Lab 7 - Ten Growth Models - Simple to Advanced Structural Equation Modeling - Instructor: Karen Nylund-Gibson

$Adam\ Garber$

May 17, 2020

Contents

1	Lab	preparation	1					
	1.1	Creating a version-controlled R-Project with Github	2					
	1.2	Data source:	2					
	1.3	LSAY data example - Math Scores across 6 timepoints	2					
2	2 Let's start modeling							
	2.1	Model 01 - Fixed time effects	4					
	2.2	Model 02 - Centering the Intercept	7					
	2.3	$\operatorname{Model} 03$ - Freely estimated time scores	8					
	2.4	$\operatorname{Model}\ 04$ - Time-invariant covariates and freely estimated time scores	10					
	2.5	Model 05 - Time-varying covariates	11					
	2.6	$\operatorname{Model}\ 06$ - Time-varying covariate with time-invariant effect	12					
	2.7	Model 07 - Quadratic Growth	13					
	2.8	Model 08 - Parallel Process Growth Model	14					
	2.9	Model 09 - Piecewise Growth Model (type I)	16					
	2.10	Model 10 - Piecewise Growth Model (type II)	18					
3 Model fit summary								
4	Ref	erences	21					
Un	iver	sity of California, Santa Barbara						

1 Lab preparation

1.1 Creating a version-controlled R-Project with Github

 $Download\ repository\ here:\ https://github.com/garberadamc/SEM-Lab7$

On the Github repository webpage:

- a. fork your own branch of the lab repository
- b. copy the repository web URL address from the clone or download menu

Within R-Studio:

- c. click "NEW PROJECT"
- d. choose option Version Control
- e. choose option Git
- f. paste the repository web URL path copied from the clone or download menu on Github page
- g. choose location of the R-Project

1.2 Data source:

All examples utilize the public use dataset, The Longitudinal Survey of American Youth (LSAY) See documentation here

Load packages

```
library(transformr)
library(gganimate)
library(hrbrthemes)
library(tidyverse)
library(haven)
library(janitor)
library(MplusAutomation)
library(rhdf5)
library(here)
library(kableExtra)
library(gtsummary)
library(semPlot)
library(naniar)
```

1.3 LSAY data example - Math Scores across 6 timepoints

Read in data

```
lsay_data <- read_spss(here("data", "LSAY_labs.sav")) %>%
  select(RURAL, GENDER, FATHED, MOTHED,
        -starts_with("AB"),
                                                                              #
          ends_with("IMP"),
          -contains("BIO"),
          -contains("PHY")) %>%
  clean_names() %>%
  rename( math_07 = amthimp ,
         math_08 = cmthimp ,
          math_09 = emthimp ,
          math_10 = gmthimp ,
         math_11 = imthimp ,
          math_12 = kmthimp ,
           sci_07 = asciimp,
           sci_08 = csciimp ,
           sci_09 = esciimp,
           sci_10 = gsciimp ,
           sci_11 = isciimp ,
           sci_12 = ksciimp ) %>%
  replace_with_na_all(condition = ~.x == 9999.00)
```

View metadeta

```
sjPlot::view_df(lsay_data)
```

Write a CSV file

```
write_csv(lsay_data, here("data", "lsay_lab7_data.csv"))
```

Read in the CSV file (SPSS labels removed)

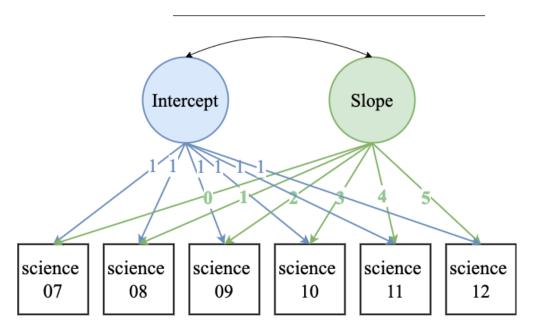
```
lsay_lab7 <- read_csv(here("data", "lsay_lab7_data.csv"))</pre>
```

2 Let's start modeling

Table. LSAY repeated measures

Name	Labels	Variable type
math_07	7th grade math score	time varying covariate
math_08	8th grade math score	
$math_09$	9th grade math score	
$math_10$	10th grade math score	
$math_11$	11th grade math score	
math_12	12th grade math score	
sci_07	7th grade science score	model indicators (outcomes)
sci_08	8th grade science score	
sci_09	9th grade science score	
sci_10	10th grade science score	
sci_11	11th grade science score	
sci_12	12th grade science score	

$2.1 \mod 01$ - Fixed time effects



```
m1_growth <- mplusObject(
  TITLE = "m01_fixed_time_scores",
  VARIABLE =
    "usevar =
    sci_07-sci_12; ",

ANALYSIS =
    "estimator = MLR",

MODEL =
    "i s | sci_07@0 sci_08@1 sci_09@2 sci_10@3 sci_11@4 sci_12@5; ",

OUTPUT = "sampstat standardized;",</pre>
```

Load in the mplus.R functions

```
source(here("mplus.R.txt"))
```

```
## [1] "Loaded rhdf5 package"
```

Plotting using gh5 data generated by Mplus

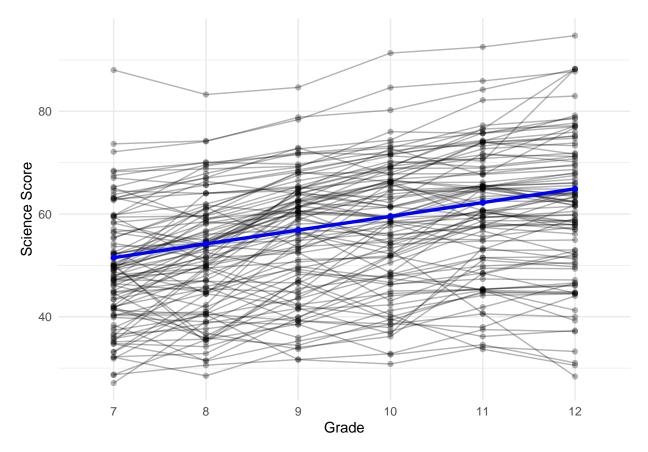
- 1. View plots available for a given model
- 2. Generate plots using the get.plot.___ function
- 3. Extract data and transform to tidy format
- 4. Plot with ggplot

```
mplus.view.plots(here("mplus_files", "m1_growth_Lab7.gh5"))
```

Prepare plot data

Plot the model estimated means superimposted on the obserbed individual values

```
growth_plot <- ggplot() +</pre>
                                                                                               #
  geom_point(data = plot_obs,
             aes(x = grade, y = value, group = rowname), alpha = .3) +
                                                                                               #
  geom_line(data = plot_obs,
            aes(x = grade, y = value, group = rowname), alpha = .3) +
                                                                                               #
  geom_point(data=mean_est,
                                                                                               #
             aes(x=grade, y = V1), color = "Blue", size = 1.5) +
                                                                                               #
 geom_line(data=mean_est,
                                                                                               #
            aes(x=grade, y = V1, group = 1), color = "Blue", size = 1.2) +
                                                                                               #
  scale_x_discrete(labels = c("7", "8", "9", "10", "11", "12")) +
                                                                                               #
  labs(x="Grade", y="Science Score") +
  theme_minimal()
growth_plot
```



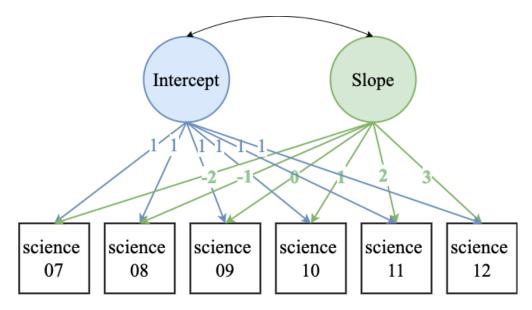
```
ggsave(here("figures", "spaghetti_p1.png"), height = 6, width = 8, dpi = "retina")
```

Animate the plot with {gganimate}

```
anim_save(here("figures", "spaghetti_plot.gif"), height = 6, width = 8, dpi = "retina")
```

$2.2 \mod 02$ - Centering the Intercept

- a. Centering determines the interpretation of the intercept growth factor
- b. The centering point is the timepoint at which the time score is zero
- c. A model can be estimated for different centering points depending on which interpretation is of interest



```
m2_growth <- mplusObject(
  TITLE = "m02_centering_time_scores",
  VARIABLE =
    "usevar =
    sci_07-sci_12; ",

ANALYSIS =
    "estimator = MLR" ,

MODEL =
    "i s | sci_07@-2 sci_08@-1 sci_09@0 sci_10@1 sci_11@2 sci_12@3; " ,

OUTPUT = "sampstat standardized;",

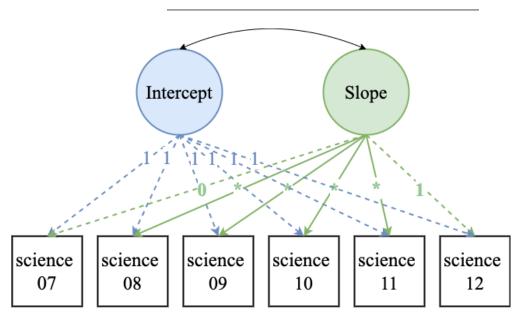
PLOT = "type=plot3;
    series = sci_07-sci_12(*)",

usevariables = colnames(lsay_lab7),
    rdata = lsay_lab7)

m2_growth_fit <- mplusModeler(m2_growth,</pre>
```

```
dataout=here("mplus_files", "Lab7.dat"),
modelout=here("mplus_files", "m2_growth_Lab7.inp"),
check=TRUE, run = TRUE, hashfilename = FALSE)
```

2.3 Model 03 - Freely estimated time scores

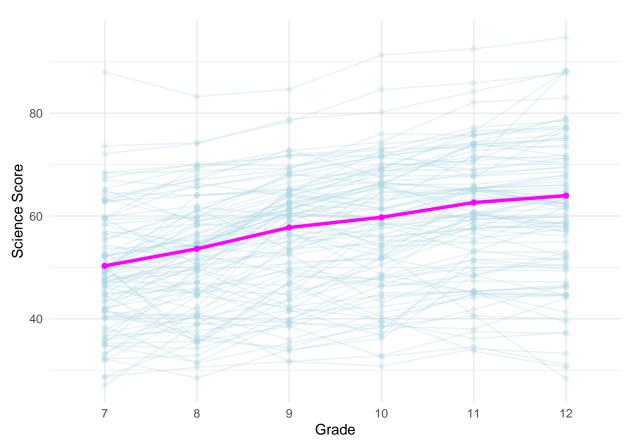


```
m3_growth <- mplusObject(</pre>
  TITLE = "m03_free_time_scores",
  VARIABLE =
    "usevar =
     sci_07-sci_12; ",
  ANALYSIS =
    "estimator = MLR" ,
  MODEL =
  "i s | sci_07@0 sci_08* sci_09* sci_10* sci_11* sci_12@1; " ,
  OUTPUT = "sampstat standardized;",
  PLOT = "type=plot3;
          series = sci_07-sci_12(*)",
  usevariables = colnames(lsay_lab7),
  rdata = lsay_lab7)
m3_growth_fit <- mplusModeler(m3_growth,
                     dataout=here("mplus_files", "Lab7.dat"),
                     modelout=here("mplus_files", "m3_growth_Lab7.inp"),
                     check=TRUE, run = TRUE, hashfilename = FALSE)
```

Prepare plot data

```
mean_est2 <- as.data.frame(mplus.get.estimated_means(here("mplus_files", "m3_growth_Lab7.gh5"))) %>%
    mutate(grade = gradelevels)
```

Plot the model estimated means superimposted on the obserbed individual values



$2.4~{ m Model}~04$ - Time-invariant covariates and freely estimated time scores

covariates:

```
gender: 1 = female
rural: 1 = rural
fathed: Father's reported education
mothed: Mother's reported education
```

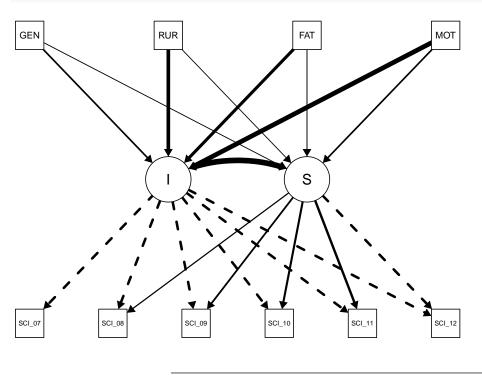
```
m4_growth <- mplusObject(
  TITLE = "m04_time-inv_covs",
  VARIABLE =
   "usevar =
    sci_07-sci_12
    gender rural fathed mothed; ",
  ANALYSIS =
    "estimator = MLR",
  DEFINE =
   "center fathed mothed (grandmean);",
  MODEL =
  "i s | sci_07@0 sci_08* sci_09* sci_10* sci_11* sci_12@1;
   is on gender rural fathed mothed;",
  OUTPUT = "sampstat standardized;",
 PLOT = "type=plot3;
          series = sci_07-sci_12(*)",
 usevariables = colnames(lsay_lab7),
 rdata = lsay_lab7)
m4_growth_fit <- mplusModeler(m4_growth,</pre>
                     dataout=here("mplus_files", "Lab7.dat"),
                     modelout=here("mplus_files", "m4_growth_Lab7.inp"),
                     check=TRUE, run = TRUE, hashfilename = FALSE)
```

Check the path diagram with semPlot

```
m4_output <- readModels(here("mplus_files", "m4_growth_Lab7.out"))</pre>
```

Reading model: /Users/agarber/Desktop/SEM_S20/Lab7_SEM/mplus_files/m4_growth_Lab7.out

```
semPaths(m4_output, "est",
    intercepts=FALSE, residuals = FALSE, fade = FALSE,
    edge.color = "black", edgeLabels = "") #
```



2.5 Model 05 - Time-varying covariates

repeated measure covariate: math scores: grades 7 to 12 time-invariant covariate: mothed

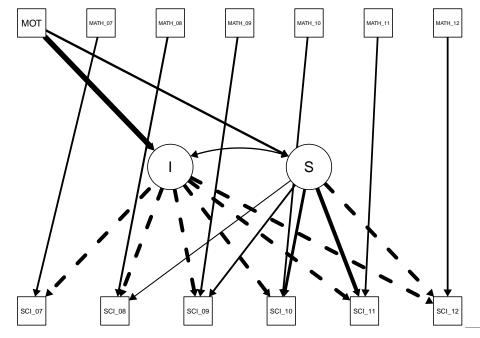
```
m5_growth <- mplusObject(
  TITLE = "m05_time-varying_covs",
  VARIABLE =
    "usevar =
    sci_07-sci_12
    math_07-math_12 mothed; ",

ANALYSIS =
    "estimator = MLR" ,

DEFINE =
    "center mothed (grandmean);",

MODEL =
    "i s | sci_07@0 sci_08* sci_09* sci_10* sci_11* sci_12@1;
    i s on mothed;
    sci_07 on math_07;
    sci_08 on math_08;
    sci_09 on math_09;</pre>
```

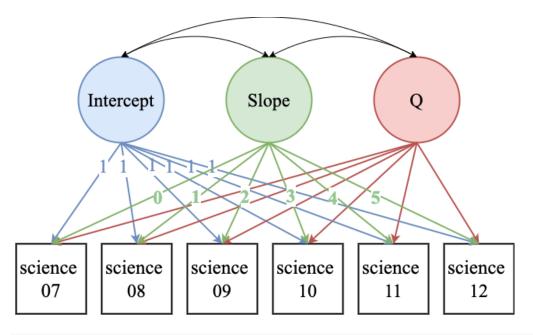
Check the path diagram



 $2.6~{
m Model}~06$ - Time-varying covariate with time-invariant effect

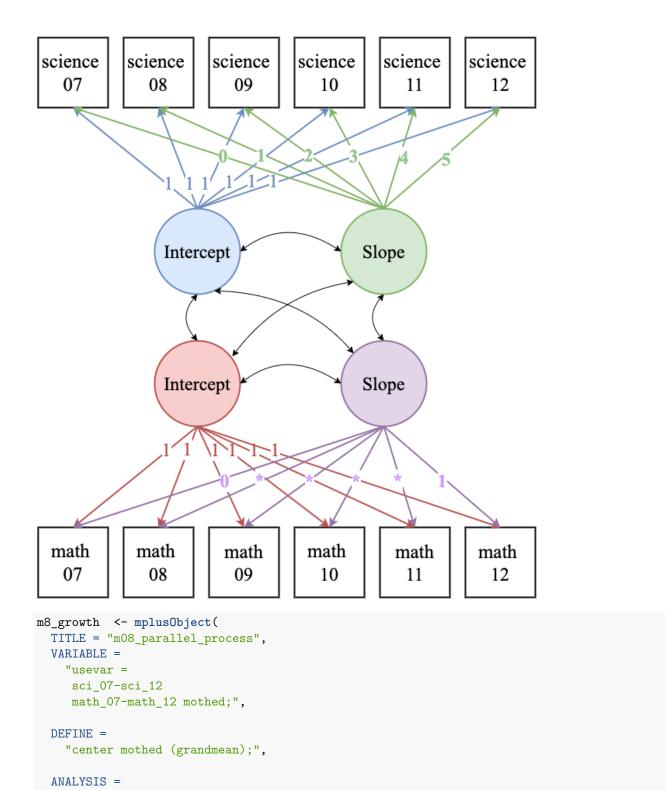
```
m6_growth <- mplusObject(</pre>
  TITLE = "m06_time-varying_covs",
  VARIABLE =
    "usevar =
    sci_07-sci_12
     math_07-math_12 mothed; ",
  ANALYSIS =
    "estimator = MLR" ,
  DEFINE =
    "center mothed (grandmean);",
  MODEL =
  "i s | sci_07@0 sci_08* sci_09* sci_10* sci_11* sci_12@1;
   i s on mothed;
   sci_07 on math_07(1); ! TIME-INVARIANT: Fixed to equality
   sci_08 on math_08(1);
   sci_09 on math_09(1);
   sci_10 on math_10(1);
   sci_11 on math_11(1);
   sci_12 on math_12(1); ",
  OUTPUT = "sampstat standardized;",
  PLOT = "type=plot3;
          series = sci_07-sci_12(*)",
  usevariables = colnames(lsay_lab7),
 rdata = lsay_lab7)
m6_growth_fit <- mplusModeler(m6_growth,</pre>
                     dataout=here("mplus_files", "Lab7.dat"),
                     modelout=here("mplus_files", "m6_growth_Lab7.inp"),
                     check=TRUE, run = TRUE, hashfilename = FALSE)
```

2.7 Model 07 - Quadratic Growth



```
m7_growth <- mplusObject(</pre>
  TITLE = "m07_quadratic",
  VARIABLE =
    "usevar =
     sci_07-sci_12; ",
  ANALYSIS =
    "estimator = MLR" ,
  MODEL =
  "i s q | sci_0700 sci_0801 sci_0902 sci_1003 sci_1104 sci_1205; ",
  OUTPUT = "sampstat standardized tech4;",
  PLOT = "type=plot3;
          series = sci_07-sci_12(*)",
  usevariables = colnames(lsay_lab7),
  rdata = lsay_lab7)
m7_growth_fit <- mplusModeler(m7_growth,</pre>
                     dataout=here("mplus_files", "Lab7.dat"),
                     modelout=here("mplus_files", "m7_growth_Lab7.inp"),
                      check=TRUE, run = TRUE, hashfilename = FALSE)
```

2.8 Model 08 - Parallel Process Growth Model



"is ss | sci_07@0 sci_08@1 sci_09@2 sci_10@3 sci_11@4 sci_12@5;

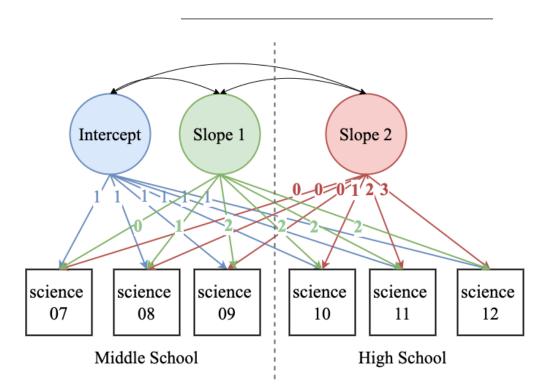
is ss im sm on mothed; ! time-invariant covariate ",

im sm | math_0700 math_0801 math_0902 math_1003 math_1104 math_1205;

"estimator = MLR" ,

MODEL =

2.9 Model 09 - Piecewise Growth Model (type I)



```
m9_growth <- mplusObject(
  TITLE = "m09_piecewise_1int_2slopes",
  VARIABLE =
    "usevar =
    sci_07-sci_12 mothed;",

ANALYSIS =
    "estimator = MLR" ,

DEFINE =</pre>
```

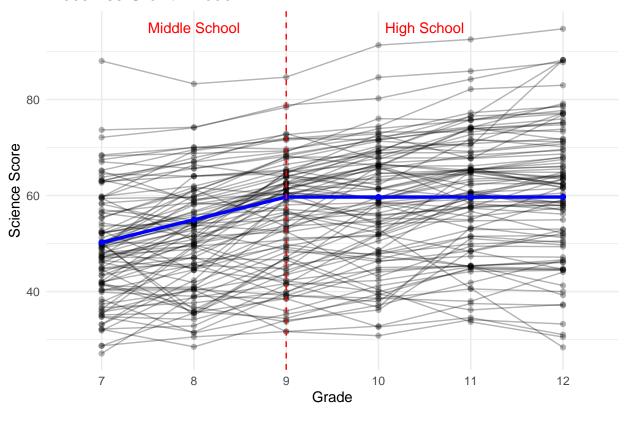
Prepare plot data

```
mean_est3 <- as.data.frame(mplus.get.estimated_means(here("mplus_files", "m9_growth_Lab7.gh5"))) %>%
    mutate(grade = gradelevels)
```

Plot the model estimated means superimposted on the obserbed individual values

```
growth_plot <- ggplot() +</pre>
                                                                                              #
  geom_point(data = plot_obs,
             aes(x = grade, y = value, group = rowname), alpha = .3) +
                                                                                              #
                                                                                              #
  geom_line(data = plot_obs,
            aes(x = grade, y = value, group = rowname), alpha = .3) +
                                                                                              #
  geom_point(data=mean_est3,
                                                                                              #
             aes(x=grade, y = V1), color = "Blue", size = 1.5) +
                                                                                              #
  geom_line(data=mean_est3,
            aes(x=grade, y = V1, group = 1), color = "Blue", size = 1.2) +
  geom_vline(xintercept=3, linetype="dashed", color = "red") +
  annotate("text", x = 2, y = 95, label = "Middle School", color="red") +
  annotate("text", x = 4.5, y = 95, label = "High School", color="red") +
  scale_x_discrete(labels = c("7", "8", "9", "10", "11", "12")) +
  labs(title= "Piecewise Growth Model", x="Grade", y="Science Score") +
  theme_minimal()
growth_plot
```

Piecewise Growth Model



$2.10 \mod 10$ - Piecewise Growth Model (type II)

```
m10_growth <- mplusObject(
    TITLE = "m10_piecewise_1int_2slopes",
    VARIABLE =
        "usevar =
            sci_07-sci_12 mothed;",

ANALYSIS =
        "estimator = MLR;" ,

DEFINE =
        "center mothed (grandmean);",

MODEL =
        "i1 s1 | sci_0700 sci_0801 sci_0902;
        i2 s2 | sci_1000 sci_1101 sci_1202;

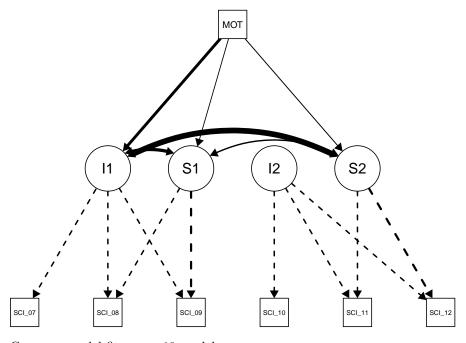
[s1] (p1);
[s2] (p2);</pre>
```

Check the path diagram

```
m10_output <- readModels(here("mplus_files", "m10_growth_Lab7.out"))</pre>
```

Reading model: /Users/agarber/Desktop/SEM_S20/Lab7_SEM/mplus_files/m10_growth_Lab7.out

```
semPaths(m10_output, "est", intercepts = FALSE, residuals = FALSE, fade = FALSE,
edge.color = "black", edgeLabels = "")
```



Compare model fit across 10 models

```
all_models <- readModels(here("mplus_files"))</pre>
```

Reading model: /Users/agarber/Desktop/SEM_S20/Lab7_SEM/mplus_files/m1_growth_Lab7.out

```
## Reading model: /Users/agarber/Desktop/SEM_S20/Lab7_SEM/mplus_files/m10_growth_Lab7.out
## Reading model: /Users/agarber/Desktop/SEM_S20/Lab7_SEM/mplus_files/m2_growth_Lab7.out
## Reading model: /Users/agarber/Desktop/SEM_S20/Lab7_SEM/mplus_files/m3_growth_Lab7.out
## Reading model: /Users/agarber/Desktop/SEM_S20/Lab7_SEM/mplus_files/m4_growth_Lab7.out
## Reading model: /Users/agarber/Desktop/SEM_S20/Lab7_SEM/mplus_files/m5_growth_Lab7.out
## Reading model: /Users/agarber/Desktop/SEM_S20/Lab7_SEM/mplus_files/m6_growth_Lab7.out
## Reading model: /Users/agarber/Desktop/SEM_S20/Lab7_SEM/mplus_files/m7_growth_Lab7.out
## Reading model: /Users/agarber/Desktop/SEM_S20/Lab7_SEM/mplus_files/m8_growth_Lab7.out
## Reading model: /Users/agarber/Desktop/SEM_S20/Lab7_SEM/mplus_files/m9_growth_Lab7.out
## Reading model: /Users/agarber/Desktop/SEM_SEM_SEM_Mplus_files/m9_growth_Lab7.out
## Reading model: /Users/agarber/Desktop/SEM_SEM_SEM_Mplus_files/m9_growth_Lab7.out
## Reading model: /Users/agarber/Desktop/SEM_SEM_Mplus_files/m9_growth_Lab7.out
## Reading m
```

3 Model fit summary

Make a table with gt::gt()

```
library(gt)

fit_table <- summary_fit %>%
  gt() %>%
  tab_header(
    title = "Fit Indices", # Add a title
    subtitle = "" # And a subtitle
) %>%
  tab_options(
    table.width = pct(80)
) %>%
  tab_footnote(
    footnote = "Longitudinal Survey American Youth (LSAY)",
    location = cells_title())
```

Fit Indices¹

Title	Parameters	ChiSqM_Value	CFI	TLI	SRMR	RMSEA_Estimate	RMSEA_
m01_fixed_time_scores	11	1064.490	0.938	0.942	0.076	0.106	
m02_centering_time_scores	11	1064.492	0.938	0.942	0.076	0.106	
$m03_free_time_scores$	15	408.801	0.977	0.971	0.063	0.075	
$m04_time-inv_covs$	23	464.948	0.971	0.960	0.041	0.075	
$m05_time-varying_covs$	23	1398.808	0.906	0.884	0.338	0.127	
$m06_time-varying_covs$	18	1540.135	0.897	0.884	0.340	0.126	
$m07$ _quadratic	15	226.244	0.987	0.984	0.036	0.055	
$m08_parallel_process$	30	3138.892	0.936	0.930	0.061	0.086	
$m09$ _piecewise_1int_2slopes	17	2405.521	0.876	0.837	0.180	0.161	
$m10_piecewise_1int_2slopes$	20	3806.656	0.803	0.682	0.397	0.226	

¹Longitudinal Survey American Youth (LSAY)

4 References

Hallquist, M. N., & Wiley, J. F. (2018). MplusAutomation: An R Package for Facilitating Large-Scale Latent Variable Analyses in Mplus. Structural equation modeling: a multidisciplinary journal, 25(4), 621-638.

Miller, J. D., Hoffer, T., Suchner, R., Brown, K., & Nelson, C. (1992). LSAY codebook. Northern Illinois University.

Muthén, B. O., Muthén, L. K., & Asparouhov, T. (2017). Regression and mediation analysis using Mplus. Los Angeles, CA: Muthén & Muthén.

Muthén, L.K. and Muthén, B.O. (1998-2017). Mplus User's Guide. Eighth Edition. Los Angeles, CA: Muthén & Muthén

R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/

Wickham et al., (2019). Welcome to the tidy verse. Journal of Open Source Software, 4(43), 1686, https://doi.org/10.21105/joss.01686

UC **SANTA BARBARA**