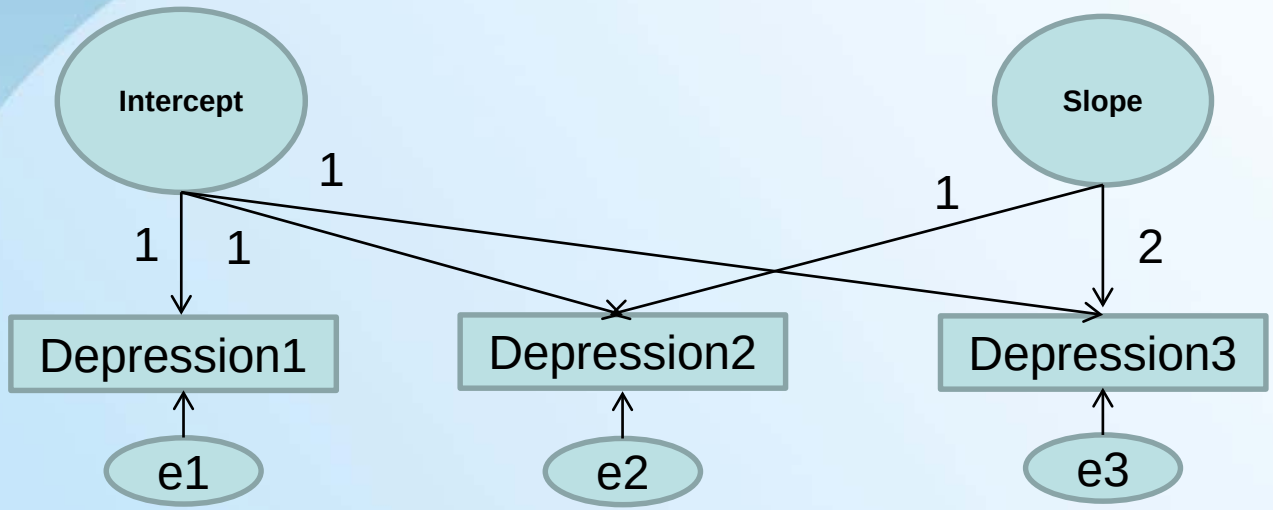
- Trajectory is a function of time.
- Trajectory can take on different functional forms (e.g., linear, curvilinear, cubic, and other forms).
- Trajectory describes whether individuals change over time (Intra-individual change) and how fast they change.
- The question of why people have different trajectories is equivalent to testing whether people with different attributes have different trajectories (i.e., inter-individual differences in the intra-individual change).
- If people have different trajectories, it indicates that they differ in at least one of the parameters that define the trajectories.

SEM and HLM Approaches to GCM

- SEM and HLM both can estimate Growth Curve Modeling, but specify the models differently.

- SEM and HLM generate the same results for two-level model (e.g., outcomes nested within individuals across time points) but not for more complex models (e.g., individuals are also nested with different geographic areas).

SEM models growth curve by conceptualizing the growth parameters as latent constructs and using factor loading as the effects of time.



- Corresponding Equations:

  Depression1 = Intercept + 0*Slope + e1

  Depression2 = Intercept + 1*Slope + e2

  Depression3 = Intercept + 2*Slope + e3

- The time variable is not an actual variable in the equations.
- These equations together indicate that depression has a linear relation with time
- SEM estimates the intercept and slope while effects of time is already specified as factor loadings.

HLM models growth curve by conceptualizing the variances of the outcome variables are embedded within individuals over time and time is an important predictor.

Level 1 Model:

$$Depression_{ij} = \beta_{0j} + \beta_{1j}(Time) + r_{ij}$$

Level 2 Model:

$$\beta_{0j} = \gamma_{00} + u_{0j}$$
$$\beta_{1j} = \gamma_{10} + u_{1j}$$

Full Model:

$$Depression_{ij} = \gamma_{00} + \gamma_{10}(Time) + u_{0j} + u_{1j}(Time) + r_{ij}$$

- Time is an actual variable in the equation.
- The effects of time is estimated by coefficients γ_{10} and $u_{1j.}$

Data formats for SEM and HLM

Table 3. SEM data format

id	female	dep1	dep2	dep3
1	0	4.9	4.7	4.6
2	0	4.4	4.3	5
3	0	4.7	5	4.9
49	1	4.1	4.8	4.6
50	1	4.1	4.4	5.1

Table 4. HLM data format

id	time	gender	depressior
1	1	0	4.9
1	2	0	4.7
1	3	0	4.6
2	1	0	4.4
2	2	0	4.3
2	3	0	5
3	1	0	4.7
3	2	0	5
3	3	0	4.9
49	1	1	4.1
49	2	1	4.8
49	3	1	4.6
50	1	1	4.1
50	2	1	4.4
50	3	1	5.1

Sample SEM and HLM Syntax in Stata

SEM:

```
sem (dep1 <- I@1 S@0 _cons@0) ///
        (dep2 <- I@1 S@1 _cons@0) ///
        (dep3 <- I@1 S@2 _cons@0), ///
         var(e.dep1@var e.dep2@var e.dep3@var) ///
         means(I S)
```

HLM:

```
mixed dep time || id:time, var cov(unstr)
```

Respective Strengths of SEM and HLM Approaches

- Strengths of SEM:
  - SEM can easily incorporate measurement models
  - SEM can easily model the mediational pathways

- Strengths of HLM:
  - HLM can easily handle large waves of data without computational problems
  - HLM can easily incorporate the time-varying variables in the growth curve model.

Conclusions

- Growth curve modeling is a statistical technique to describe and explain an individual's change over time
- Growth curve modeling requires at least three waves of panel data.
- Growth curve modeling can be estimated either by SEM or HLM approaches. These approaches will generate the same results if you just try to model the trajectory.
- The decision between using SEM or HLM may hinge on other considerations, including if an measurement model is involved, if mediational relations are modeled, how many waves of data are available, and whether researchers want to establish the association or causal relation among variables.
- If you have any questions about growth curve modeling, please come see me at 5D, Williams Hall or send me an email (wuh@bgsu.edu).