Using Mplus To Do Latent Transition Analysis And Random Intercept Latent Transition Analysis

Bengt Muthén
bmuthen@statmodel.com
Mplus: www.statmodel.com

Mplus Web Talks: No. 2

Recorded February 2021

This web talk can be referred to as:

Muthén, Bengt [Mplus]. (2021, February 28). Using Mplus To Do Latent Transition Analysis And Random Intercept Latent Transition Analysis [Video file]. Retrieved from https://www.youtube.com/c/MplusVideos

I thank Tihomir Asparouhov and Thuy Nguyen for expert programming of the new features in this talk, Tihomir for insightful contributions to the statistics, Linda Muthén for helpful comments, and Noah Hastings for great assistance with the production.

Mplus Web Talks Page

- YouTube video segments and handouts available at `http://www.statmodel.com/MplusWebTalks.shtml`
- Companion piece: Mplus Web Talk No. 1
  - LTA and RI-LTA analyses: Segments 11 - 14

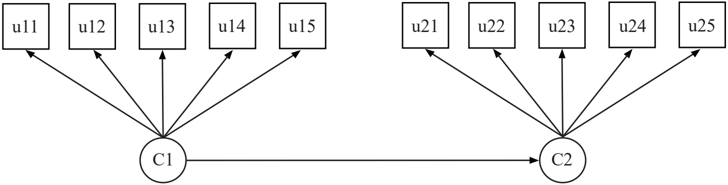
Outline

- An example
- Basic building blocks
- Analysis without covariates
    - Regular LTA
    - RI-LTA
    - Comparing results. Deciding on the number of classes
    - Checking model fit and model modifications
        - Checking response pattern fit and bivariate fit
        - Measurement invariance across time
        - Residual associations across time
        - Lag-2 modeling
- Adding covariates. Covariate effects on transition probabilities
    - RI-LTA. Comparing results with Regular LTA
- Measurement invariance across individuals
    - Multiple-group analysis and direct effects of covariates on indicators
- Special topics
    - Modeling stationarity of transition probabilities
    - Mover-Stayer modeling
    - Distal outcome
- Error messages

Reading Proficiency from Kindergarten to First Grade

- Early Childhood Longitudinal Study (ECLS-K), N = 3574
- T=4: Fall and Spring of Kindergarten and Fall and Spring of Grade 1
- 5 binary items representing mastery of:
  - Basic reading skills of letter recognition
  - Beginning sounds
  - Ending sounds
  - Sight words
  - Words in context
- 3 latent classes corresponding to 3 stages of learning:
  - Low alphabet knowledge, early word reading, early reading comprehension
- Covariate: Poverty. Child's household is below (Poverty=1) or above (Poverty = 0) the U.S. census poverty threshold
- Kaplan (2008). An overview of Markov chain methods for the study of stage-sequential developmental processes. Developmental Psychology

Features of Regular LTA



1. Measurement probabilities: $P(U_t|C_t)$ - LCA for each time point t, measurement invariance across time
2. Initial status probabilities: $P(C_1)$
3. Transition probabilities: $P(C_2|C_1)$

Reading Data Measurement Probability Estimates. Regular LTA, T=4

Class 1 = low alphabet knowledge, Class 2 = early word reading, Class 3 = early reading comprehension

	Classes		
	1	2	3
Letrec	0.505	0.994	1.000
Begin	0.066	0.917	0.984
Ending	0.013	0.660	0.972
Sight	0.000	0.051	0.985
WIC	0.000	0.000	0.509

Same for all time points. Provided in the output under the heading LATENT CLASS INDICATOR MEANS AND PROBABILITIES FOR EACH LATENT CLASS

Reading Data Transition Probabilities. Regular LTA, T=4

- Estimated transition matrix for the first two time points, transitioning from Fall K to Spring K (C1 $\rightarrow$ C2):

Latent classes	1	2	3
1	0.338	0.649	0.012
2	0.001	0.652	0.348
3	0.000	0.000	1.000

Provided in the output under the heading
LATENT TRANSITION PROBABILITIES BASED ON THE ESTIMATED MODEL

Time 1, Time 2 and Transition Probabilities

Latent classes	Time 1 (C1)	Transition matrix 1	2	3	Time 2 (C2)
1	0.694	0.338	0.649	0.012	0.235
2	0.284	0.001	0.652	0.348	0.635
3	0.023	0.000	0.000	1.000	0.130

$$\hat{P}(C2=1) = 0.694 \times 0.338 + 0.284 \times 0.001 + 0.023 \times 0.000 = 0.235$$
$$\hat{P}(C2=2) = 0.694 \times 0.649 + 0.284 \times 0.652 + 0.023 \times 0.000 = 0.635$$
$$\hat{P}(C2=3) = 0.694 \times 0.012 + 0.284 \times 0.348 + 0.023 \times 1.000 = 0.130$$

Provided in the output under the heading
FINAL CLASS COUNTS AND PROPORTIONS FOR EACH LATENT CLASS VARIABLE BASED ON THE ESTIMATED MODEL

Typical Transition Paths. Regular LTA, T=4

- Estimated frequencies for the 3 most frequent paths through the latent class variables over the 4 time points (entries are latent classes):

Path	Fall K	Spring K	Fall 1st	Spring 1st	Freq.	% of N
1	1	2	2	3	1163	33
2	2	2	2	3	477	13
3	2	3	3	3	350	10

Provided in the TECH15 output under the heading
FINAL CLASS COUNTS AND PROPORTIONS FOR THE LATENT CLASSES BASED ON THE ESTIMATED MODEL

Outline

- An example
- **Basic building blocks**
- Analysis without covariates
  - Regular LTA
  - RI-LTA
  - Comparing results. Deciding on the number of classes
  - Checking model fit and model modifications
    - Checking response pattern fit and bivariate fit
    - Measurement invariance across time
    - Residual associations across time
    - Lag-2 modeling
- Adding covariates. Covariate effects on transition probabilities
  - RI-LTA. Comparing results with Regular LTA
- Measurement invariance across individuals
  - Multiple-group analysis and direct effects of covariates on indicators
- Special topics
  - Modeling stationarity of transition probabilities
  - Mover-Stayer modeling
  - Distal outcome

Basic Building Blocks

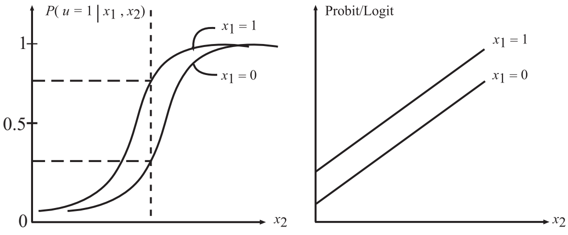
- Logistic regression (logit parameterization): Binary, ordinal, and nominal DV
- Probabilities
- Odds: Single variable, ratio of probabilities
- Odds ratio: Two variables, ratio of odds

The Mplus modeling uses a logit parameterization. Probabilities, odds, and odds ratios provided in the output

Sampling Distributions of Logits, Probabilities, Odds, and Odds Ratios

- Logits vary from $-\infty$ to $+\infty$ and are approximately normally distributed:
  - Can use symmetric confidence intervals (estimate $\pm 1.96 \times SE$)
- Probabilities vary from 0 to 1 with non-normal distributions:
  - Need to use non-symmetric confidence intervals
  - Non-symmetric interval: Distances between the limits and the estimate are different
- Odds and odds ratios vary from 0 to $+\infty$ with 1 being the neutral point. They have non-normal distributions:
  - Need to use non-symmetric confidence intervals
  - Significant if the confidence interval does not contain 1
- Non-symmetric confidence interval limits for odds and odds ratios are obtained by exponentiating the confidence interval limits of logits

A Quick Reminder of Logistic Regression



- Logistic regression where U is a binary DV, X_1 a binary covariate, and X_2 a continuous covariate (e.g. cancer, smoking, age)

$$P(U_i = 1 | X_{1i}, X_{2i}) = \frac{1}{1 + e^{-Logit}} = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i})}}.$$

- Two other examples:
    - DV is a binary latent class indicator, X_1 a binary latent class variable, and X_2 a direct effect from a covariate
    - DV is a latent class variable with 2 classes and X_1, X_2 are covariates predicting the latent class variable

Observed or Latent Nominal Dependent Variable

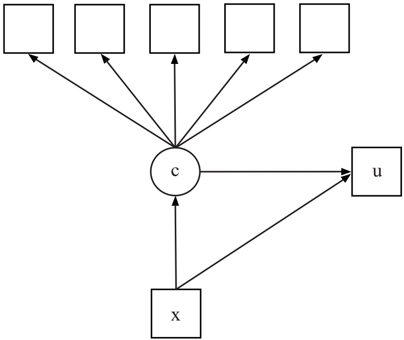
- How do you generalize the case of a binary DV to more than 2 unordered categories, that is, a nominal variable?
- Note that $P(U = 1|X) = 1/[1 + e^{-(a+bX)}] = e^{a+b\,X}/[e^{a+bX} + 1]$
- Regression of a nominal variable U with R categories on a covariate X (multinominal logistic regression):
  - $P(U = r|X) = e^{a_r+b_r\,X}/[e^{a_1+b_1X} + e^{a_2+b_2X} + \ldots + e^{a_{R-1}+b_{R-1}X} + 1]$
  - The last category is the reference category so that $a_R = 0, b_R = 0$, that is, $e^{a_R+b_RX} = e^0 = 1$
  - Odds ratio: e^{b_r} representing the odds $P(U = r|X)/P(U = R|X)$ for X=x divided by the odds for X = x-1 (a one-unit change), e.g., X=1 versus X=0 for a binary X
- The same multinomial logistic regression formulation is used for regression with a latent nominal latent class variable as DV

- C1: Latent class variable at first time point of an LTA. Not influenced by other C's
- $P(C1 = r|X) = e^{a_r + b_r X} / [e^{a_1 + b_1 X} + e^{a_2 + b_2 X} + \ldots + e^{a_{R-1} + b_{R-1} X} + 1]$
- Odds ratio: e^{b_r} representing the odds $P(C1 = r|X)/P(C1 = R|X)$ for X=x divided by the odds for X = x-1 (a one-unit change), e.g., X=1 versus X=0 for a binary X
- Changing reference class from the last to the first class - subtract $a_1 + b_1 X$ so that the logit is zero for the first class:
  - $P(C1 = r|X) = e^{a_r - a_1 + (b_r - b_1)X} / [e^{a_1 - a_1 + (b_1 - b_1)X} + e^{a_2 - a1 + (b_2 - b_1)X} + \ldots + e^{a_{R-1} - a_1 + (b_{R-1} - b_1)X} + e^{-a_1 - b_1 X}]$
  - Can be computed in MODEL CONSTRAINT but can also be done automatically by re-ordering the latent classes

Re-Ordering the Latent Classes for C1 ON X

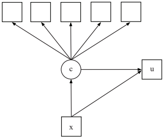
- Assume that we want to see the classes in the order Low, Medium, High that we refer to as 1, 2, 3
- Assume that the run gets the order Medium, High, Low (2 3 1)
- Re-order: 2 3 1 $\rightarrow$ 3 1 2 to get the order Low, Medium, High
- This is done using OPTSEED and SVALUES (3 1 2)
- Re-ordering can be used to highlight different parts of the modeling. Example of C1 regressed on the binary X variable Pov (Pov=0/1):
  - Low, Medium, High (High is the reference class): OR for C1#1 ON Pov = 8.101 [2.539 25.850] ("significant", that is, not covering 1)
    - Odds of being in Low class relative to High class at time 1 is much higher for Pov=1 than Pov=0 (8.101 times higher)
  - Keeping the order in the first run of Medium, High, Low (Low is the reference class): OR for C1#1 ON Pov = 0.232 [0.174 0.311] ("significant", that is, not covering 1)
    - Odds of being in Medium class relative to Low class at time 1 is much lower for Pov=1 than Pov=0

Example: Distal Outcome in Latent Class Analysis ORs and Their Non-Symmetric Confidence Intervals



Distal Outcome Continued

- Odds and odds ratios with two predictors:
  - General case - logit regression (logit = logodds):
  - $Logodds(U|X_1, X_2) = a + b_1\, X_1 + b_2\, X_2$
  - Odds ratios interpretation:
    - OR effect of X_1 is e^{b_1} irrespective of the value of X_2
    - OR effect of X_2 is e^{b_2} irrespective of the value of X_1
  - Latent class model:



  - Logit regression with X and a latent class predictor:
    - $logodds(U|X, C = c) = a_c + b\, X = -t_c + b\, X$ for latent classes $c = 1, 2 \ldots C$
    - OR effect of X is e^b irrespective of the latent class
    - OR effect of latent class is expressed in terms of the U odds for $c = 1$ divided by the U odds for $c = C$ and is not influenced by the value of X

OR for Class Effect: Mplus Script with X = Pov, U = Begin2, Latent Class Indicators = Lectrec1-Wic1

```
VARIABLE:
        ...
        USEVARIABLES = pov
        letrec1-wic1 begin2;
        CATEGORICAL =
        letrec1-wic1 begin2;
        CLASSES = c(3);
ANALYSIS:
        TYPE = MIXTURE;
        STARTS = 160 40;
        PROCESSORS = 8;
MODEL:
        %OVERALL%
        c ON pov;
        begin2 ON pov (b);
        %c#1%
        [begin2$1] (t1);
        [letrec1$1-wic1$1];
        %c#2%
        [begin2$1] (t2);
        [letrec1$1-wic1$1];
        %c#3%
        [begin2$1] (t3);
        [letrec1$1-wic1$1];
MODEL CONSTRAINT:
        NEW(p1 p2 p3 odds1 odds2 odds3 or13
        diff13);
        ! Pov = 0:
        p1 = 1/(1+exp(t1));
        p2 = 1/(1+exp(t2));
        p3 = 1/(1+exp(t3));
        odds1 = p1/(1-p1);
        odds2 = p2/(1-p2);
        odds3 = p3/(1-p3);
        or13 = odds1/odds3;
        ! Non-symmetric CI for or13 based on
        ! the estimate and SE of:
        ! log or13 = log odds1 - log odds3 =
        ! = -t1 - (-t3)
        diff13 = -t1 + t3;
        ! diff13 is the same for Pov = 1
        ! because the slope b cancels out
```

Computing Non-Symmetric Confidence Intervals for ORs by Exponentiating the Confidence Limits of the Logit Difference for the Distal Outcome

- Estimated diff13 = -0.948, SE = 0.431
- Symmetric 95 percent confidence interval for diff13:
  $-0.948 \pm 1.96 \times 0.431$, that is, CI = [-1.793, -0.103]
- Exp(diff13) = OR13 = 0.388. Non-symmetric confidence interval needed - compute them by exponentiating the symmetric limits:
  - Lower confidence interval limit = exp(-1.793) = 0.166
  - Upper confidence interval limit = exp(-0.103) = 0.902
  - Because the CI does not cover 1, OR13 is significant
- The output shows these results under the heading:
  - LATENT CLASS INDICATOR ODDS RATIOS FOR THE LATENT CLASSES

C2 Regressed on C1: Nominal DV and Predictor Mplus Default: Logit Parameterization

		C_2		
		1	2	3
	1	$a_1 + b_{11}$	$a_2 + b_{21}$	0
C_1	2	$a_1 + b_{12}$	$a_2 + b_{22}$	0
	3	a_1	a_2	0

a_1: [C2#1]
a_2: [C2#2]
b_{11}: C2#1 ON C1#1
b_{12}: C2#1 ON C1#2
b_{21}: C2#2 ON C1#1
b_{22}: C2#2 ON C1#2

- Mplus parameterization:
  - a and b logit parameters
- Short-hand specification:
  - C2 ON C1
  - Gives all 6 parameters

C2 Regressed on C1: What does P(C2 | C1) Mean? Logit Parameterization: Last Class is Reference Class

		C_2		
		1	2	3
	1	$a_1 + b_{11}$	$a_2 + b_{21}$	0
C_1	2	$a_1 + b_{12}$	$a_2 + b_{22}$	0
	3	a_1	a_2	0

- Each row is a multinomial model parameterized by logits, producing transition probabilities for each cell in the 3 x 3 table
- For example, consider $a_2 + b_{21}$ in row 1 and column 2:
  - $P(C_2 = 2 \mid C_1 = 1) = e^{a_2+b_{21}}/(e^{a_1+b_{11}} + e^{a_2+b_{21}} + 1)$
- Odds $P(C_2 = 2 \mid C_1 = 1)/P(C_2 = 3 \mid C_1 = 1) = e^{a_2+b_{21}}$
  - Odds of transitioning from class 1 to class 2 versus transitioning from class 1 to class 3 (very high for Reading example)

Changing Reference Class to Stayers: Zero Logits on the Diagonal of the Transition Table

		C_2		
		1	2	3
	1	0	$a_2 + b_{21}$ - $(a_1 + b_{11})$	-$(a_1 + b_{11})$
C_1	2	$a_1 + b_{12}$ -($a_2 + b_{22}$)	0	-($a_2 + b_{22}$)
	3	a_1	a_2	0

- Suitable for considering the odds
  - $P(C_2 = 2 \mid C_1 = 1)/P(C_2 = 1 \mid C_1 = 1)$
    Odds of transitioning from class 1 to class 2 versus staying in class 1
- Produced automatically in the output

Example: Transition Odds for $C1 \rightarrow C2$

- Estimated transition matrix for the first two time points. Probability of transitioning from Fall K to Spring K (C1 $\rightarrow$ C2):

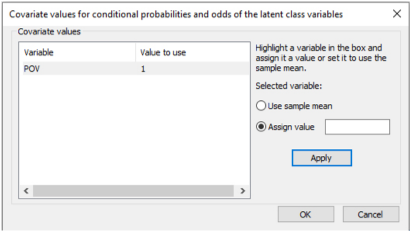
Latent classes	1	2	3
1	0.338	**0.649**	0.012
2	0.001	0.652	0.348
3	0.000	0.000	1.000

Transition table odds and 95% confidence intervals for C1 to C2		
1.000 (1.000,1.000)	**1.920 (1.751,2.104)**	0.037 (0.022,0.062)
0.001 (0.000,0.055)	1.000 (1.000,1.000)	0.534 (0.458,0.621)
0.000 (0.000,0.000)	0.000 (0.000,0.000)	1.000 (1.000,1.000)

The odds of transitioning from class 1 to class 2 is 1.920 (0.649/0.338)

Calculator: Transition Probabilities for Different Covariate Values

- Calculator formerly called LTA Calculator
- Available from the Mplus menu



Calculator: Transition Probabilities for Different Covariate Values

- Example using regular LTA with C ON X:
  - Results for Pov=1 and Pov=0, transitioning from C3 to C4:

$C3 \to C4$	Pov=1 Fall 1st-Spring 1st			Pov=0 Fall 1st - Spring 1 st		
	1	2	3	1	2	3
1	0.301	0.542	0.157	0.232	0.464	0.305
2	0.010	**0.226**	**0.764**	0.004	**0.115**	**0.881**
3	0.001	0.000	0.999	0.000	0.000	1.000

- OR = $\frac{0.764/0.226}{0.881/0.115} = \frac{3.381}{7.66} = 0.44$
- Produced automatically in the output with CIs; see next slide

Example Continued: Odds Ratio Output for Effects of Pov Covariate

Transition table odds ratio and 95% confidence intervals for C3 to C4

1.000 (1.000,1.000)	0.901 (0.593, 1.370)	0.396 (0.251, 0.626)
1.109 (0.730, 1.685)	1.000 (1.000,1.000)	**0.440 (0.332, 0.583)**
2.522 (1.598, 3.981)	2.274 (1.716, 3.013)	1.000 (1.000,1.000)

- Interpretation:
  - The odds of transitioning to the third class at time 4 relative to staying in the second class is 0.440 times lower for Pov=1 than for Pov=0 and significantly so (OR = 0.440 [0.332, 0.583])
- Given in the output under the heading COVARIATE EFFECTS ON TRANSITION PROBABILITY ODDS RATIOS

Outline

- An example
- Basic building blocks
- **Analysis without covariates**
  - Regular LTA
  - RI-LTA
  - Comparing results. Deciding on the number of classes
  - Checking model fit and model modifications
    - Checking response pattern fit and bivariate fit
    - Measurement invariance across time
    - Residual associations across time
    - Lag-2 modeling
- Adding covariates. Covariate effects on transition probabilities
  - RI-LTA. Comparing results with Regular LTA
- Measurement invariance across individuals
  - Multiple-group analysis and direct effects of covariates on indicators
- Special topics
  - Modeling stationarity of transition probabilities
  - Mover-Stayer modeling
  - Distal outcome

Analysis Without Covariates: Regular LTA

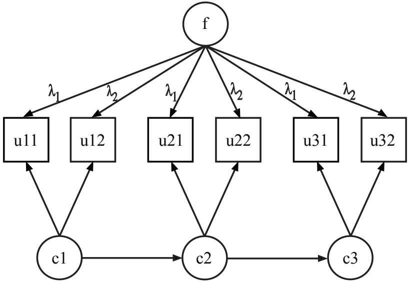
- Running Mplus:
  - Specify html output
  - Check replications of best loglikelihood (LL) and error messages
  - Note number of parameters, LL, BIC, class probabilities, entropy
  - Interpret measurement parameter results in probability scale
  - Automatic class reordering (introduced in version 8.5) using OPTSEED and new SVALUES feature. See the Mplus Version 8.5 Language Addendum at `http://www.statmodel.com/ugexcerpts.shtml`
  - Interpret class and transition probability results in odds and odds ratio scales
  - Check TECH10 and TECH15

Go to outputs 1 and 2 for Mplus Web Talk No. 2 at `www.statmodel.com`

Outline

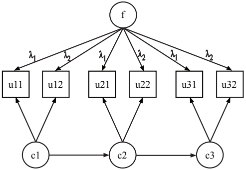
- An example
- Basic building blocks
- Analysis without covariates
  - Regular LTA
  - **RI-LTA**
  - Comparing results. Deciding on the number of classes
  - Checking model fit and model modifications
    - Checking response pattern fit and bivariate fit
    - Measurement invariance across time
    - Residual associations across time
    - Lag-2 modeling
- Adding covariates. Covariate effects on transition probabilities
  - RI-LTA. Comparing results with Regular LTA
- Measurement invariance across individuals
  - Multiple-group analysis and direct effects of covariates on indicators
- Special topics
  - Modeling stationarity of transition probabilities
  - Mover-Stayer modeling
  - Distal outcome

Random Intercept LTA (RI-LTA)



- Muthén & Asparouhov (2020). Forthcoming in Psychological Methods
- See also Mplus Web Talk No. 1
- Adding a random intercept factor f with equal loadings over time that captures time-invariant between-subject differences (trait differences; measurement non-invariance)

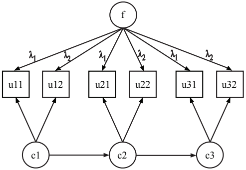
Random Intercept LTA (RI-LTA)



- Fits the data much better than regular LTA which is unnecessarily restrictive and gives distorted results
- Does not confound between- and within-subject sources of variation: $C_{t-1} -> C_t$ represents a within-subject process free of time-invariant between-subject differences (trait differences; meas. non-invariance)
- Latent class indicators correlate over time beyond what is captured by the latent class correlations over time:
  - Tends to reduce the probability of subjects staying in the same class as compared to regular LTA
  - Reduces the need for Mover-Stayer modeling because Movers and Stayers can be captured by different random intercept values

Key RI-LTA Features Running Mplus

- MODEL command for the random intercept factor f:




```
f BY u11-u12* (p1-p2)
     u21-u22* (p1-p2)
     u31-u32* (p1-p2);
f@1; [f@0];
```

- RI-LTA typically needs more random starts than regular LTA
- ML requires numerical integration over f which can be time-consuming (but only 1 dimension of integration):
  - ANALYSIS command: ALGORITHM = INTEGRATION;
  - Default number of integration points = 15
  - In some cases 30 points are needed to increase numerical precision and replicate best logL: INTEGRATION = 30;

Software Improvements for LTA and RI-LTA

RI-LTA can be time consuming due to numerical integration and needing many random starts to find the global maximum.
Mplus Version 8.4 released November 2019:

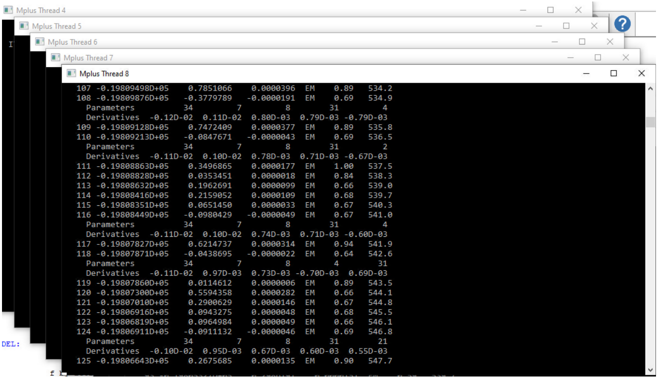
- Significant speed improvements for computationally demanding mixture models such as with LTA and RI-LTA using a new three-stage random starts search and using specialized algorithms drawing on Baum-Welch ideas
- Asparouhov & Muthén (2019). Random Starting Values and Multistage Optimization (Technical Report: `http://www.statmodel.com/download/StartsUpdate.pdf`)
    - ”A 20 hours computation in Mplus 8.3 can be done in Mplus 8.4 in less than 15 minutes, by utilizing the advantages of the three-stage estimation, the Baum-Welch algorithm, as well as updated hardware (i9-9900k Intel CPU)”
- Substantially simplified output for mixture models with multiple latent class variables

RI-LTA Speed Aspects

i7-7700k		i9-9900k					
8.3 Proc=8	8.4 Proc=8	8.3 Proc=8	8.3 Proc=10	8.3 Proc=12	8.4 Proc=8	8.4 Proc=10	8.4 Proc=12
Example 1							
23:12	12:55	13:21	13:16	12:21	7:26	7:36	6:41
Example 2							
6:08	00:28	3:37	5:09	05:43	00:18	00:17	00:17
Example 3							
1:30:08	10:36	1:01:46	47:34	45:26	6:03	5:36	5:47
Example 4							
12:30:04	14:36	8:47:42	7:36:54	7:15:02	8:11	7:44	7:40
Example 5							
20:25:17	27:22	15:32:38	12:17:19	12:17:12	15:01	14:51	14:47

- Example 1. This a two-level growth mixture model with 1 latent class variable with 4 categories, 4 observed continuous variables, 2 covariates, 2 dimensions of numerical integration, sample size=1500, and STARTS=250 100
- Example 2. This is an LTA model with 3 latent class variables with 5 categories each, 12 observed categorical variables, 0 dimensions of numerical integration, sample size=2933, and STARTS=100 20
- Example 3. This is a binary RI-LTA model with 4 latent class variables, one has 2 categories and the other three have 5 categories, 12 observed categorical variables, 0 dimensions of numerical integration, sample size=2933, and STARTS=320 80
- Example 4. This is a continuous RI-LTA model with 3 latent class variables with 5 categories each, 12 observed categorical variables, 1 dimension of numerical integration, sample size=2933, and STARTS=400 80
- Example 5. This is a continuous RI-LTA model with 3 latent class variables with 5 categories each, 9 observed categorical variables, 1 dimension of numerical integration, sample size=2933, and STARTS=320 80

Column 3 in TECH8 onscreen printing:



Negative ABS Changes: Technical Settings for Integration

- Better numerical precision can avoid negative abs changes but leads to slower calculations
  - Default number of integration points = 15
  - Change to INTEGRATION = 30
- Negative abs changes may appear for poor solutions and are of less concern if not present for the best logL solution
- Can be avoided also by ADAPTIVE = OFF to force monotonic logL improvement although some precision is lost

Sample Size (N) and Number of Timepoints (T)

- With categorical latent class indicators, large samples are recommended for LTA and RI-LTA: N in the order of a couple of thousand (Reading data has N=3574)
- With binary indicators, LTA - when it is the correct model - may be alright already at N = 500 but power low for finding effects on transitions
- RI-LTA performs well when $T \geq 3$ as long as $N \geq 500$
- With binary indicators and T=2, RI-LTA may need $N \geq 4000$
- With binary indicators, N=500, T=2, and where RI-LTA is the true model, RI-LTA has less bias than regular LTA
- With continuous indicators, RI-LTA with N=500, T=2 gets good results

Go to outputs 3 and 4 for Mplus Web Talk No. 2 at `www.statmodel.com`

Outline

- An example
- Basic building blocks
- Analysis without covariates
  - Regular LTA
  - RI-LTA
  - **Comparing results. Deciding on the number of classes**
  - Checking model fit and model modifications
    - Checking response pattern fit and bivariate fit
    - Measurement invariance across time
    - Residual associations across time
    - Lag-2 modeling
- Adding covariates. Covariate effects on transition probabilities
  - RI-LTA. Comparing results with Regular LTA
- Measurement invariance across individuals
  - Multiple-group analysis and direct effects of covariates on indicators
- Special topics
  - Modeling stationarity of transition probabilities
  - Mover-Stayer modeling
  - Distal outcome

Comparing Results
Regular LTA versus RI-LTA: Loglikelihood and BIC

Model	# parameters	log likelihood	BIC
Regular LTA	35	-21793	43873
RI-LTA	40	-20329	40984

- Likelihood-ratio chi-square difference testing not quite kosher due to the two models differing with respect to a factor (testing on the border of the parameter space). 2 times logL difference is not exactly chi-square distributed
- Chi-square test not needed given large logL and BIC differences
- BIC ok to use whenever the same DVs are considered

Reading Data Measurement Probability Estimates

Class 1 = low alphabet knowledge, Class 2 = early word reading, Class 3 = early reading comprehension

	Regular LTA			RI-LTA		
	Classes			Classes		
	1	2	3	1	2	3
Letrec	0.505	0.994	1.000	0.627	0.939	0.979
Begin	0.066	0.917	0.984	0.303	0.806	0.941
Ending	0.013	0.661	0.972	0.167	0.630	0.904
Sight	0.000	0.051	0.985	0.020	0.208	0.808
WIC	0.000	0.001	0.509	0.005	0.058	0.460

Reading Data Latent Class Probabilities

	Regular LTA			
	Fall K	Spring K	Fall 1st	Spring 1st
Class Probabilities				
1	0.694	0.235	0.142	0.041
2	0.284	0.635	0.627	0.154
3	0.023	0.130	0.232	0.805
	RI-LTA			
	Fall K	Spring K	Fall 1st	Spring 1st
Class Probabilities				
1	0.948	0.161	0.040	0.010
2	0.049	0.818	0.880	0.017
3	0.003	0.022	0.080	0.973

Reading Data Transition Probabilities

		Regular LTA		
		1	2	3
		Spring K		
Fall K	1	**0.338**	0.649	0.012
	2	0.001	**0.652**	0.348
	3	0.000	0.000	**1.000**
		RI-LTA		
		1	2	3
		Spring K		
Fall K	1	**0.170**	0.820	0.010
	2	0.000	**0.815**	0.185
	3	0.000	0.000	**1.000**

		Regular LTA		
		1	2	3
		Spring 1st		
Fall 1st	1	**0.263**	0.505	0.232
	2	0.005	**0.132**	0.863
	3	0.001	0.000	**0.999**
		RI-LTA		
		1	2	3
		Spring 1st		
Fall 1st	1	**0.155**	0.002	0.843
	2	0.004	**0.019**	0.977
	3	0.009	0.000	**0.991**

Reading Data Typical Transition Paths

- Estimated frequencies for the 3 most frequent paths through the latent class variables over the four time points (entries are latent classes):

Fall K	Spring K	Fall 1^{st}	Spring 1^{st}	Frequency	% of N
			Regular LTA		
1	2	2	3	1163	33
2	2	2	3	477	13
2	3	3	3	350	10
			RI-LTA		
1	2	2	3	2528	71
1	1	2	3	416	12
1	2	3	3	187	5

- Shown in TECH15 under the heading FINAL CLASS COUNTS AND PROPORTIONS FOR THE LATENT CLASSES BASED ON THE ESTIMATED MODEL

Deciding on the Number of Classes Using BIC
Two Approaches: One Timepoint at a Time or All Jointly

- Analysis without covariates

	Number of classes per time point			
	2	3	4	5
LCA Fall K	11709	**11476**	11523	
LCA Spring K	13740	**13034**	13075	
LCA Fall 1st	14030	**13091**	13131	
LCA Spring 1st	11464	**10939**	10960	
Regular LTA: All time points jointly	53649	43873	42017	**40694**
RI-LTA: All time points jointly	47742	40984	40511	**40328**

- Unlike regular LTA, a model for one timepoint at a time is not a subset of the RI-LTA model: LCA+1 factor is not a subset of the full model because it doesn't capture time invariance of the factor loadings
- BIC not showing a minimum is an indication that the model may need modification (see TECH10; try more flexible or alternative modeling)

Outline

- An example
- Basic building blocks
- Analysis without covariates
  - Regular LTA
  - RI-LTA
  - Comparing results. Deciding on the number of classes
  - **Checking model fit and model modifications**
    - Checking response pattern fit and bivariate fit
    - Measurement invariance across time
    - Residual associations across time
    - Lag-2 modeling
- Adding covariates. Covariate effects on transition probabilities
  - RI-LTA. Comparing results with Regular LTA
- Measurement invariance across individuals
  - Multiple-group analysis and direct effects of covariates on indicators
- Special topics
  - Modeling stationarity of transition probabilities
  - Mover-Stayer modeling
  - Distal outcome

Checking Model Fit and Model Modifications

- Frequency table test of model fit:
    - With categorical latent class indicators, the model can be tested against data using Pearson and likelihood-ratio chi-square frequency table tests. Summing over the cells of the table:

$$Pearson: \sum_j (o_j - e_j)^2 / e_j$$

$$Likelihood\ ratio: \ 2 \sum_j o_j \, log(o_j / e_j)$$

    - With LTA, there are typically too many frequency table cells with many cells having estimated frequencies close to zero, invalidating the tests:
        - THE CHI-SQUARE TEST CANNOT BE COMPUTED BECAUSE THE FREQUENCY TABLE FOR THE LATENT CLASS INDICATOR MODEL PART IS TOO LARGE.
        - Reading data has 5 binary indicators at 4 time points: $2^{20} = 1,048,576$ cells, where 604 cells have observations
- Alternative checks: Most frequent response patterns, bivariate tables; TECH10

Checking Fit of Response Patterns in TECH10

MOST FREQUENT RESPONSE PATTERNS AND CHI-SQUARE CONTRIBUTIONS

Response Pattern	Frequency Observed	Frequency Estimated	Standardized Residual (z-score)	Chi-square Pearson	Contribution Loglikelihood
			Regular LTA		
29	145.00	86.67	6.34	39.26	149.24
24	134.00	84.44	5.46	29.09	123.76
49	96.00	44.58	7.75	59.32	147.29
69	92.00	27.82	12.21	148.04	220.06
201	**91.00**	**5.74**	35.62	1266.63	**502.97**
61	81.00	6.66	28.82	829.11	404.62
44	73.00	20.28	11.74	137.00	186.97
67	69.00	10.28	18.34	335.39	262.73
28	69.00	81.01	-1.35	1.78	-22.15
1	67.00	22.89	9.25	85.01	143.92
			RI-LTA		
29	145.00	139.70	0.46	0.20	10.80
24	134.00	128.47	0.50	0.24	11.29
49	96.00	54.32	5.70	31.97	109.32
69	92.00	101.42	-0.95	0.88	-17.94
201	**91.00**	**30.84**	10.88	117.37	**196.94**
61	81.00	53.72	3.75	13.85	66.52
44	73.00	40.58	5.12	25.90	85.73
67	69.00	35.97	5.53	30.33	89.89
28	69.00	45.24	3.56	12.48	58.26
1	67.00	33.20	5.89	34.43	94.11

- Pattern 201 has all 0's for all items at all time points

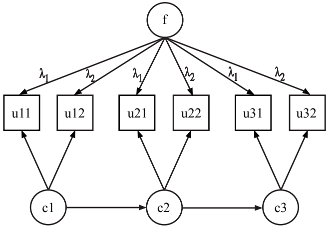
Checking Bivariate Fit in TECH10

- Results presented under the heading BIVARIATE MODEL FIT INFORMATION for each bivariate table
- E.g. LETREC1 BEGIN1: Bivariate Log-Likelihood Chi-Square 27.083

	Fall K		Spring 1st	
Variable pair	Regular LTA	RI-LTA	Regular LTA	RI-LTA
LETREC, BEGIN	27	6	30	72
ENDING	36	7	51	42
SIGHT	7	1	30	40
WIC	2	2	46	10
BEGIN, ENDING	48	37	14	14
SIGHT	7	4	4	14
WIC	4	7	37	6
ENDING, SIGHT	32	3	24	10
WIC	28	5	23	41
SIGHT, WIC	8	21	21	10

Go to output 2 for Mplus Web Talk 2

Measurement Invariance Across Time: Two Approaches (N=3574)



- Problematic: All indicators invariant across all time points with respect to measuring the latent classes for each class versus no such invariance:
    - Regular LTA: Chi-square (45) = 1850
    - RI-LTA: Chi-square (45) = 1340
    - Strong rejection but no information about which indicators contribute the most and class sizes and interpretation change
- Better: Allow non-invariance for one indicator at a time across all time points for each class. Speeded up by using SVALUES with STARTS=0

Measurement Invariance Across Time Allowing Non-Invariance for One Indicator at a Time: Chi-square Difference Testing (9 df), N=3574

Model	Letrec	Begin	Ending	Sight	WIC
Regular LTA	86	62	170	222	88
RI-LTA	88	88	178	322	82

Letrec: Basic reading skills of letter recognition
Begin: Beginning sounds
Ending: Ending sounds
Sight: Sight words (commonly used words that young children are encouraged to memorize as a whole by sight, so that they can automatically recognize these words in print without having to use any strategies to decode).
WIC: Words in context

- Let the Sight indicator be non-invariant and re-estimate for the other indicators

Input for Non-Invariance of the Sight Indicator

```
MODEL c1:
        %c1#1%
        [letrec1$1-wic1$1] (1-5);
        %c1#2%
        [letrec1$1-wic1$1] (6-10);
        %c1#3%
        [letrec1$1-wic1$1] (11-15);
MODEL c2:
        %c2#1%
        [letrec2$1-wic2$1] (1-5);
        [sight2$1]; ! over-rides the previous statement for sight2
        %c2#2%
        [letrec2$1-wic2$1] (6-10);
        [sight2$1];
        %c2#3%
        [letrec2$1-wic2$1] (11-15);
        [sight2$1];
```

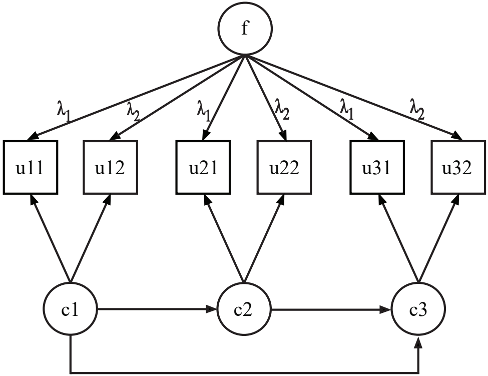
Etc for sight3 in MODEL c3 and sight4 in MODEL c4

Residual Correlation Over Time

- Asparouhov & Muthen (2015). Residual associations in latent class and latent transition analysis. Structural Equation Modeling
- Each indicator is allowed to have residual correlations across time
- Regular LTA, no covariates:
  - Zero residual correlations: 35 par's, logL = -21793, BIC = 43873
  - Residual correlations: 50 par's, logL = -20699, BIC = 41808
- Not available with RI-LTA (40 par's, logL = -20329, BIC = 40984)

```
ANALYSIS:      TYPE = MIXTURE;
               STARTS = 100 20;
               PROCESSORS = 8;
               PARAMETERIZATION = RESCOV;
MODEL:         %OVERALL%
               c2 ON c1;
               c3 ON c2;
               c4 ON c3;
               letrec1 WITH letrec2; letrec2 WITH letrec3; letrec3 WITH letrec4;
               begin1 WITH begin2; begin2 WITH begin3; begin3 WITH begin4;
               ending1 WITH ending2; ending2 WITH ending3; ending3 WITH ending4;
               sight1 WITH sight2; sight2 WITH sight3; sight3 WITH sight4;
               wic1 WITH wic2; wic2 WITH wic3; wic3 WITH wic4
```

Lag-2 Modeling: C3 ON C1



Lag-2 Model Test Results

Model	# par's	LL	BIC	Test	Chi-square[1]	df
Regular LTA						
1. Lag-1	35	-21793	43873			
2. Lag-2	43	-21655	**43661**	1 vs **2**	276	8
RI-LTA						
3. Lag-1	40	-20329	**40984**			
4. Lag-2	48	-20319	41030	3 vs **4**	20	8

- Lag-2 model needed for both regular LTA and RI-LTA
  - - but there is a much stronger need for lag-2 with regular LTA
- What do the lag-2 effects look like? Look at TECH15
- Does lag-2 modeling result in a different decision on the number of classes? Lag-3?

Lag-2 Effects in Probability Terms Combining Results from TECH15

- Effect of C1 on C3, marginalized over C2:

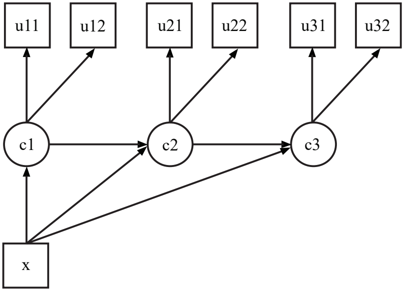
$$
\begin{aligned}
P(C3 = a|C1 = b) = & \\
& = P(C3 = a|C1 = b, C2 = 1) \times P(C2 = 1|C1 = b) \\
& + P(C3 = a|C1 = b, C2 = 2) \times P(C2 = 2|C1 = b) \\
& + P(C3 = a|C1 = b, C2 = 3) \times P(C2 = 3|C1 = b)
\end{aligned}
$$

- Note that in lag-1 models:
  - $P(C3 = a|C1 = b, C2 = c)$ is independent of b
  - But $P(C3 = a|C1 = b)$ is not independent of b
- That is, such a table could be useful in lag 1 models as well answering the question: How does starting point affect ending point?

Outline

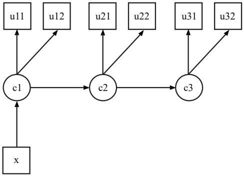
- An example
- Basic building blocks
- Analysis without covariates
  - Regular LTA
  - RI-LTA
  - Comparing results. Deciding on the number of classes
  - Checking model fit and model modifications
    - Checking response pattern fit and bivariate fit
    - Measurement invariance across time
    - Residual associations across time
    - Lag-2 modeling
- **Adding covariates. Covariate effects on transition probabilities**
  - RI-LTA. Comparing results with Regular LTA
- Measurement invariance across individuals
  - Multiple-group analysis and direct effects of covariates on indicators
- Special topics
  - Modeling stationarity of transition probabilities
  - Mover-Stayer modeling
  - Distal outcome

Latent Class Variables Influenced by Covariate



- Three types of models:
  - C1 ON X
  - C1-C3 ON X (main effect model)
  - C1 ON X, C1-specific C2 ON X, C2-specific C3 ON X (interaction model)

C1 ON X: Covariate Influencing the Starting Point of the Transitions



- In this model, transition probabilities do not change as a function of X, only C1, but C2 and C3 are influenced indirectly
- Good model to use as a baseline model to test the need for effects on transition probabilities
- C1 ON X is a multinomial regression where the last class (R) is the reference class :
  - $P(C1 = r|X) = e^{a_r + b_r X} / [e^{a_1 + b_1 X} + e^{a_2 + b_2 X} + \ldots + e^{a_{R-1} + b_{R-1} X} + 1]$
  - Odds ratio: e^{b_r} represents the ratio of the odds $P(C1 = r|X)/P(C1 = R|X)$ for X=x divided by the odds for X = x-1 (a one-unit change), e.g., X=1 versus X=0 for a binary X

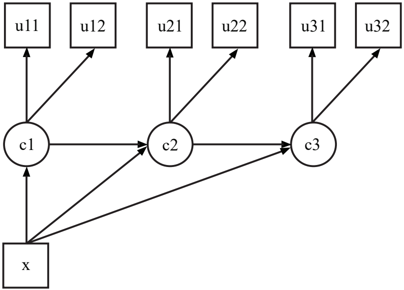
Logistic Regression Odds Ratio Results for C1

- $P(C1 = r|X) = e^{a_r + b_r X} / [e^{a_1 + b_1 X} + e^{a_2 + b_2 X} + \ldots + e^{a_{R-1} + b_{R-1} X} + 1]$
- Odds ratio: e^{b_r} represents the ratio of the odds $P(C1 = r|X)/P(C1 = R|X)$ for X=x divided by the odds for X = x-1 (a one-unit change), e.g., X=1 versus X=0 for a binary X

			95% C.I.	
	Estimate	S.E.	Lower 2.5%	Upper 2.5%
Categorical Latent Variables				
C1#1 ON				
POV	8.101	4.796	2.539	25.850
C1#2 ON				
POV	1.881	1.143	0.571	6.189

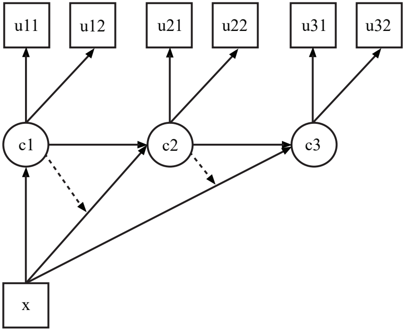
- Results given in the output under the heading LOGISTIC REGRESSION ODDS RATIO RESULTS

Transition Probabilities Influenced by Covariate: Main Effect Model: $C_1 - C_3$ ON X



C1-C3 ON X;
C2 ON C1;
C3 ON C2;

Transition Probabilities Influenced by Covariate: Interaction Effect Model: $C_1 - C_3$ ON X, C_1-Specific C_2 ON X, and C_2-Specific C_3 ON X



User's Guide ex 8.14 (two time points)

Logits and Transition Probabilities Influenced by Covariate

Main effect model: C_2 ON X

		C_2 1	2	3
C_1	1	$a_1 + b_{11} + g_1 x$	$a_2 + b_{21} + g_2 x$	0
	2	$a_1 + b_{12} + g_1 x$	$a_2 + b_{22} + g_2 x$	0
	3	$a_1 + g_1 x$	$a_2 + g_2 x$	0

Interaction model: C_1-specific C_2 ON X

		C_2 1	2	3
C_1	1	$a_1 + b_{11} + g_{11} x$	$a_2 + b_{21} + g_{21} x$	0
	2	$a_1 + b_{12} + g_{12} x$	$a_2 + b_{22} + g_{22} x$	0
	3	$a_1 + g_{13} x$	$a_2 + g_{23} x$	0

Effects of a Covariate on Transitions: Odds Ratios

Interaction model: C_1-specific C_2 ON X

		C_2 1	2	3
C_1	1	$a_1 + b_{11} + g_{11}\mathrm{x}$	$a_2 + b_{21} + g_{21}\mathrm{x}$	0
	2	$a_1 + b_{12} + g_{12}\mathrm{x}$	$a_2 + b_{22} + g_{22}\mathrm{x}$	0
	3	$a_1 + g_{13}\mathrm{x}$	$a_2 + g_{23}\mathrm{x}$	0

- Each row represents a regression with a nominal DV. The Mplus modeling uses the last class as reference class. Take, for example, the interpretation of g_{21}:
  - Consider row 1: $a_2 + b_{21} + g_{21}\mathrm{x}$ is the log of the odds of $P(C2 = 2|C1 = 1, X)/P(C2 = 3|C1 = 1, X)$
  - Consider the ratio of this odds for X=1 and X=0. The log of this odds ratio is g_{21} (because the $a_2 + b_{21}$ terms cancel out)
  - This means that $\exp(g_{21})$ is an odds ratio that describes the effect of X on the odds of being in C2=2 versus C2=3 for C1=1
  - In the main effect model, this odds ratio effect of X is the same for all rows, i.e., all classes of C1. The OR is shown under the heading LOGISTIC REGRESSION ODDS RATIO RESULTS

Effects of a Covariate on Transitions: Odds Ratios Cont'd

Interaction model: C_1-specific C_2 ON X

		C_2 1	2	3
C_1	1	$a_1 + b_{11} + g_{11}\mathrm{x}$	$a_2 + b_{21} + g_{21}\mathrm{x}$	0
	2	$a_1 + b_{12} + g_{12}\mathrm{x}$	$a_2 + b_{22} + g_{22}\mathrm{x}$	0
	3	$a_1 + g_{13}\mathrm{x}$	$a_2 + g_{23}\mathrm{x}$	0

$$e^{g_{21}} = odds_{x=1}/odds_{x=0} = \frac{P(C2=2|C1=1,X=1)/P(C2=3|C1=1,X=1)}{P(C2=2|C1=1,X=0)/P(C2=3|C1=1,X=0)}$$

Instead of using the last class as reference, C2=3, transitions are often more easily understood by using the diagonal, stayer class as reference, C2=1 :

$$e^{g_{21}} = odds_{x=1}/odds_{x=0} = \frac{P(C2=2|C1=1,X=1)/P(C2=1|C1=1,X=1)}{P(C2=2|C1=1,X=0)/P(C2=1|C1=1,X=0)}$$

Using Diagonal Class as Reference Class

- Instead of using the last class as reference, transitions are often more easily understood by using the diagonal, stayer class as reference
- Logit pattern for transition table:

		C_2		
		1	2	3
	1	0		
C_1	2		0	
	3			0

- The reference class for each row has odds ratio 1 because both the numerator and denominator odds are 1. E.g. for row 1:

$$Odds_{x=1}/odds_{x=0} = \frac{P(C2=1|C1=1,X=1)/P(C2=1|C1=1,X=1)}{P(C2=1|C1=1,X=0)/P(C2=1|C1=1,X=0)}$$

Effects of a Covariate on Transitions: Odds Ratios Cont'd

- Example: Pov=1 compared to Pov=0 for regular LTA with C1-C4 ON Pov:

Transition table odds ratio and 95% confidence intervals for C1 to C2

1.000 (1.000,1.000)	**0.305** (0.248, 0.376)	0.459 (0.267, 0.790)
3.273 (2.659, 4.032)	1.000 (1.000,1.000)	1.503 (0.914, 2.473)
2.178 (1.267, 3.7344)	0.665 (0.404, 1.094)	1.000 (1.000,1.000)

- Interpretation:
  - The odds of transitioning to the second class relative to staying in the first class is 0.305 times lower for Pov=1 than for Pov=0 and significantly so (OR = 0.305 [0.248, 0.376])
- Given in the output under the heading COVARIATE EFFECTS ON TRANSITION PROBABILITY ODDS RATIOS

Go to outputs 5 - 7 for Mplus Web Talk No. 2

Model Tests with Covariates: Regular LTA (N=3574)

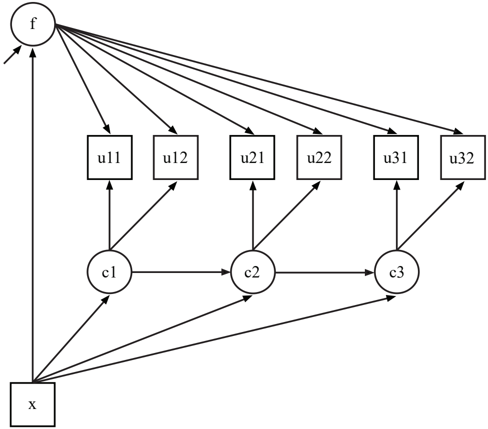
Model	# par's	LL	BIC	Test	Chi-2[1]	df
1. No Pov covariate	35	-21793	43873			
2. Baseline:						
C1 ON Pov	37	-21708	43718	1 vs **2**	310	2
Effects on transitions:						
Main effects:						
3. C1-C4 ON Pov	43	-21584	**43519**	2 vs **3**	248	6
Interaction Effects:						
4.	55	-21579	43608	**3** vs 5	10	12

Scaling correction factors not applied (close to 1).
For scaling correction, see chi-2 diff testing using the loglikelihood at `http://www.statmodel.com/chidiff.shtml`

Class Percentages: Regular LTA

		Time Point			
Model	Class	Fall K	Spring K	Fall 1st	Spring 1st
No covariate	1	69	23	14	5
	2	28	64	63	15
	3	2	13	23	81
Adding Poverty	1	69	23	14	4
as a covariate	2	29	64	63	15
(Model 3)	3	2	13	23	81

Transition Probabilities Influenced by Covariate: RI-LTA



Go to outputs 8 and 9 for Mplus Web Talk No. 2 at
`www.statmodel.com`

Model Tests with Covariates: RI-LTA (N=3574)

Model	# par's	LL	BIC	Test	Chi-2[1]	df
1. No Pov	40	-20329	40984			
2. f ON Pov	41	-20128	**40591**	1 vs **2**	402	1
3. C1 ON Pov	42	-20319	40981	1 vs **3**	20	2
4. C1 ON Pov, f ON Pov	43	-20127	40604	**2** vs 4	2	2
Effects on transitions:						
Main effects:						
5. C2-C4 ON Pov, f ON Pov	47	-20106	40597	2 vs **5**	44	6
6. C1-C4 ON Pov, f ON Pov (output 8)	49	-20104	40608	**5** vs 6	4	2
Interaction effects:						
7. (output 9)	61	-20098	40695	**5** vs 7	16	14

[1] Scaling correction factors not applied (close to 1).

Best BIC with regular LTA: 43519

Class Percentages: RI-LTA

		Time Point			
Model	Class	Fall K	Spring K	Fall 1st	Spring 1st
No covariate	1	95	16	4	1
	2	5	82	88	2
	3	0	2	8	97
Adding Poverty as a covariate (Model 6)	1	94	17	5	2
	2	5	80	86	2
	3	1	3	9	96

Class Percentages without and with Poverty Covariate

Model	Class	Fall K	Spring K	Fall 1st	Spring 1st
		Time Point			
Regular LTA	1	69	23	14	5
No covariate	2	28	64	63	15
	3	2	13	23	81
Adding Poverty	1	69	23	14	4
as a covariate	2	29	64	63	15
(Model 3)	3	2	13	23	81
RI-LTA	1	95	16	4	1
No covariate	2	5	82	88	2
	3	0	2	8	97
Adding Poverty	1	94	17	5	2
as a covariate	2	5	80	86	2
(Model 6)	3	1	3	9	96

Transition Probabilities from LTA Calculator: Regular LTA

- Using the Calculator to look at more details of Model 3, run 6:

$C3 \rightarrow C4$	Pov=1 Fall $1^{st} \rightarrow$ Spring 1^{st}			Pov=0 Fall $1^{st} \rightarrow$ Spring 1^{st}		
	1	2	3	1	2	3
1	0.301	0.542	0.157	0.232	0.464	0.305
2	0.010	**0.226**	**0.764**	0.004	**0.115**	**0.881**
3	0.001	0.000	0.999	0.000	0.000	1.000

- OR for C4=2 | C3=2: $\frac{0.226/0.764}{0.115/0.881} = 2.27$. Log(2.27) = 0.82 = $\hat{g}_2$, where g_2 is C4#2 ON POV
- Interpretation: The odds of staying in class 2 versus moving to class 3 is 2.27 times larger for Pov=1 than Pov=0
- Instead using the diagonal as reference class: $\frac{0.764/0.226}{0.881/0.115} = 0.44$
  - This is the OR for moving to class 3 versus staying in class 2
  - The ORs and their CIs are given under the heading COVARIATE EFFECTS ON TRANSITION PROBABILITY ODDS RATIOS
- 5 of 7 coefficients for C2-C4 ON Pov are significant

Transition Probabilities from LTA Calculator: RI-LTA

- Using the Calculator to look at more details of Model 6, run 8

$C3 \rightarrow C4$	Pov=1 Fall $1^{st} \rightarrow$ Spring 1^{st}			Pov=0 Fall $1^{st} \rightarrow$ Spring 1^{st}		
	1	2	3	1	2	3
1	0.288	0.059	0.653	0.256	0.014	0.730
2	0.004	**0.052**	**0.944**	0.004	**0.011**	**0.985**
3	0.010	0.000	0.990	0.008	0.000	0.992

- Only 2 of 7 C2-C4 ON Pov coefficients are significant
- Most of the effect of Pov is on the random intercept factor f
- This indicates measurement non-invariance with respect to poverty (time-invariant measurement non-invariance)

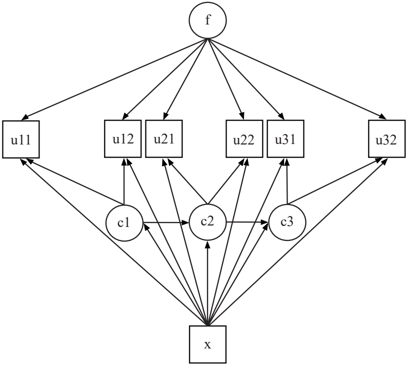
Outline

- An example
- Basic building blocks
- Analysis without covariates
    - Regular LTA
    - RI-LTA
    - Comparing results. Deciding on the number of classes
    - Checking model fit and model modifications
        - Checking response pattern fit and bivariate fit
        - Measurement invariance across time
        - Residual associations across time
        - Lag-2 modeling
- Adding covariates. Covariate effects on transition probabilities
    - RI-LTA. Comparing results with Regular LTA
- **Measurement invariance across individuals**
    - Multiple-group analysis and direct effects of covariates on indicators
- Special topics
    - Modeling stationarity of transition probabilities
    - Mover-Stayer modeling
    - Distal outcome

Allowing Measurement Non-Invariance Across Individuals

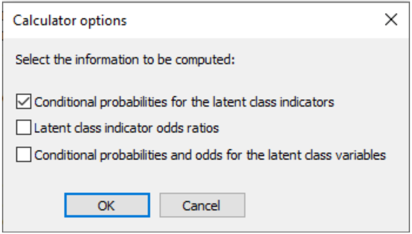
- Multiple-group approach using the Knownclass option
  - Adds one more class which makes computations slower
  - Measurement probabilities given for each Knownclass under the heading RESULTS IN PROBABILITY SCALE
- Direct effects of covariates on latent class indicators
  - Faster than using Knownclass (in this example by a factor of about 2 for regular LTA and by a factor > 10 for RI-LTA)
  - Significance of non-invariance for each covariate-indicator pair given directly through the estimated direct effects
  - The Calculator can give measurement probabilities for any combination of covariate values

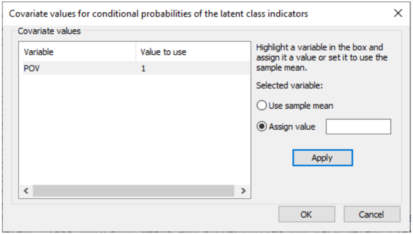
Direct Effects of Covariates on Latent Class Indicators for the RI-LTA Model (Regular LTA Excludes the f Factor)



- Note that f ON x is not needed/should not be included because it is not identified when all direct effects are present

Direct Effect Approach Using the Calculator





- Go to outputs 10-13

Measurement Invariance Testing for Poverty (N = 3574)

Model	# par's	LL	BIC	Test	Chi-2[1]	df
		Regular LTA, C1-C4 ON POV				
1. Multiple-group (non-inv),						
Pov as Knownclass	59	-23176	46834			
Adjusted[2]	58	-21428	43330			
Pov as covariate:						
2. Direct effects (non-inv)	58	-21428	**43330**			
3. No direct effects (inv)	43	-21584	43519	3 vs **2**	312	15
		RI-LTA, C1-C4 ON Pov, f not ON Pov/f ON Pov[3]				
2. Direct effects	63	-20090	40695			
3. No direct effects[3]	49	-20104	**40608**	3 vs 2 (?)	28	14

[1] Two times logL difference. Scaling correction factors not applied (close to 1).
[2] Adjustment for Knownclass part of LL: $N[p\ log\ p + (1-p)\ log\ (1-p)] = -1748$.
[3] Model 3 is the same as Model 6 on slide 71.

- In RI-LTA, the random intercept factor f takes care of measurement non-invariance that is time-invariant; direct effects not needed

Measurement Non-Invariance Using Multiple-Group (Knownclass)

Regular LTA

	Pov=1: Poverty (19%)			Pov=0: Non-poverty (81%)		
Letrec	0.300	0.987	1.000	0.593	0.995	1.000
Begin	0.023	0.807	0.971	0.117	0.930	0.986
Ending	0.007	0.496	0.971	0.021	0.698	0.973
Sight	0.000	0.023	0.893	0.000	0.063	0.986
WIC	0.001	0.001	0.310	0.000	0.001	0.537

RI-LTA

	Pov=1: Poverty (19%)			Pov=0: Non-poverty (81%)		
Letrec	0.371	0.880	0.977	0.692	0.969	0.990
Begin	0.116	0.641	0.911	0.343	0.858	0.961
Ending	0.052	0.421	0.840	0.187	0.687	0.932
Sight	0.003	0.082	0.653	0.019	0.230	0.857
WIC	0.000	0.011	0.237	0.003	0.062	0.512

Meas. Non-Inv. and Covariate Effects on Latent Classes

Regular LTA, no poverty covariate

	1	2	3
C1 class sizes:	0.694	0.284	0.023
C4 class sizes:	0.041	0.154	0.805

Regular LTA, poverty covariate, measurement invariance

C1 class sizes:	0.692	0.286	0.023
C4 class sizes:	0.041	0.154	0.805
$C1^{\#1}$ ON POV:	2.092	(3.534)	
$C1^{\#2}$ ON POV:	0.632	(1.039)	

Regular LTA, poverty covariate, measurement non-invariance (all direct effects)

C1 class sizes:	0.585	0.184	0.231
C4 class sizes:	0.069	0.711	0.220
$C1^{\#1}$ ON POV:	2.436	(4.163)	
$C1^{\#2}$ ON POV:	7.504	(12.949)	

- RI-LTA with measurement invariance has insignificant effects of Pov on C1

1-Step and Multi-Step Analysis

1-step, Regular LTA, poverty covariate
measurement non-invariance (all direct effects)

	1	2	3
C1 class sizes:	0.585	0.184	0.231
C4 class sizes:	0.069	0.711	0.220
$C1^{\#1}$ ON POV:	2.436	(4.163)	
$C1^{\#2}$ ON POV:	7.504	(12.949)	

2-step, fixed measurement model (meas. inv.) [1]

C1 class sizes:	0.693	0.284	0.023
C4 class sizes:	0.041	0.154	0.805
$C1^{\#1}$ ON POV:	2.090	(3.534)	
$C1^{\#2}$ ON POV:	0.626	(1.031)	

3-step BCH, Mplus Web Note 21 (meas. inv.)

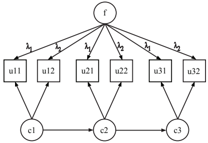
C1 class sizes:	0.688	0.294	0.018
C4 class sizes:	0.041	0.153	0.807
$C1^{\#1}$ ON POV:	3.966	(1.198)	
$C1^{\#2}$ ON POV:	0.824	(0.232)	

[1] Bakk & Kuha, 2018, Psychometrika

Measurement Non-Invariance Issues

- Direct effect modeling is important to explore the need to allow for measurement non-invariance
  - Including all direct effects can be difficult with many covariates due to needing many starts. Alternatives:
    - Regressing one latent class indicator at a time on all covariates
    - Using a few key candidate covariates as with gender in the Dating example of the Muthén-Asparouhov (2021) Psych Meth paper
- Multi-step methods to study the effects of covariates on the latent class variables (C ON X) are risky when there is measurement non-invariance
  - Mplus Web Note 15 (Asparouhov-Muthén, 2014 in Structural Equation Modeling) shows a simulation study where bias for C ON X is found with low entropy and many direct effects
  - Also pointed out in Janssen et al. (2019). The detection and modeling of direct effects in latent class analysis. Structural Equation Modeling
  - What to conclude?
    - When you really need multi-step, it doesn't work
    - When it works, you don't really need it

The Meaning of the Random Intercept Factor: The Measurement Non-Invariance Perspective



- For binary latent class indicator U_r, individual i, at time t, and class c,

$$logit[P(U_{rit} = 1|C_{it} = c, f_i) = -\tau_{rc} + \lambda_r f_i, \quad (1)$$

  for threshold τ and random intercept factor loading λ. A high logit value implies a high probability of $U = 1$, e.g. $\lambda > 0, f_i > 0$.

- In the presentation of measurement probabilities, the random intercept factor f was integrated out, that is, considering the overall conditional probability $P(U_{rit} = 1|C_{it} = c)$
- Individuals with different f_i values have different conditional probabilities $P(U_{rit} = 1|C_{it} = c, f_i)$ as shown in (1)
- This means that the random intercept factor f can be seen as capturing measurement non-invariance

Explicating Implicit RI-LTA Measurement Non-Invariance

RI-LTA at different random intercept factor values (no Poverty covariate)						
	$f = -0.5$			$f = +0.5$		
Letrec	0.384	0.987	0.998	0.954	1.000	1.000
Begin	0.045	0.809	0.977	0.425	0.985	0.999
Ending	0.013	0.439	0.942	0.151	0.911	0.995
Sight	0.000	0.005	0.850	0.001	0.185	0.996
WIC	0.000	0.000	0.091	0.001	0.010	0.811
RI-LTA, allowing measurement non-invariance across poverty groups						
	Poverty (19%)			Non-poverty (81%)		
Letrec	0.371	0.880	0.977	0.692	0.969	0.990
Begin	0.116	0.641	0.911	0.343	0.858	0.961
Ending	0.052	0.421	0.840	0.187	0.687	0.932
Sight	0.003	0.082	0.653	0.019	0.230	0.857
WIC	0.000	0.011	0.237	0.003	0.062	0.512

- The random intercept factor f
  - - captures much of the poverty non-invariance (time-invariant part)
  - - is strongly related to poverty (f ON poverty)
  - - can perhaps be viewed as a reading preparedness dimension

Outline

- An example
- Basic building blocks
- Analysis without covariates
  - Regular LTA
  - RI-LTA
  - Comparing results. Deciding on the number of classes
  - Checking model fit and model modifications
    - Checking response pattern fit and bivariate fit
    - Measurement invariance across time
    - Residual associations across time
    - Lag-2 modeling
- Adding covariates. Covariate effects on transition probabilities
  - RI-LTA. Comparing results with Regular LTA
- Measurement invariance across individuals
  - Multiple-group analysis and direct effects of covariates on indicators
- **Special topics**
  - Modeling stationarity of transition probabilities
  - Mover-Stayer modeling
  - Distal outcome

Special Topics: Modeling Stationarity of Transition Probabilities

		C_2		
		1	2	3
	1	$a_1 + b_{11}$	$a_2 + b_{21}$	0
C_1	2	$a_1 + b_{12}$	$a_2 + b_{22}$	0
	3	a_1	a_2	0

		C_3		
		1	2	3
	1	$a_1 + b_{11}$	$a_2 + b_{21}$	0
C_2	2	$a_1 + b_{12}$	$a_2 + b_{22}$	0
	3	a_1	a_2	0

MODEL:

```
%OVERALL%
[c2#1 c3#1] (a1);
[c2#2 c3#2] (a2);
c2 ON c1 (b1-b4);
c3 ON c2 (b1-b4);
```

If x covariates are involved, their effects have to be held equal across time.

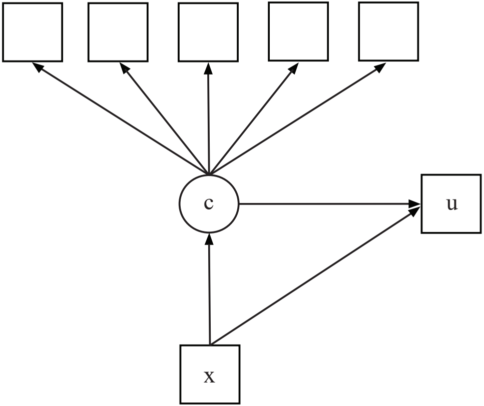
Mover-Stayer Modeling

- Adding a latent class of Stayers who have zero probability of changing class; probability=1 on the diagonal of the transition table. Discussed further in Mplus Web Talk No. 1
- Regular LTA, no covariates:
  - No Mover-Stayer component: 35 par's, logL = -21793, BIC = 43873
  - Mover-Stayer: 38 par's, logL = -21791, BIC = 43892 (6 % Stayers)
- RI-LTA, no covariates:
  - No Mover-Stayer component: 40 par's, logL = -20329, BIC = 40984
  - Mover-Stayer: zero percent in Stayer class (M-S not needed). Stayers captured by large loadings on the random intercept factor

```
ANALYSIS:   TYPE = MIXTURE;
            STARTS = 200 40;
            PROCESSORS = 8;
            PARAMETERIZATION = PROBABILITY;
MODEL:      %OVERALL%
            c1 ON c;
MODEL c:    %c#1%
            ! Mover class (regular class)
            c2 ON c1;
            c3 ON c2;
            c4 ON c3;
```

```
%c#2%
! Stayer class
c2#1 ON c1#1@1; c2#2 ON c1#1@0;
c2#1 ON c1#2@0; c2#2 ON c1#2@1;
c3#1 ON c2#1@1; c3#2 ON c2#1@0;
c3#1 ON c2#2@0; c3#2 ON c2#2@1;
c4#1 ON c3#1@1; c4#2 ON c3#1@0;
c4#1 ON c3#2 @0; c4#2 ON c3#2@1;
```

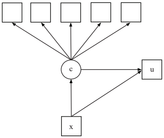
Distal Outcome in Latent Class Analysis



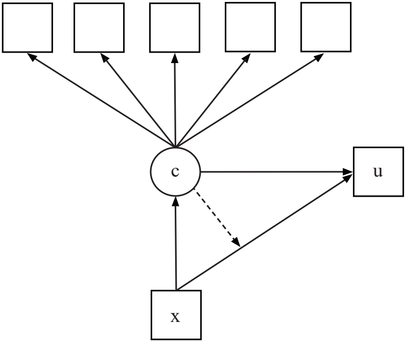
- See also slides 17-20

Distal Outcome Continued

- Odds and odds ratios with two predictors:
  - General case - logit regression (logit = logodds):
  - $Logodds(U|X_1, X_2) = a + b_1\, X_1 + b_2\, X_2$
  - Odds ratios interpretation:
    - OR effect of X_1 is e^{b_1} irrespective of the value of X_2
    - OR effect of X_2 is e^{b_2} irrespective of the value of X_1
  - Latent class model (main effect model):



  - Logit regression with X and a latent class predictor:
    - $logodds(U|X, C = c) = a_c + b\, X = -t_c + b\, X$ for latent classes $c = 1, 2 \ldots C$
    - OR effect of X is e^b irrespective of the latent class
    - OR effect of latent class is expressed in terms of the U odds for $c = 1$ divided by the U odds for $c = C$ and is not influenced by the value of X



- Logit regression:
  - $logodds(U) = a_c + b_c\,X = -t_c + b_c\,X$ for $c = 1, 2 \ldots C$

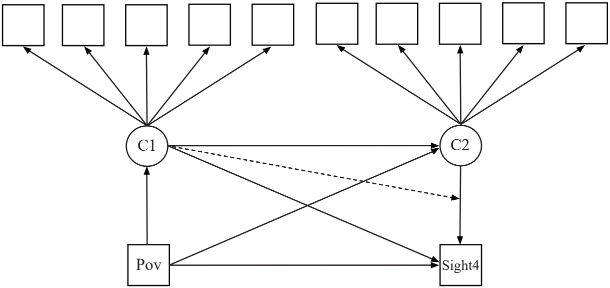
Reading Example: Distal Outcome Prediction in LTA

Class 1 = low alphabet knowledge, Class 2 = early word reading, Class 3 = early reading comprehension (regular LTA estimates)

	Classes		
	1	2	3
Letrec	0.505	0.994	1.000
Begin	0.066	0.917	0.984
Ending	0.013	0.660	0.972
Sight	0.000	0.051	**0.985**
WIC	0.000	0.000	0.509

- Sight (Sight words) is a key indicator of belonging to class 3
- What can we say about the likelihood of mastering Sight at Spring 1st grade based on development from K Fall to K Spring?

Distal Outcome Sight at Time 4 Predicted by Pov and LTA Transition Paths from Time 1 to Time 2



- Transitions follow the main effect model
- Distal outcome influenced by Pov and C1, C2 interaction effects, that is, transition paths

Distal Outcome Predicted by Transition Paths

Latent Class				Prob (Sight4 = 1) Spring 1st	
Fall K	Spring K	Freq	%	Pov = 1	Pov = 0
1	1	752	21%	0.338	0.498
1	2	1573	44%	0.773	0.869
2	2	709	20%	0.944	0.970

- The modeling uses the dot approach which refers to a combination of latent classes
- A check confirms that the latent class percentages and the latent class regressions on Pov are unchanged by the addition of Sight4
- The probabilities for Sight4 can be calculated via thresholds and the slope in the regression of Sight4 on Pov using the Model Constraint command as shown on the next slides (not available via the Calculator for items influenced by more than one latent class variable)

Distal Outcome Predicted by Transition Paths: Regular LTA, Main Effect Model

MODEL:

```
%OVERALL%
c2 ON c1;
c1 c2 ON pov;
sight4 ON pov (b);

%c1#1.c2#1%
[sight4$1] (t11);
[letrec1$1-wic1$1] (1-5) ;
[letrec2$1-wic2$1] (1-5) ;

%c1#1.c2#2%
[sight4$1] (t12);
[letrec1$1-wic1$1] (1-5) ;
[letrec2$1-wic2$1] (6-10);

%c1#1.c2#3%
[sight4$1] (t13);
[letrec1$1-wic1$1] (1-5) ;
[letrec2$1-wic2$1] (11-15);

%c1#2.c2#1%
[sight4$1] (t21);
[letrec1$1-wic1$1] (6-10);
[letrec2$1-wic2$1] (1-5) ;

%c1#2.c2#2%
[sight4$1] (t22);
[letrec1$1-wic1$1] (6-10);
[letrec2$1-wic2$1] (6-10);

%c1#2.c2#3%
[sight4$1] (t23);
[letrec1$1-wic1$1] (6-10);
[letrec2$1-wic2$1] (11-15);

%c1#3.c2#1%
[sight4$1] (t31);
[letrec1$1-wic1$1] (11-15);
[letrec2$1-wic2$1] (1-5) ;

%c1#3.c2#2%
[sight4$1] (t32);
[letrec1$1-wic1$1] (11-15);
[letrec2$1-wic2$1] (6-10);

%c1#3.c2#3%
[sight4$1] (t33);
[letrec1$1-wic1$1] (11-15);
[letrec2$1-wic2$1] (11-15);
```

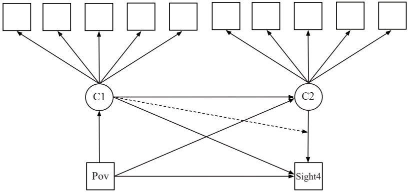
Model Constraint Statements: Excerpts

```
! Pov = 1:
! Prob = 1/ exp(-logit), logit = -threshold + b Pov
p111 = 1/(1+exp(t11-b));
p112 = 1/(1+exp(t12-b));
p113 = 1/(1+exp(t13-b));
p121 = 1/(1+exp(t21-b));
p122 = 1/(1+exp(t22-b));
p123 = 1/(1+exp(t23-b));
p131 = 1/(1+exp(t31-b));
p132 = 1/(1+exp(t32-b));
p133 = 1/(1+exp(t33-b));
```

ORs for Effects of Transitions on Sight4

- OR's and their CIs for effects of transitions on Sight4 can be computed in line with slides 19-20 but are given in TECH15 under the heading
  Latent class indicator odds ratios for the latent classes
- ORs for different transition examples:
  - 1, 1 (21 %) versus 1,2 (44 %) has OR = 0.150 [0.118, 0.190]
  - 1,2 versus 1, 1 not printed but obtained as the inverse of the above estimate and CI limits (switched), that is, OR = 6.67 [5.26, 8.47]
    - If you progress from class 1 to class 2 by end of Kindergarten, your odds of mastering Sight words at the end of First grade is almost 7 times higher than if you are still in class 1 by end of Kindergarten
  - 2, 2 (20 %) versus 1, 1 has OR = 33.33 [19.23. 55.55]
  - 2, 2 versus 1, 2 has OR = 4.90 [2.88, 8.40]
- OR for the effect of Pov on Sight4 is 0.514 [0.409, 0.645] given under the heading
  Logistic regression odds ratio results

Does the Fall Kindergarten Status Matter?



- Effects of latent classes on Sight4:

Model	# parameters	log likelihood	BIC
1. Transition (C1, C2 interaction)	37	-12644	25592
2 . C2 only	31	-12673	25599
3. C1 and C2 main effect	33	-12645	**25559**

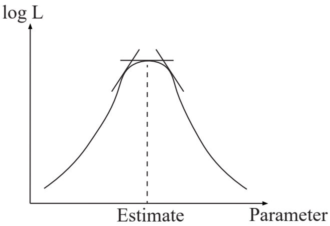
- Testing model 2 against model 1: Chi-square (6) = 56 $\rightarrow$ model 1
- Testing model 3 against model 1: Chi-square (4) = 2 $\rightarrow$ model 3
- It matters where you start in Fall of Kindergarten

- RI-LTA version is possible
- To ensure that latent class membership doesn't change, for instance with non-normal continuous distal outcomes, distal outcome analysis can be handled by multi-step methods such as BCH; see Mplus Web Note 21
- Go to full outputs 14, 15, 16 (models 1, 2, 3)
- RI-LTA model 3: Best BIC. Output 17

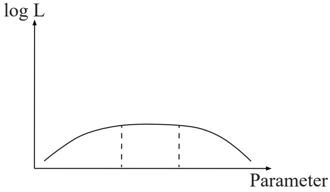
Outline

- An example
- Basic building blocks
- Analysis without covariates
  - Regular LTA
  - RI-LTA
  - Comparing results. Deciding on the number of classes
  - Checking model fit and model modifications
    - Checking response pattern fit and bivariate fit
    - Measurement invariance across time
    - Residual associations across time
    - Lag-2 modeling
- Adding covariates. Covariate effects on transition probabilities
  - RI-LTA. Comparing results with Regular LTA
- Measurement invariance across individuals
  - Multiple-group analysis and direct effects of covariates on indicators
- Special topics
  - Modeling stationarity of transition probabilities
  - Mover-Stayer modeling
  - Distal outcome
- **Error messages**

- ML maximizes the log likelihood function by finding an estimate of a parameter that gives a zero first-order derivative and a negative second-order derivative



(a) Peaked logL. Small SE



(b) Flat logL. Large SE or non-identified

- Real-world situation for LTA/RI-LTA:
  - Bumpy logL due to mixtures (use many STARTS). See slide 116 of Short Course Topic 5 on the Mplus website
  - Models with many parameters (63 for output 12)

Error Messages. ML Estimation. Non-Identified Models

- The ML estimate of the parameter vector θ has covariance matrix $V(\hat{\theta}) = I^{-1}$, where I is the information matrix. The SEs are the square root of the diagonal values. TECH3 gives the covariance and correlation matrices
- The information matrix can be approximated by a first-order derivative product matrix (MLF), a negative second-order derivative matrix (ML), or a sandwich combining the two (MLR). It must be positive definite ($>$0 in the single parameter case)
- The more peaked the logL is, the smaller the parameter estimate variance (and the SE), that is, the better determined the estimate is
- When the logL is flat, the second-order derivative is zero and the parameter is not identified
- A model is not identified if 2 sets of parameter values given the same log likelihood value
- Empirical non-identification: Not identified in a certain data set

Error Messages. Non-Identified Models Continued

- A non-identified model has a singular information matrix (equals 0 in the single parameter case)
- Because the information matrix needs to be inverted (1/i in the single parameter case), it cannot be singular (you can't divide by zero)
- A singular matrix has at least one zero eigenvalue
- Condition number is a good indicator of a close to singular matrix: It is the ratio of smallest to largest eigenvalue ($< 10^{-10}$ typical threshold for numerically determining that the matrix is singular; $10^{-10} = 1/10^{10}$)
- During the inversion of the information matrix, Mplus finds the row, that is, the parameter that triggers falling below the condition number threshold and prints this parameter with its number from TECH1
- The first-order derivative product version of the information matrix is particularly sensitive to singularity/non-identification and is always checked in Mplus
- Asparouhov & Muthén (2021). First-order derivative warning message, condition number, and non-identification. FAQ: `http://www.statmodel.com/download/ConditionNumber.pdf`

Error Messages. Non-Identified Models

- Example:
  - 1-factor model with free loadings (λ) and free factor variance (ψ): $\psi \times c$, $\lambda/\sqrt(c)$ give the same logL for all non-zero c values
  - The logL is flat in the direction of ψ - all $\psi \times c$ values give the same logL value. Fixing ψ at any value gives the same logL

- Example message from RI-LTA where the factor variance has not been fixed at 1:
  - THE STANDARD ERRORS OF THE MODEL PARAMETER ESTIMATES MAY NOT BE TRUSTWORTHY FOR SOME PARAMETERS DUE TO A **NON-POSITIVE DEFINITE FIRST-ORDER DERIVATIVE PRODUCT MATRIX**. THIS MAY BE DUE TO THE STARTING VALUES BUT MAY ALSO BE AN INDICATION OF MODEL NONIDENTIFICATION. THE CONDITION NUMBER IS **0.687D-16**. PROBLEM INVOLVING THE FOLLOWING PARAMETER: Parameter 6, $\%C1\#1.C2\#1.C3\#1.C4\#1\%$: **F** (equality/label)
  - 0.1D-1 = 0.01 = 10^{-2} where D means double precision (15 digits)

Error Messages. An LTA Example. Output 2

- IN THE OPTIMIZATION, ONE OR MORE **LOGIT THRESHOLDS** APPROACHED EXTREME VALUES OF -15.000 AND 15.000 AND WERE FIXED TO STABILIZE MODEL ESTIMATION. THESE VALUES IMPLY PROBABILITIES OF 0 AND 1. IN THE MODEL RESULTS SECTION, THESE PARAMETERS HAVE 0 STANDARD ERRORS AND 999 IN THE Z-SCORE AND P-VALUE COLUMNS.
- ONE OR MORE **MULTINOMIAL LOGIT PARAMETERS** WERE FIXED TO AVOID **SINGULARITY OF THE INFORMATION MATRIX**. THE SINGULARITY IS MOST LIKELY BECAUSE THE MODEL IS NOT IDENTIFIED, OR BECAUSE OF EMPTY CELLS IN THE JOINT DISTRIBUTION OF THE CATEGORICAL LATENT VARIABLES AND ANY INDEPENDENT VARIABLES. THE FOLLOWING PARAMETERS WERE FIXED:
  - Parameter 26, C2#1 ON C1#2
  - Parameter 27, C2#2 ON C1#2
  - Parameter 35, C4#2 ON C3#2

Error Messages. An LTA Example Continued

- Measurement model estimates: Large logit thresholds and 0/1 probabilities

	Classes		
	1	2	3
Letrec	0.505	0.994	1.000
Begin	0.066	0.917	0.984
Ending	0.013	0.660	0.972
Sight	0.000	0.051	0.985
WIC	0.000	0.000	0.509

Latent Class C1#1			
Thresholds			
LETREC1$1	-0.018	0.048	-0.383
BEGIN1$1	2.657	0.133	19.992
ENDING1$1	4.360	0.208	20.933
SIGHT1$1	15.000	0.000	999.000
WIC1$1	15.000	0.000	999.000
Latent Class C1#2			
Thresholds			
LETREC1$1	-5.102	0.191	-26.645
BEGIN1$1	-2.396	0.062	-38.389
ENDING1$1	-0.665	0.038	-17.448
SIGHT1$1	2.919	0.113	25.842
WIC1$1	6.747	0.492	13.720
Latent Class C1#3			
Thresholds			
LETREC1$1	-15.000	0.000	999.000
BEGIN1$1	-4.101	0.133	-30.783
ENDING1$1	-3.563	0.112	-31.829
SIGHT1$1	-4.164	0.198	-21.026
WIC1$1	-0.036	0.037	-0.976

Error Messages. An LTA Example Continued

- Transition matrix with 0/1 probabilities

ONE OR MORE MULTINOMIAL LOGIT PARAMETERS WERE FIXED TO AVOID SINGULARITY OF THE INFORMATION MATRIX. THE SINGULARITY IS MOST LIKELY BECAUSE THE MODEL IS NOT IDENTIFIED, OR BECAUSE OF EMPTY CELLS IN THE JOINT DISTRIBUTION OF THE CATEGORICAL LATENT VARIABLES AND ANY INDEPENDENT VARIABLES. THE FOLLOWING PARAMETERS WERE FIXED:
Parameter 26, C2#1 ON C1#2
Parameter 27, C2#2 ON C1#2
Parameter 35, C4#2 ON C3#2

LATENT TRANSITION PROBABILITIES BASED ON THE ESTIMATED MODEL

C1 Classes (Rows) by C2 Classes (Columns)

	1	2	3
1	0.338	0.649	0.012
2	0.001	0.652	0.348
3	0.000	0.000	1.000

C2 Classes (Rows) by C3 Classes (Columns)

	1	2	3
1	0.596	0.402	0.002
2	0.002	0.837	0.161
3	0.002	0.003	0.994

C3 Classes (Rows) by C4 Classes (Columns)

	1	2	3
1	0.263	0.505	0.232
2	0.005	0.132	0.863
3	0.001	0.000	0.999

Error Messages. An LTA Example Continued

```
Categorical Latent Variables

 C2#1      ON
    C1#1              31.015       2.166      14.316       0.000
    C1#2              21.237       0.000     999.000     999.000

 C2#2      ON
    C1#1              29.948       0.284     105.304       0.000
    C1#2              26.619       0.000     999.000     999.000

 C3#1      ON
    C2#1              11.617       1.440       8.066       0.000
    C2#2               1.718       1.325       1.297       0.195

 C3#2      ON
    C2#1              10.836       1.652       6.560       0.000
    C2#2               7.307       1.287       5.678       0.000

 C4#1      ON
    C3#1               7.401       2.611       2.835       0.005
    C3#2               2.161       2.735       0.790       0.429

 C4#2      ON
    C3#1              27.555       0.173     159.627       0.000
    C3#2              24.900       0.000     999.000     999.000

 Means
    C1#1               3.420       0.121      28.300       0.000
    C1#2               2.526       0.126      20.097       0.000
    C2#1             -27.709       2.149     -12.896       0.000
    C2#2             -25.990       0.078    -334.999       0.000
    C3#1              -6.042       1.004      -6.018       0.000
    C3#2              -5.657       1.282      -4.414       0.000
    C4#1              -7.278       2.604      -2.795       0.005
    C4#2             -26.779       0.077    -348.807       0.000
```

```
PARAMETER SPECIFICATION FOR LATENT CLASS REGRESSION MODEL PART


           ALPHA(C)
              C1#1          C1#2          C1#3          C2#1          C2#2
              ________      ________      ________      ________      ________
                   16            17             0            18            19


           ALPHA(C)
              C2#3          C3#1          C3#2          C3#3          C4#1
              ________      ________      ________      ________      ________
                    0            20            21             0            22


           ALPHA(C)
              C4#2          C4#3
              ________      ________
                   23             0


           BETA(C)
              C1#1          C1#2          C1#3
              ________      ________      ________
 C2#1              24            26             0
 C2#2              25            27             0
 C2#3               0             0             0


           BETA(C)
              C2#1          C2#2          C2#3
              ________      ________      ________
 C3#1              28            30             0
 C3#2              29            31             0
 C3#3               0             0             0


           BETA(C)
              C3#1          C3#2          C3#3
              ________      ________      ________
 C4#1              32            34             0
 C4#2              33            35             0
 C4#3               0             0             0
```

Error Messages. An RI-LTA Example. Output 9

- RI-LTA with a covariate, interaction model (classes not reordered):
  - WARNING: THE **SAMPLE COVARIANCE OF THE INDEPENDENT VARIABLES** IN CLASS 4 IS SINGULAR.
  - THE STANDARD ERRORS OF THE MODEL PARAMETER ESTIMATES MAY NOT BE TRUSTWORTHY FOR SOME PARAMETERS DUE TO A **NON-POSITIVE DEFINITE FIRST-ORDER DERIVATIVE PRODUCT MATRIX**. THIS MAY BE DUE TO THE STARTING VALUES BUT MAY ALSO BE AN INDICATION OF MODEL NONIDENTIFICATION. THE CONDITION NUMBER IS **0.269D-15**. PROBLEM INVOLVING THE FOLLOWING PARAMETER: Parameter 32, C2#1 ON C1#1
  - ONE OR MORE MULTINOMIAL LOGIT PARAMETERS WERE FIXED TO AVOID SINGULARITY OF THE **INFORMATION MATRIX**. THE SINGULARITY IS MOST LIKELY BECAUSE THE MODEL IS NOT IDENTIFIED, OR BECAUSE OF EMPTY CELLS IN THE JOINT DISTRIBUTION OF THE CATEGORICAL LATENT VARIABLES AND ANY INDEPENDENT VARIABLES. THE FOLLOWING PARAMETERS WERE FIXED ...

Error Messages. An RI-LTA Example Continued

- WARNING: THE SAMPLE COVARIANCE OF THE INDEPENDENT VARIABLES IN CLASS 4 IS SINGULAR.
- A class-specific regression on an X variable is non-identified when the class-specific variance of X is zero
- TECH7 gives information on class-specific variances of X variables:
  - Class 4 refers to the 4th latent class pattern that is listed in TECH7: (1) 1 1 1 1, (2) 1 1 1 2, (3) 1 1 1 3, (4) 1 1 2 1
  - SAMPLE STATISTICS WEIGHTED BY ESTIMATED CLASS PROBABILITIES FOR PATTERN 1 1 2 1
Means POV 0.000
Covariances POV 0.000
  - The latent class pattern 1 1 2 1 (not reordered) has 28% of the children (according to TECH15), none with Pov=1

Error Messages. An RI-LTA Example Continued

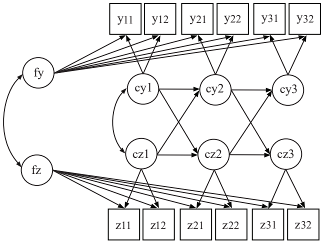
- THE FOLLOWING PARAMETERS WERE FIXED:
  - Parameter 33, $C2\#2$ ON $C1\#1$
  - Parameter 44, MODEL C1: $\%C1\#1\%$: $C2\#1$ ON POV
  - Parameter 45, MODEL C1: $\%C1\#1\%$: $C2\#2$ ON POV
  - Parameter 49, MODEL C1: $\%C1\#3\%$: $C2\#2$ ON POV
  - Parameter 50, MODEL C2: $\%C2\#1\%$: $C3\#1$ ON POV
  - Parameter 57, MODEL C3: $\%C3\#1\%$: $C4\#2$ ON POV
  - Parameter 51, MODEL C2: $\%C2\#1\%$: $C3\#2$ ON POV
  - Parameter 61, MODEL C3: $\%C3\#3\%$: $C4\#2$ ON POV
  - Parameter 41, $C4\#2$ ON $C3\#1$
  - Parameter 37, $C3\#2$ ON $C2\#1$
  - Parameter 35, $C2\#2$ ON $C1\#2$
- C ON C: See LATENT TRANSITION PROBABILITIES BASED ON THE ESTIMATED MODEL
- C ON Pov: See MODEL RESULTS and the TECH15 section FINAL CLASS COUNTS AND PROPORTIONS FOR THE LATENT CLASSES BASED ON THE ESTIMATED MODEL - Output 9

Final Comments

- Advantages of RI-LTA over regular LTA:
  - Better fit to data
  - More information extracted from data
  - Less concern about the need for measurement non-invariance across groups of individuals
  - Less concern about the need for indicator-specific residual correlation across time
  - Less concern about the neeed for lag-2 modeling
  - Less concern about the need for Mover-Stayer modeling
- Disadvantages:
  - Requires longer computing time
  - Requires larger samples

Final Comments: Model Extensions

- Latent class indicators can be binary, ordinal, nominal, count, continuous and combinations
- More than one process, e.g. RI-LTA-CLPM



- Trends - random slopes in addition to random intercepts
- Multilevel RI-LTA: individuals within schools, organizations, communities
- Multi-step methods not discussed in detail here but are useful for distal outcomes - maybe less so for exploring covariate effects on latent classes

Final Comments Continued

- Latent class representation is not always the best model for the data
  - Mplus Web Talk No. 1, Segment 14, slides 69-70 show that a longitudinal factor model with a random intercept fits the Reading data better
  - `http://www.statmodel.com/MplusWebTalks.shtml`
- This web talk can be referred to as:

  Muthén, Bengt [Mplus]. (2021, February 28). Using Mplus To Do Latent Transition Analysis And Random Intercept Latent Transition Analysis [Video file]. Retrieved from `https://www.youtube.com/c/MplusVideos`.