JOHN FLOURNOY
Research Scientist
Harvard University
john_flournoy@g.harvard.edu

ETHAN MCCORMICK
Postdoctoral Researcher
Rutgers University
emccormick@unc.edu

Introduction to Longitudinal Structural Equation Modeling: Theory

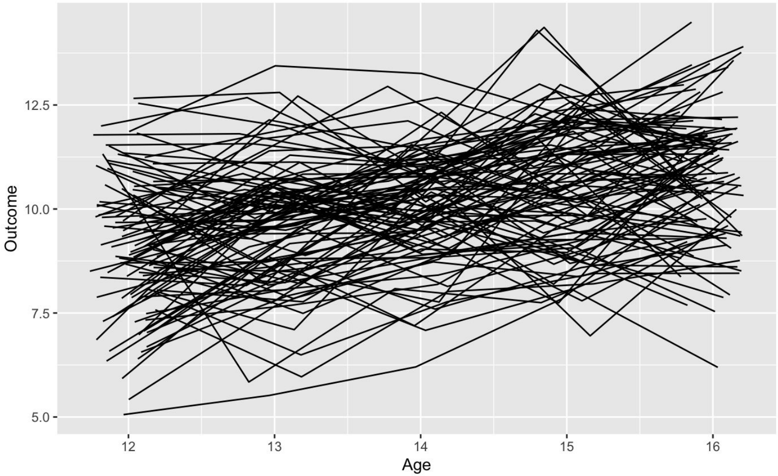
ABCD Workshop on Brain Development and Mental Health

Sponsored by NIMH

June 22nd - July 22nd, 2021

Topics we will cover

We have data:



But which model do we pick??

- Track I
  - Raw versus residualized change
  - Path Diagrams
  - Autoregressive Cross-Lag Panel Models + extensions
  - Latent Curve Model
    - Fixed and random effects
- Track II
  - Latent Curve Models review + extensions
  - Multivariate LCMs + extensions
  - Latent Change Score Models
    - Univariate
    - Multivariate

Introduction to Longitudinal Structural Equation Modeling: Theory
Track I

Different Kinds of Change over Time

Consider y_1 and y_2 as repeated measures

1. Residualized Change

- Differences in residuals (conditioning on prior information)

2. Raw Score Change

- Difference scores
- Smooth trajectories
- Latent change scores
  - A special case of both approaches

Different Kinds of Change over Time

Residualized Change

- Change in value of y_2 beyond what is expected based on y_1
  - Can use a regression expression to predict y_2 from y_1
  - $y_{2i} = \beta_0 + \beta_1 y_{1i} + \varepsilon_i$
  - $\hat{y}_{2i} = \beta_0 + \beta_1 y_{1i}$
  - $y_{2i} = \hat{y}_{2i} + \varepsilon_i$

- So residualized change is $\varepsilon_i = y_{2i} - \hat{y}_{2i}$
  - i.e., change from the model's expectation

- Used in ANCOVA and ARCL models

Different Kinds of Change over Time

Residualized Change

- Pros
  - y_1 and y_2 do **not** even have to be the same measure (remember it's just a regression)
    - May be important in contexts of measurement non-invariance (a fun term that we won't have time to go into, but bother Ethan or John about it sometime if you suffer from insomnia (why do we love what hurts us?))
  - Can help control for background characteristic differences in some contexts

- Cons
  - Inconsistent with most longitudinal theories which concern absolute change (i.e., trajectories)
  - Regressions link pairs of timepoints at a time rather than the full timeseries
    - Parameter rich models

Different Kinds of Change over Time

Raw Score Change

- Quite literally based on the difference in magnitude between y_2 and y_1
  - $d_i = y_{2i} - y_{1i}$

- Forms the basis of most longitudinal models
  - Paired t-test (which is **barely** a longitudinal model), repeated measures ANOVA, growth models, latent change score models
    - "**two** waves of data **are better than one, but** maybe **not much better**" (Rogosa et al., 1982)
  - For most of these models we don't directly compute difference scores, but change on the raw metric is the unit of analysis

Different Kinds of Change over Time

Raw Score Change

- Pros
  - Intuitive and what you would likely think of in terms of longitudinal modeling
- Cons
  - If there is measurement error on y_1 and/or y_2 (and there will be; heartbreak #1 of the day), that will get compounded in the difference score
  - Unreliability will decrease the power of our model test of change (heartbreak #2)
  - Fair warning: we'll generally accept these limitations for reasons (we'll explain when you're older; in approx 20 minutes)

Different Kinds of Change over Time

- Of course, people have calm, rational disagreements about this
    - Just kidding, quantitative methodologists are **dramatic**, especially in the 20th century

- Cronbach & Furby (1970); (yes that Cronbach)
    - "...they [raw change scores] are still employed, even by some otherwise sophisticated investigators."

- Willett (1997)
    - "When discussing residualized change scores, methodologists disagree as to what exactly is being estimated, how well it is estimated, and how it can be interpreted."
    - So that's all…

- Lord's paradox (as solved by Judea Pearl, 2016)

Different Kinds of Change over Time: Readings

- Cronbach, L. J., & Furby, L. (1970). How we should measure" change": Or should we?. *Psychological Bulletin*, *74*(1), 68.
- Willett, J. B. (1997). Measuring change: What individual growth modeling buys you. *Change and Development: Issues of theory, method, and application*, *213*, 243.
- Rausch, J.R., Maxwell, S.E. & Kelley, K. (2003). Analytic methods for questions pertaining to a randomized pretest, posttest, follow-up design. *Journal of Clinical and Consulting Psychology, 32*, 467-486.
- Rogosa, D. R. (1995). Myths and methods: "Myths about longitudinal research," plus supplemental questions. In J. M. Gottman (Ed.) *The Analysis of Change* (pp. 3-65). Hillsdale, New Jersey: Lawrence Erlbaum.
- Willett, J.B., Singer, J.D., and Martin, N.C. (1998). The design and analysis of longitudinal studies of development and psychopathology in context: Statistical models and methodological recommendations. *Development and Psychopathology, 10*, 395-426.
- Hendrix, L. J., Carter, M. W., & Hintze, J. L. (1979). A comparison of five statistical methods for analyzing pretest–posttest designs. *Journal of Experimental Education, 47*, 96–102

Equations and Path Diagrams in SEM

- SEM can be entirely expressed in terms of equations of course, but a handy visual shortcut has been developed to represent models that is often more intuitive for users
  - The important thing to realize though is that the visuals are **isomorphic** with the underlying equations, meaning that a full diagram implies all the same information as the equations (although there are some shortcuts we'll take to reduce visual clutter)

- We'll talk about 2 broad types of path diagrams that we'll see variations of
  - Path models (e.g., observed mediation analysis, autoregressive cross-lag models)
  - Factor models (e.g., growth models being a special case)

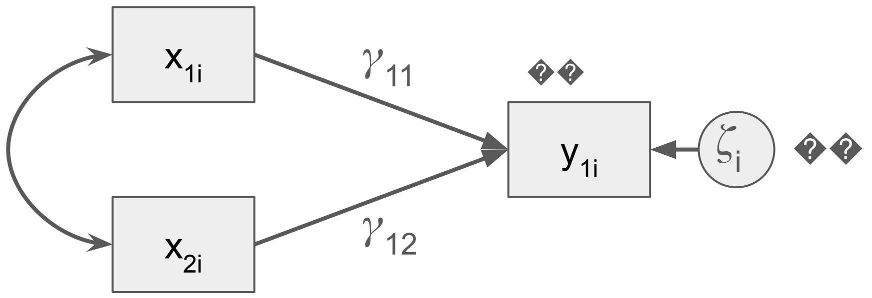
Equations and Path Diagrams in SEM

- If we consider a simple 2-predictor regression, the equation is as follows:
  - $y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \varepsilon_i$ with a normally distributed error term, $\varepsilon_i \sim N(0, \sigma^2)$
  - In SEM, we generally use slightly different notations
    - Intercepts: α
    - Regression coefficients of y's on x's: γ
    - Regression coefficients of y's on other y's: β
    - Residuals: ζ with variance Ψ
  - So now we express the above as: $y_i = \alpha + \gamma_1 x_{1i} + \gamma_2 x_{2i} + \zeta_i$ where $\zeta_i \sim N(0, \Psi)$

Equations and Path Diagrams in SEM

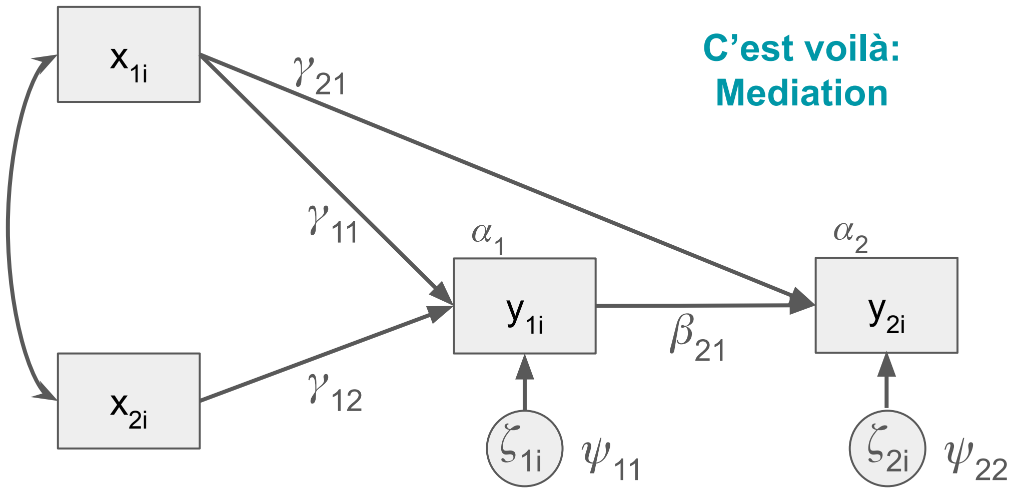
- However we can represent $y_i = \alpha + \gamma_1 x_{1i} + \gamma_2 x_{2i} + \zeta_i$ where $\zeta_i \sim N(0, \Psi)$ visually too

- Diagram conventions
  - Observed variables are rectangles
  - Latent variables (including residuals, remember we don't directly observe them) are circles
  - Single-headed arrows are regression coefficients and double-headed arrows are covariances



Equations and Path Diagrams in SEM

- Of course path models are usually more complicated than simple regression and allow for multiple dependent (y) variables
    - A simple case below:



C’est voilà: Mediation

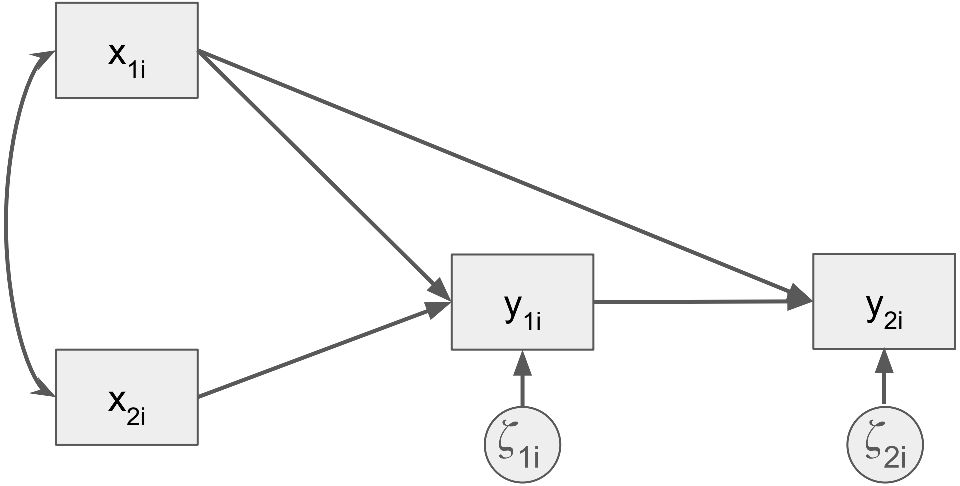
Model Equations:

$$y_{1i} = \alpha_1 + \gamma_{11}x_{1i} + \gamma_{12}x_{2i} + \zeta_{1i}$$

$$y_{2i} = \alpha_2 + \gamma_{21}x_{1i} + \beta_{21}y_{1i} + \zeta_{2i}$$

Equations and Path Diagrams in SEM

- Very often we drop alphas, psi's and gammas/betas and imply their existence
  - This would be to display your theoretical model; you could also put model estimates in place of the parameters in the results



Equations and Factor Diagrams in SEM

A factor in SEM represents a latent variable

A latent variable is some quantity that we *infer* from a set of observations.

Examples you've seen in this workshop and in the wild:

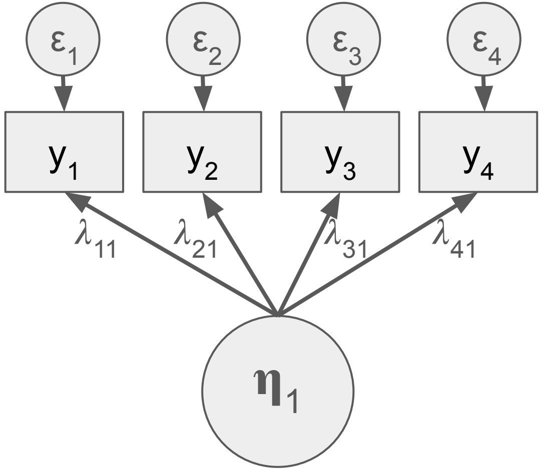
- Literally any computed score from a questionnaire
- The slope and intercept in a MLM growth model
    - In MLM, we're focus on the population mean of these (i.e., the fixed effects)
    - But we get the variance of the latent variables as the SD of the random effects

We'll cover the SEM growth model later, but it's really a restricted (i.e., constrained) form of a factor model.

Equations and Factor Diagrams in SEM

- The equation (bold lowercase terms are vectors; bold uppercase are matrices)
  - $\mathbf{y}_i = \boldsymbol{\Lambda}\boldsymbol{\eta}_i + \boldsymbol{\varepsilon}_i$

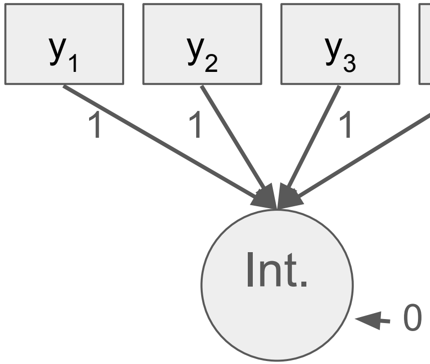
- Remember that we do not observe $\boldsymbol{\eta}$, but can infer its existence if we assume this model



Equations and Factor Diagrams in SEM

The scale score example

- When you take the average of a set of items and call that “Internalizing”, you’re creating a sort of latent variable model with some very strong constraints.
    - It’s a formative rather than reflective model
- Every item is equally formative of the latent variable.
- Every item and the construct is measured *without any error.*



Equations and Factor Diagrams in SEM

Latent variables (in SEM, MLM, and other contexts) allow us to model the true score of a variable free from measurement error.

This is awesome.

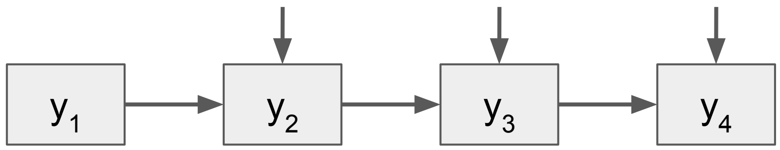
Again, feel free to bug us about measurement anytime.

Equations and Diagrams: Readings

- Bollen, K. A. (1989). *Structural Equations with Latent Variables*. New York: Wiley
- Bollen, K. A., & Davis, W. R. (2009). Two rules of identification for structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal, 16*, 523-536.
- MacCallum, R.C., & Austin, J.T. (2000). Applications of structural equation modeling in psychological research. *Annual Review of Psychology, 51*, 201-226.
- Raykov, T., & Marcoulides, G.A. (2006). *A First Course in Structural Equation Modeling, Second Edition*. Mahwah, NJ: Lawrence Earlbaum & Associates.

Autoregressive Cross-Lag Panel Models: Theory

- We saw before that **residualized change** is just a form of regression of a given measure on itself at a previous time
  - With SEM's ability to have multiple dependent variables, we can string several repeated measures together using a path model

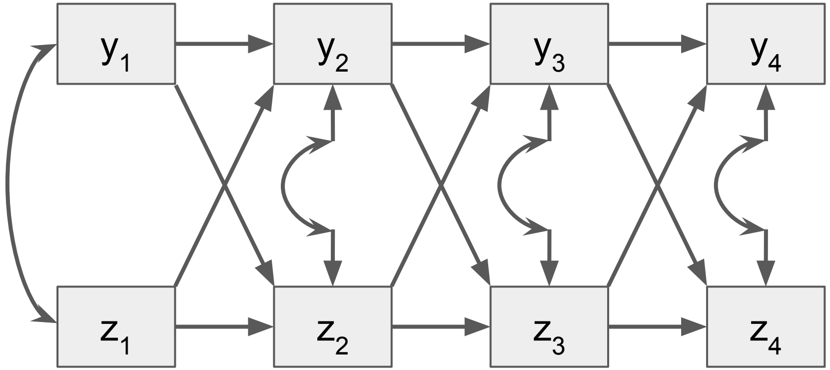


Note that we simplify the path model further by having orphaned arrows indicate the residuals.

Autoregressive Cross-Lag Panel Models: Theory

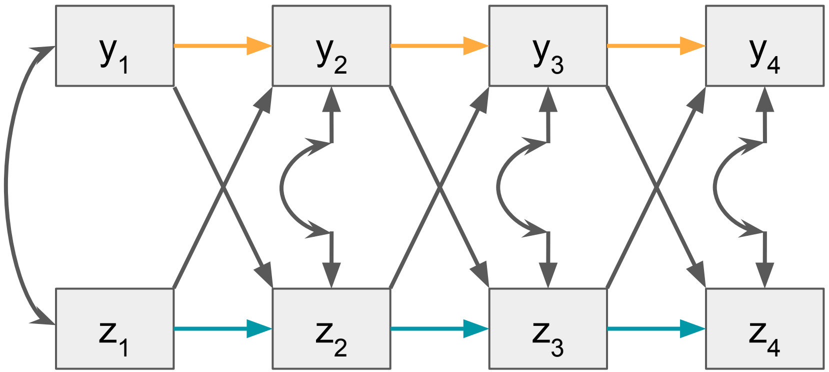
- Of course this is somewhat boring with a single variable
  - Raw change scores are often univariate (i.e., using a change score as the outcome of another variable), but ARCLs are almost never
  - Can test how two variables “travel together” over time
    - autoregressive and cross-lagged effects

One of the ONLY times quantitative methods people name something exactly what it is. Of course they also call these cross-lag panel models



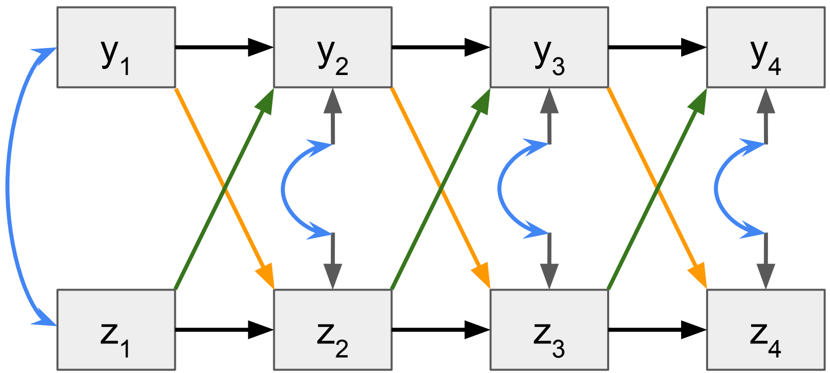
Autoregressive Cross-Lag Panel Models: Path Diagram

- Autoregressive Effects estimate the stability of a construct across time
  - Estimated separately for each construct (orange for y's, green for z's)
  - And for each lag, unless you impose equality constraints



Autoregressive Cross-Lag Panel Models: Path Diagram

- Cross-Lag Effects estimate the prospective prediction of y on z (orange), net the effect of y at the previous time point
  - And vice versa for z on y (green)
- Within time measures (or their residuals) are allowed to correlate (blue)



Autoregressive Cross-Lag Panel Models: Expansions

- Can include
  - additional repeated measures and look at chains of regression across time
  - exogenous predictors
  - Higher order lagged regression paths (e.g., lag-2)

- Instead of observations at each time point, could model a latent variable
  - VERY parameter intensive

- Can impose equality constraints to reduce parameter overload (a technical term)

Autoregressive Cross-Lag Panel Models: Limitations & Extensions

- Without equality constraints, ARCLs are highly parameterized models
  - Difficult to interpret change over time
  - No best practices for ordering of equality constraints

- Does not model change in means **at all** (can’t get a trajectory out of these models)

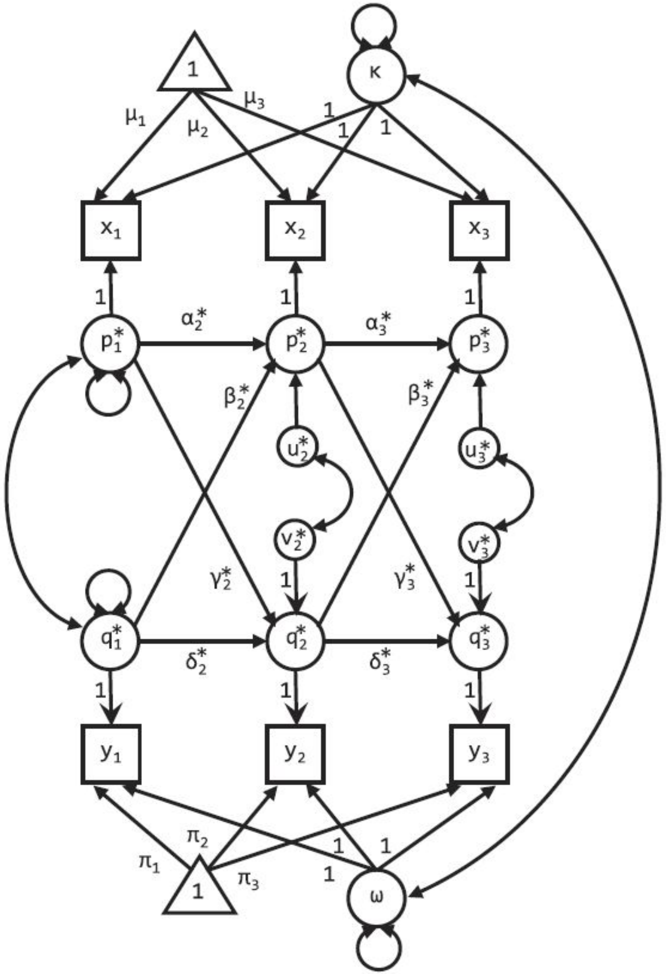
- Fails to disaggregate within-person change from between person differences
  - There is a fix (known as RI-ARCL [aka RI-CLPM] or more generally, structured residual growth models)
  - See more in Track II

- Probably don’t ever use a standard ARCL model. Please.

Random Intercept ARCL model

- Never use the standard ARCL model (aka CLPM).
- Use this model, the RI-ARCL (RI-CLPM) which separates out between and within person variance.
- You can easily specify and fit this model with the riclpm package (handles more than 2 vars!)
- Here's a more in-depth tutorial: https://johnflournoy.science/2017/10/20/riclpm-lavaan-demo/

RI-CLPM with means



Autoregressive Cross-Lag Panel Models: Readings

- Cole, D.A., & Maxwell, S.E. (2003). Testing mediational models with longitudinal data: questions and tips in the use of structural equation modeling. *Journal of Abnormal Psychology, 112*, 558-577.
- Rogosa, D., & Willett, J. B. (1985). Satisfying a simplex structure is simpler than it should be. *Journal of Educational and Behavioral Statistics, 10*, 99-107.
- Maxwell, S.E., Cole, D.A., & Mitchell, M.A. (2011). Bias in cross-sectional analyses of longitudinal mediation: Partial and complete mediation under an autoregressive model. *Multivariate Behavioral Research, 46*, 816-841.
- Curran, P.J., & Willoughby, M.T. (2003). Implications of latent trajectory models for the study of developmental psychopathology. *Development and Psychopathology*, *15*, 581-612
- Hamaker, E.L., Kuiper, R.M., & Grasman, R.P. (2015). A critique of the cross-lagged panel model. *Psychological Methods, 20*, 102-116.

Latent Curve Model: Theory

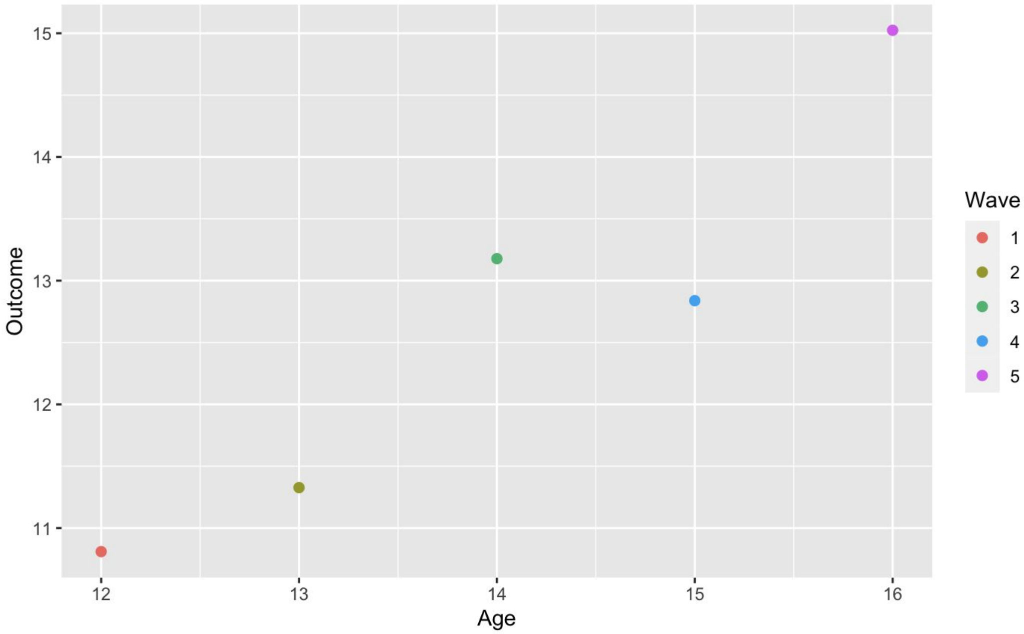
Seriously, read everything this man writes about SEM

- Latent variables are all around (Bollen, 2002)
  - Constructs that we do not (or cannot) directly measure but believe to exist
    - Actually, we cannot directly measure anything
      - Do you remember high-school physics, measuring everything 3 times?
  - Depression, cognitive control, etc

- In the context of growth, we assume that people start somewhere and grow in a certain direction
  - BUT, we can't actually measure true growth. We use repeated observations to infer the trajectory
  - If you've ever fit a random slope mixed effect growth model, you've already used latent variables!
  - Let's look at some pictures

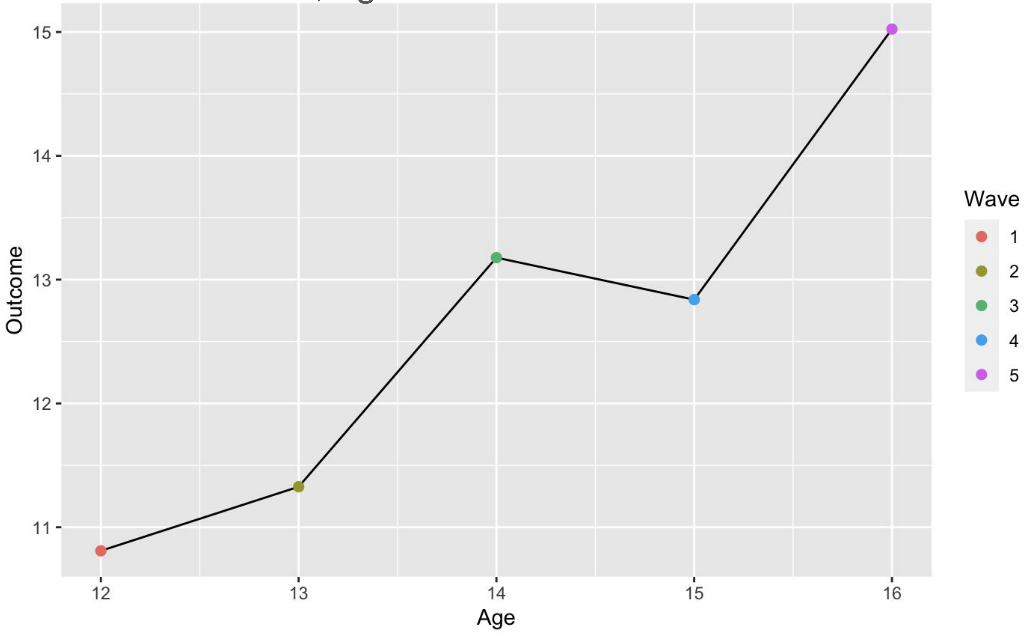
Latent Curve Model: Theory

- We have repeated measures, but how can we assess change?



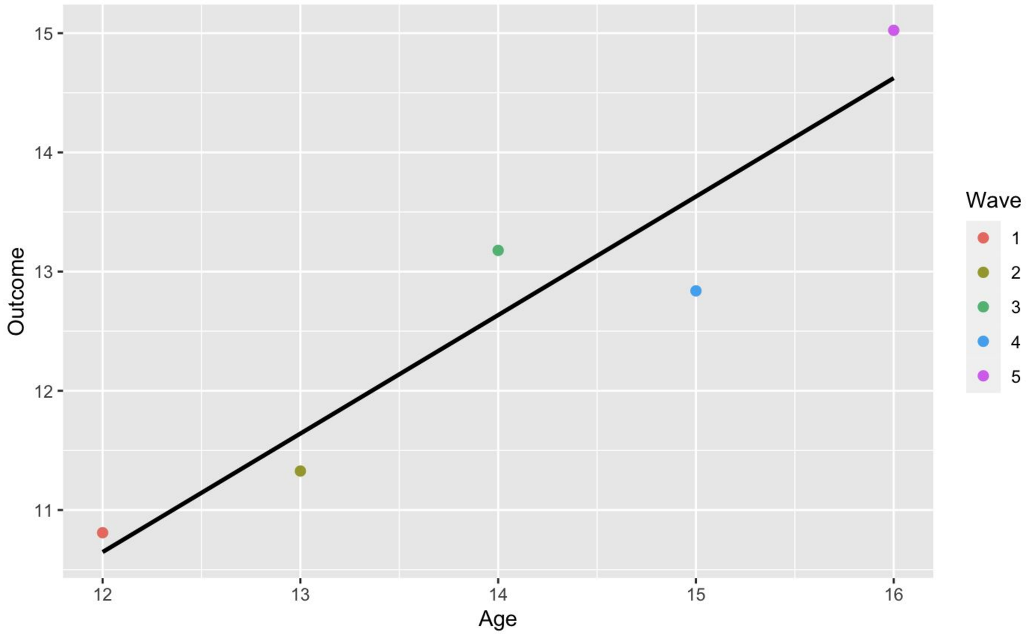
Latent Curve Model: Theory

- Could connect each time point
  - This would assess change interval by interval (a collection of difference scores)
  - This assumes that each observation is measured without error, and that change is linear between observations. Weird, right?



Latent Curve Model: Theory

- OR, could fit a smooth line to the data
  - Creates a trajectory with a model-implied intercept and slope
  - Separates true growth from measurement error
  - Rests on the assumption of linearity (first degree)



Latent Curve Model: Theory

- Trajectories attempt to capture underlying development in repeated measures
    - Views repeated measures as composed of both true change (a smooth underlying trajectory) and noise (time specific deviation/residuals of the observed scores from the trajectory)

- For a broad class of basic models, is mathematically identical to the mixed effect growth model
    - Some good reading: Curran, 2003; Bauer, 2003
    - SEM can be expanded to much more complex models that no longer have mixed effect equivalents (see Track II for some examples)

Latent Curve Model: Path Diagram

- To build this trajectory, we use a highly constrained confirmatory factor model
  - For a first-degree linear model, we define two factors (1 intercept and 1 linear slope)
  - Identity of the factor is determined by setting factor loadings to specific values

- Wait, what's it mean to identify a factor? (very briefly)
  - SEM is a system of multiple equations.
  - The number of unknown parameters must be less than the number of pieces of information you give the model. Setting loadings to specific values leaves fewer parameters unknown (to be estimated)

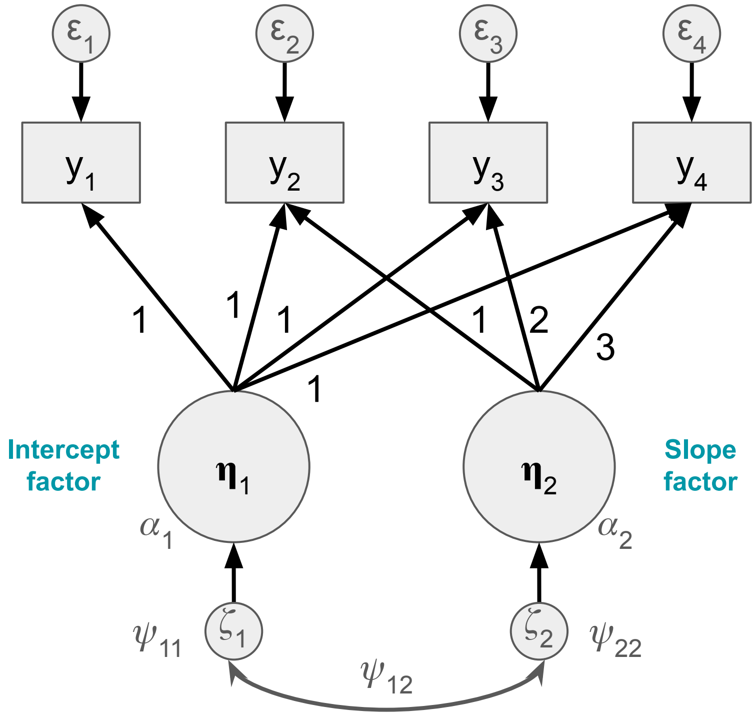
Latent Curve Model: Path Diagram

Measurement Equation

$$y_i = \Lambda \eta_i + \varepsilon_i$$

Structural Equation

$$\eta_i = \alpha + \zeta_i$$



Model Implied Mean Structure

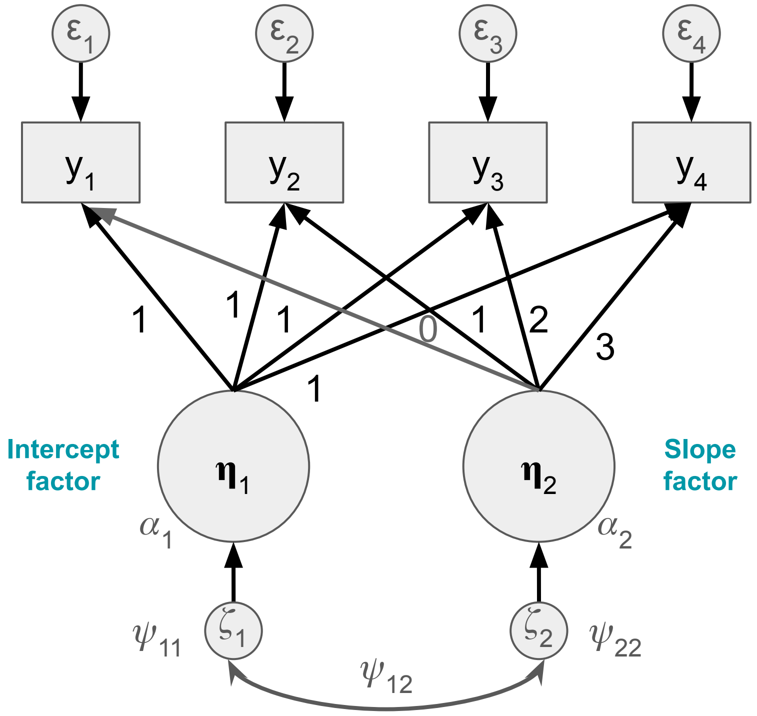
$$\mu(\theta) = \Lambda\alpha$$

Model Implied Covariance Structure

$$\Sigma(\theta) = \Lambda\Psi\Lambda' + \Theta$$

This part is basically magic. We're trying to reproduce the means and covariance of the observed variables with only the model structure….

Latent Curve Model: It's very similar to MLM



This is essentially:

$y_{ij} = \gamma_{00} + \gamma_{10}\text{TIME} + \nu_{00} + \nu_{00} + \varepsilon_{ij}$,

where TIME is 0, 1, 2, 3 (structured appropriately for the observations)

Notice that TIME is centered at the first timepoint implicitly in this model (see track 2 for more).

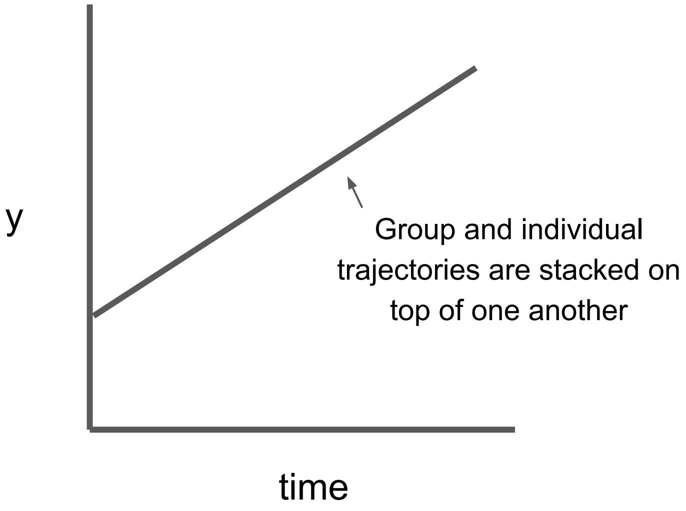
Latent Curve Model: Fixed vs Random Effects

- Just like in mixed-effect growth models we can have fixed (population mean) and random (individual) trajectories
  - We either set the variance of the factor to zero (fixed only) or freely estimate (fixed + random)
  - We never **estimate** individual trajectories, but we can compute model-implied trajectories on the back end

- Probably easier to see visually

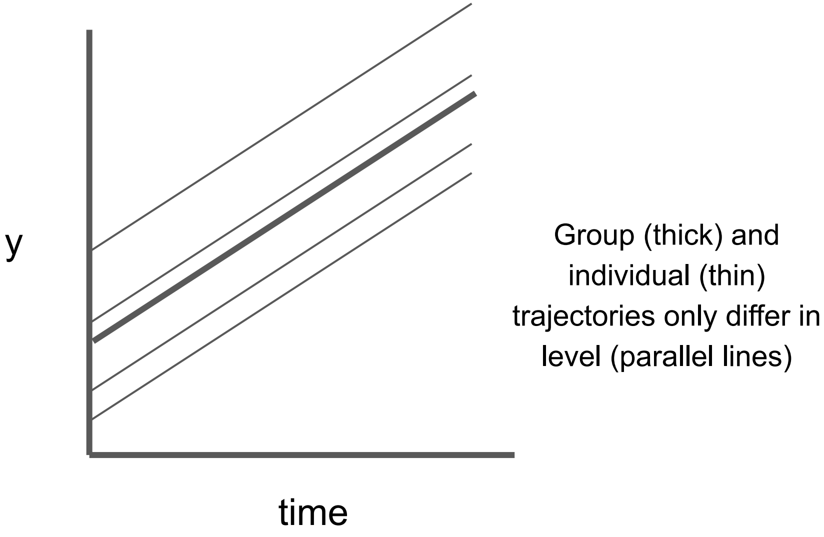
Latent Curve Model: Fixed vs Random Effects

Fixed Intercept, Fixed Slope



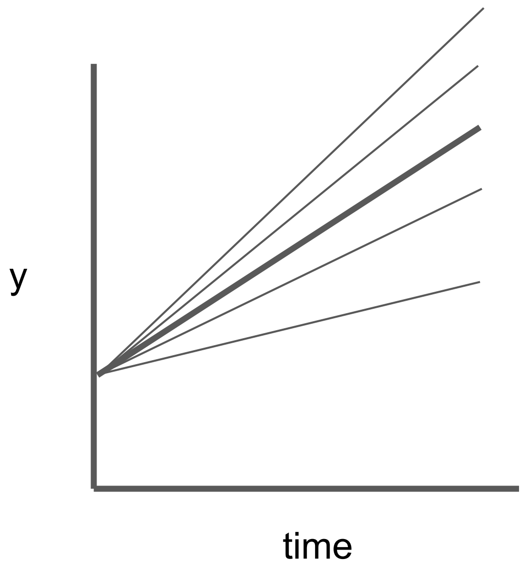
You would never do this, but it's basically OLS regression

Random Intercept, Fixed Slope



Latent Curve Model: Fixed vs Random Effects

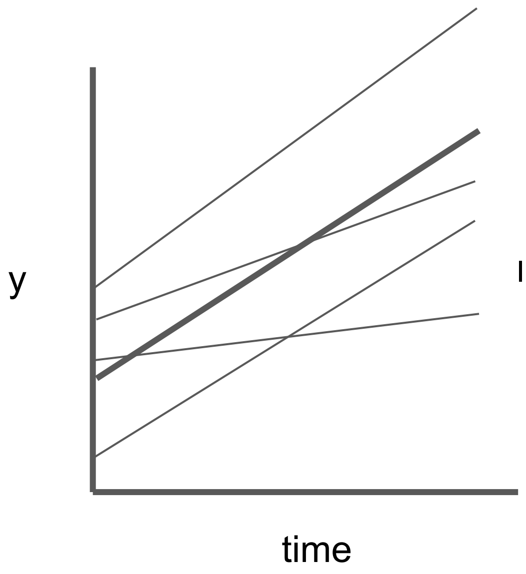
Fixed Intercept, Random Slope



Non-parallel lines radiate from a common intercept

Relatively uncommon but sometimes makes sense

Random Intercept, Random Slope



Individuals differ both in intercept and rate of change over time

Notice that rank order is NOT preserved across time

Latent Curve Model: Extensions

- Can vary the number/nature of factors
  - Intercept only, linear, quadratic, splines, latent basis
- Can bring in predictors and outcomes at the structural and measurement levels
- Multivariate Models
  - The **real** unique strength of SEM
  - but see also Bayesian modeling packages like brms and languages like Stan, too
- Structuring the residuals
  - Separation of within and between person variance
  - Combines the strengths of the ARCL without many of the limitations

Latent Curve Model: Readings

- Willett, J. B., & Sayer, A. G. (1994). Using covariance structure analysis to detect correlates and predictors of change. *Psychological Bulletin, 116*, 363-381.
- Bollen, K.A., & Curran, P.J. (2006). *Latent Curve Models: A Structural Equation Approach*. Wiley Series on Probability and Mathematical Statistics. John Wiley & Sons: New Jersey.
- Biesanz, J.C., Deeb-Sossa, N., Aubrecht, A.M., Bollen, K.A., & Curran, P.J. (2004). The role of coding time in estimating and interpreting growth curve models. *Psychological Methods, 9*, 30-52.
- Neale, M. C., Aggen, S. H., Maes, H. H., Kubarych, T. S., & Schmitt, J. E. (2006). Methodological issues in the assessment of substance use phenotypes. *Addictive Behaviors, 31*, 1010-1034.
- Curran, P. J. (2003). Have multilevel models been structural equation models all along?. *Multivariate Behavioral Research*, *38*(4), 529-569.
- Curran, P. J., Obeidat, K., & Losardo, D. (2010). Twelve frequently asked questions about growth curve modeling. *Journal of Cognition and Development*, *11*(2), 121-136.

Introduction to Longitudinal Structural Equation Modeling: Theory Track II

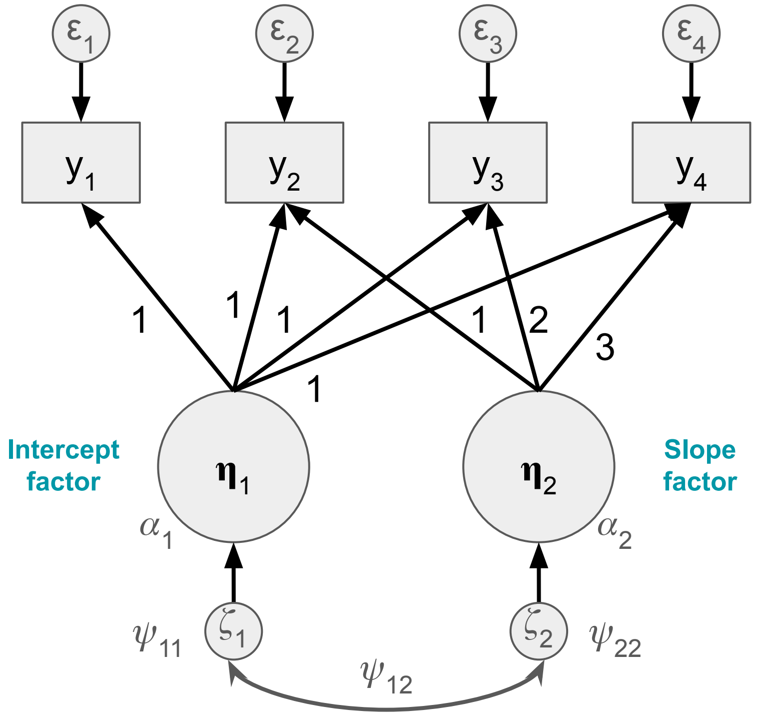
Latent Curve Model: A quick review

Measurement Equation

$$\mathbf{y}_i = \boldsymbol{\Lambda}\boldsymbol{\eta}_i + \boldsymbol{\varepsilon}_i$$

Structural Equation

$$\boldsymbol{\eta}_i = \boldsymbol{\alpha} + \boldsymbol{\zeta}_i$$



Model Implied Mean Structure

$$\boldsymbol{\mu}(\boldsymbol{\theta}) = \boldsymbol{\Lambda}\boldsymbol{\alpha}$$

Model Implied Covariance Structure

$$\boldsymbol{\Sigma}(\boldsymbol{\theta}) = \boldsymbol{\Lambda}\boldsymbol{\Psi}\boldsymbol{\Lambda}' + \boldsymbol{\Theta}$$

This part is basically magic. We're reproducing the means and covariance of the observed variables with only the factor structure…(okay so we may be a bit lame [really me mainly])

Latent Curve Model: A quick review

- Centering
  - Where we code time as zero influences the interpretation of the intercept and covariance of the latent factors, but model fit will be identical
  - In higher order models, centering will affect all but the highest-order factor

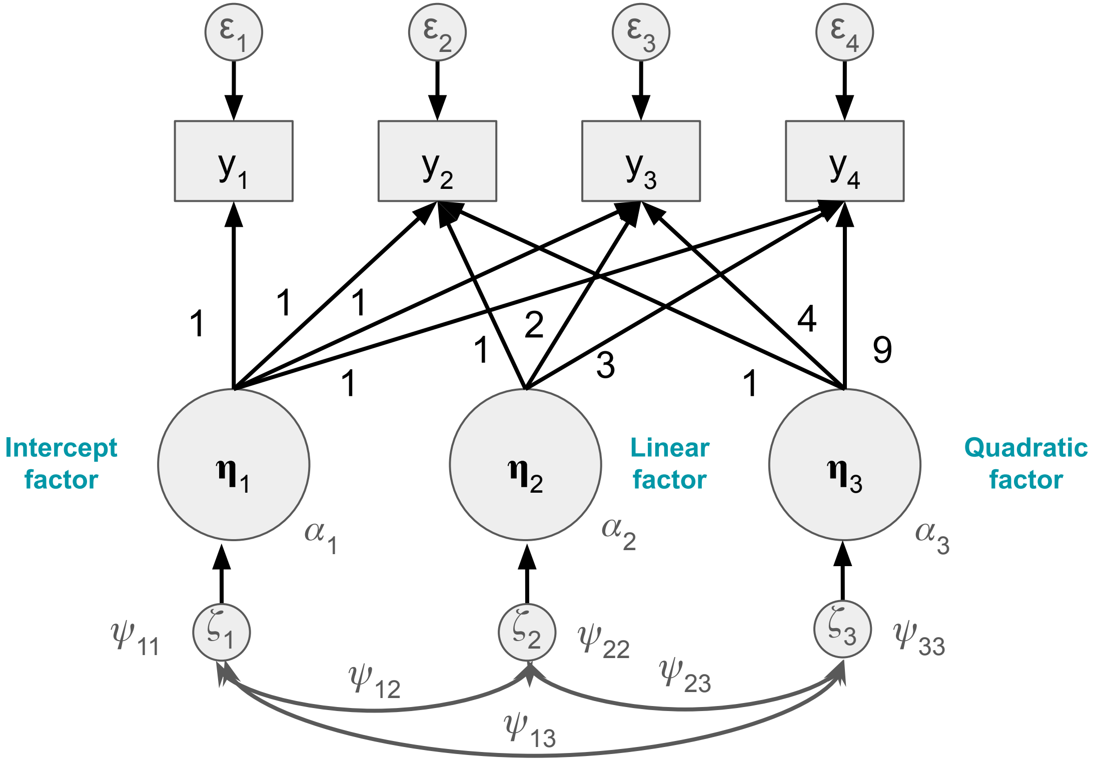
- Time structure
  - Time is coded directly into the factor loadings and in traditional LCMs we need observed variables at discrete time points
  - Missing data is easily accommodated but we need “sufficient” observations at each discrete time point
  - Continuous time can be included with definition variables/tscore model (need other software)

Latent Curve Model: Readings

- Willett, J. B., & Sayer, A. G. (1994). Using covariance structure analysis to detect correlates and predictors of change. *Psychological Bulletin, 116*, 363-381.
- Bollen, K.A., & Curran, P.J. (2006). *Latent Curve Models: A Structural Equation Approach*. Wiley Series on Probability and Mathematical Statistics. John Wiley & Sons: New Jersey.
- Biesanz, J.C., Deeb-Sossa, N., Aubrecht, A.M., Bollen, K.A., & Curran, P.J. (2004). The role of coding time in estimating and interpreting growth curve models. *Psychological Methods, 9*, 30-52.
- Neale, M. C., Aggen, S. H., Maes, H. H., Kubarych, T. S., & Schmitt, J. E. (2006). Methodological issues in the assessment of substance use phenotypes. *Addictive Behaviors, 31*, 1010-1034.
- Curran, P. J. (2003). Have multilevel models been structural equation models all along?. *Multivariate Behavioral Research*, *38*(4), 529-569.
- Curran, P. J., Obeidat, K., & Losardo, D. (2010). Twelve frequently asked questions about growth curve modeling. *Journal of Cognition and Development*, *11*(2), 121-136.

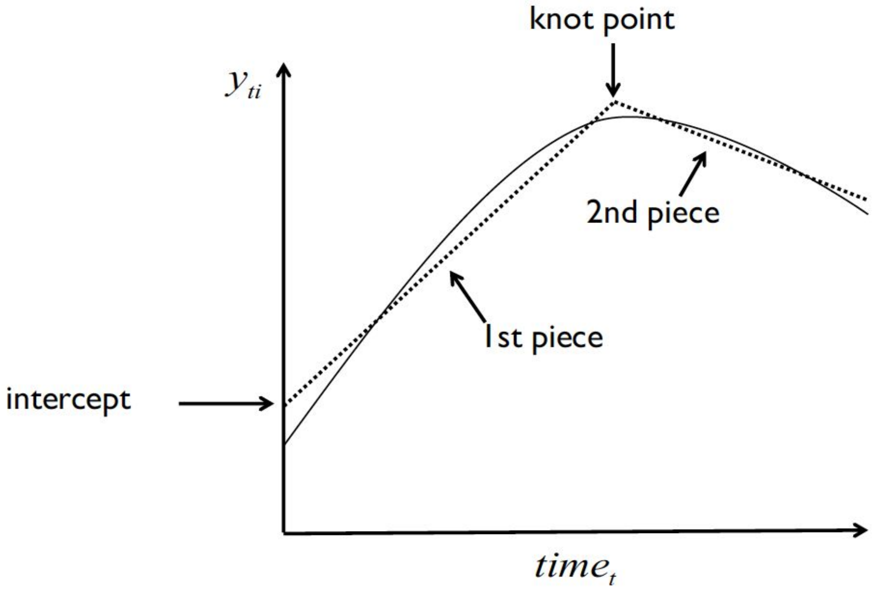
Latent Curve Model: Nonlinear Trajectories

- Higher-order Polynomials
  - A quadratic trend can be imposed by squaring the linear factor loadings
  - Often very difficult to estimate a quadratic random effect
    - That’s fine, can just fix the variance to zero
  - Even higher polynomials can be specified by this is almost never done

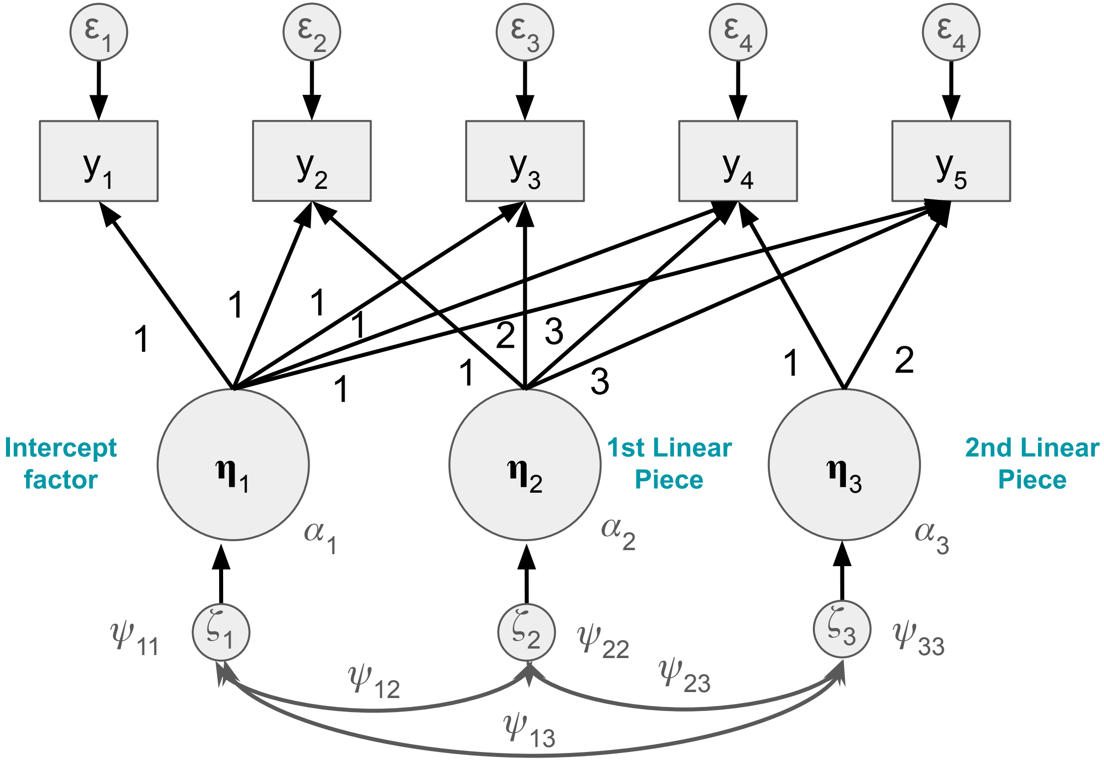


Latent Curve Model: Nonlinear Trajectories

- Piecewise Splines
  - One alternative to a smooth polynomial is to do local curve approximation with discontinuous functions
  - For linear piecewise models, you need at least 5 timepoints (3 for each line w/ 1 shared “knot point”)
  - Advanced applications can have non-linear pieces and can even estimate the knot point

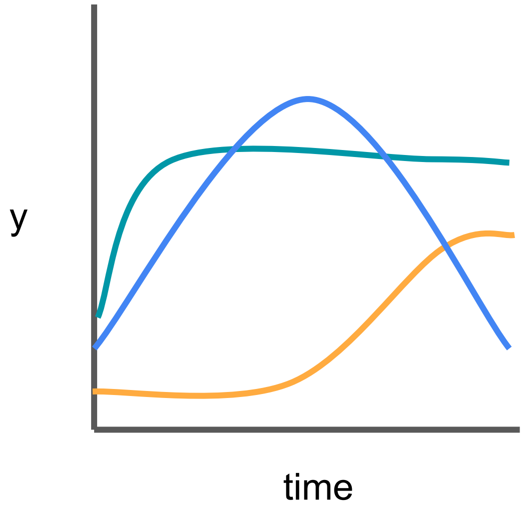


Latent Curve Model: Nonlinear Trajectories

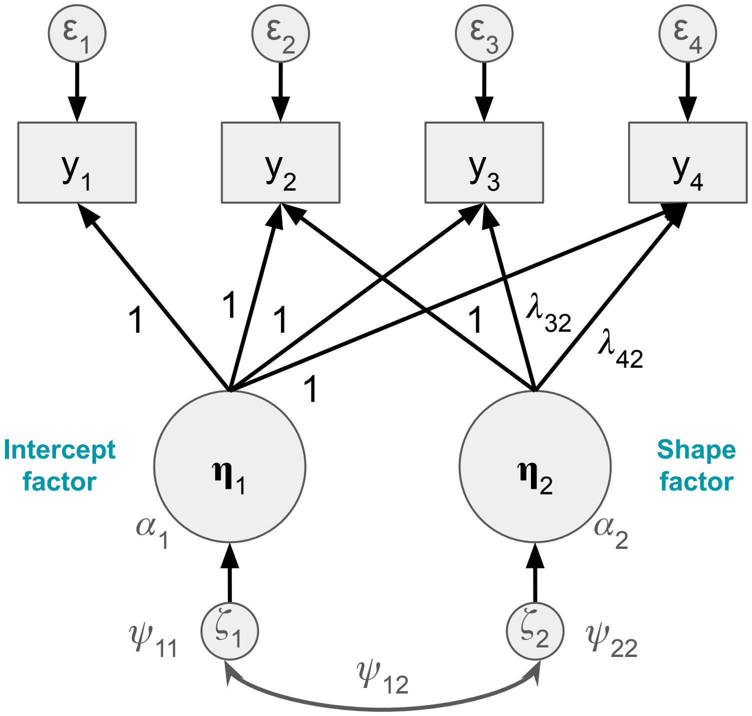


Latent Curve Model: Nonlinear Trajectories

- Latent Basis Models
  - When we think developmental patterns may be complex non-linear, we might fit a free-loading model
    - Allows a subset of factor loadings to be freely estimated
    - Can fit a complex developmental surface
  - Comes with some limitations
    - Can no longer interpret beta as a per-unit change in y
    - Random effects of “the shape”?
    - Lots of potential to overfit the data



Latent Curve Model: Nonlinear Trajectories



Nonlinear Latent Curve Model: Readings

- Biesanz, J.C., Deeb-Sossa, N., Aubrecht, A.M., Bollen, K.A., & Curran, P.J. (2004). The role of coding time in estimating and interpreting growth curve models. *Psychological Methods, 9*, 30-52.
- Flora, D. B. (2008). Specifying piecewise latent trajectory models for longitudinal data. *Structural Equation Modeling, 15*, 513-533.
- Meredith, W., & Tisak, J. (1990). Latent curve analysis. *Psychometrika, 55*, 107-122.
- McArdle, J. J. (2009). Latent variable modeling of differences and changes with longitudinal data. *Annual Review of Psychology, 60*, 577-605
- Cudeck, R., & Harring, J.R. (2007). Analysis of nonlinear patterns of change with random coefficient models. *Annual Review of Psychology, 58,* 615-637.

LCM: Time-Invariant Covariates (TICs)

- Predictors that vary at the person level, but do not take on unique values over time
  - Theoretical: age at first child birth, childhood adversity, treatment group
  - Convenience: brain function at age 10, self-identified sex

- Enter the model at the latent variable level
  - TICs can predict conditional changes in the intercept and slope
  - Just a quick note: intercept and slope contain **no** temporal information

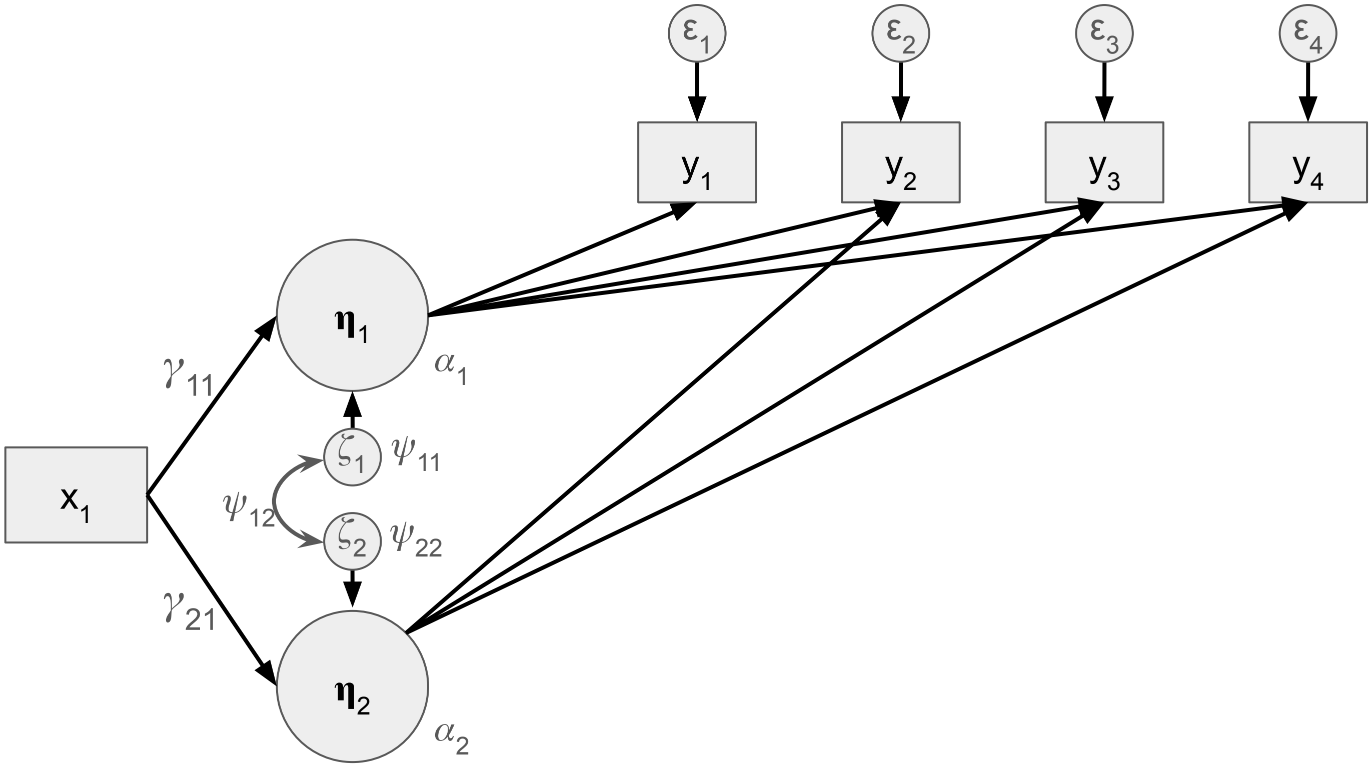
LCM: Time-Invariant Covariates

Measurement Equation

$$\mathbf{y}_i = \Lambda\boldsymbol{\eta}_i + \boldsymbol{\varepsilon}_i$$

Structural Equation

$$\boldsymbol{\eta}_i = \boldsymbol{\alpha} + \Gamma\mathbf{x}_i + \boldsymbol{\zeta}_i$$



LCM: Time-Varying Covariates (TVCs)

- Predictors that take on unique values over time
  - Truly time-specific effects: weather, treatment dosage
  - Contain person-level information too: almost **everything** we do (e.g., depression on stress)

- Enter the model at the manifest/observed variable level
  - TVCs predict time-specific deviations from the underlying trajectory
    - Can include contemporaneous and lagged effects
  - But also conditions trajectory on the TVC so inclusion can change the implied trajectory
    - E.g., from quadratic to linear

- TVCs should **not** be systematically changing over time

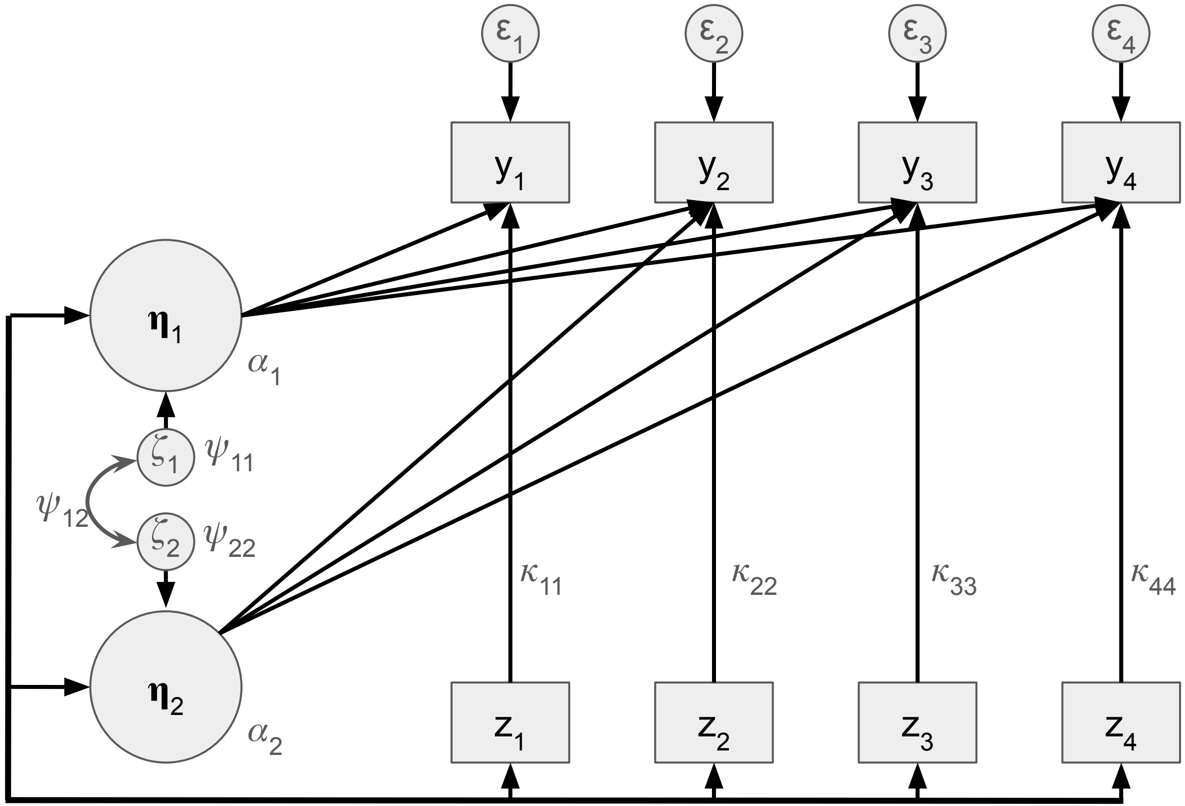
LCM: Time-Varying Covariates

Measurement Equation

$$\mathbf{y}_i = \boldsymbol{\Lambda}\boldsymbol{\eta}_i + \mathbf{K}\mathbf{z}_i + \boldsymbol{\varepsilon}_i$$

Structural Equation

$$\boldsymbol{\eta}_i = \boldsymbol{\alpha} + \boldsymbol{\zeta}_i$$



Latent Curve Model with Covariates: Readings

- Curran, P.J., Bauer, D.J., & Willoughby, M.T. (2004). Testing main effects and interactions in latent curve analysis. *Psychological Methods, 9,* 220-237.
- Bollen, K.A., & Curran, P.J. (2006). *Latent Curve Models: A Structural Equation Approach*. Wiley Series on Probability and Mathematical Statistics. John Wiley & Sons: New Jersey.
- Preacher, K. J., Curran, P. J., & Bauer, D. J. (2006). Computational tools for probing interactions in multiple linear regression, multilevel modeling, and latent curve analysis. *Journal of Educational and Behavioral Statistics*, *31*(4), 437-448.
- Curran, P. J., & Bauer, D. J. (2011). The disaggregation of within-person and between-person effects in longitudinal models of change. *Annual Review of Psychology*, *62*, 583-619.
- Curran, P. J., Lee, T., Howard, A. L., Lane, S., & MacCallum, R. (2012). Disaggregating within-person and between-person effects in multilevel and structural equation growth models. In J. R. Harring & G. R. Hancock (Eds.), *Advances in Longitudinal Methods in the Social and Behavioral Sciences* (pp. 217–253). IAP Information Age Publishing.

Multivariate LCM

- What if we think our TVC is systematically changing over time?
  - Or just contains person-level information
  - We can model a growth process on y and z at the same time

- A raw score change approach to modeling how two or more variables travel together over time
  - Co-development: Hancock & Curran, 2021

Multivariate LCM

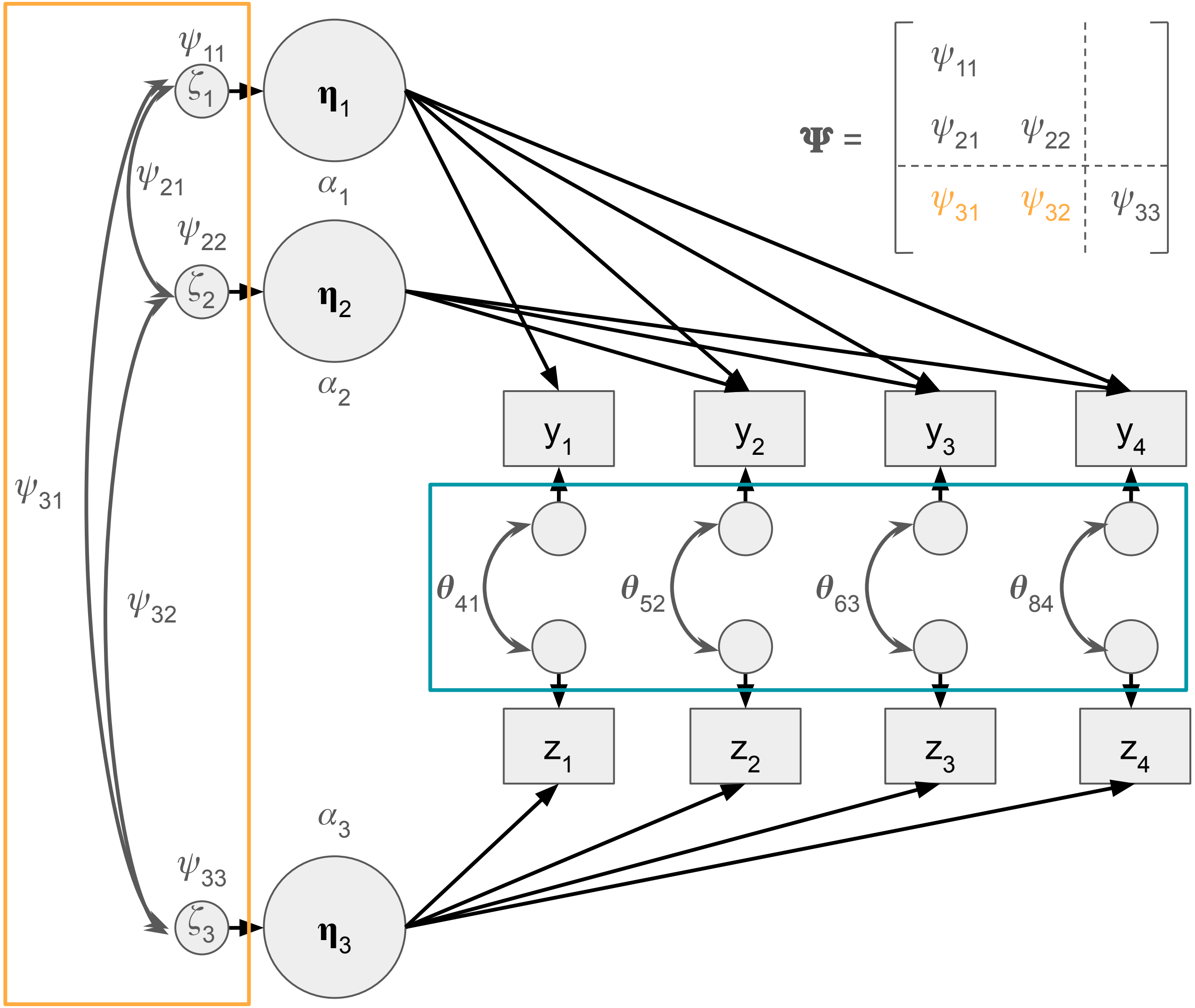
Measurement Equation

$$\begin{matrix} \mathbf{y}_i \\ - \\ \mathbf{z}_i \end{matrix} = \boldsymbol{\Lambda}\boldsymbol{\eta}_i + \boldsymbol{\varepsilon}_i$$

Structural Equation

$$\boldsymbol{\eta}_i = \boldsymbol{\alpha} + \boldsymbol{\zeta}_i$$

$$\boldsymbol{\Theta} = \left[\begin{array}{cccc|cccc} \theta_{y1} & & & & & & & \\ 0 & \theta_{y2} & & & & & & \\ 0 & 0 & \theta_{y3} & & & & & \\ 0 & 0 & 0 & \theta_{y4} & & & & \\ \hline \theta_{41} & 0 & 0 & 0 & \theta_{z1} & & & \\ 0 & \theta_{52} & 0 & 0 & 0 & \theta_{z2} & & \\ 0 & 0 & \theta_{63} & 0 & 0 & 0 & \theta_{z3} & \\ 0 & 0 & 0 & \theta_{84} & 0 & 0 & 0 & \theta_{z4} \end{array}\right]$$



$$\boldsymbol{\Psi} = \left[\begin{array}{cc|c} \psi_{11} & & \\ \psi_{21} & \psi_{22} & \\ \hline \psi_{31} & \psi_{32} & \psi_{33} \end{array}\right]$$

Multivariate LCM: Extensions

- Can bring in TICs to predict growth factors
    - **Could** do cross-construct regression among factors, but I recommend against
- Can have different function forms for each outcome
    - Example was intercept only for z, intercept + linear slope for y
    - Could have higher polynomials
- Can separate within- and between-person effects by structuring residuals

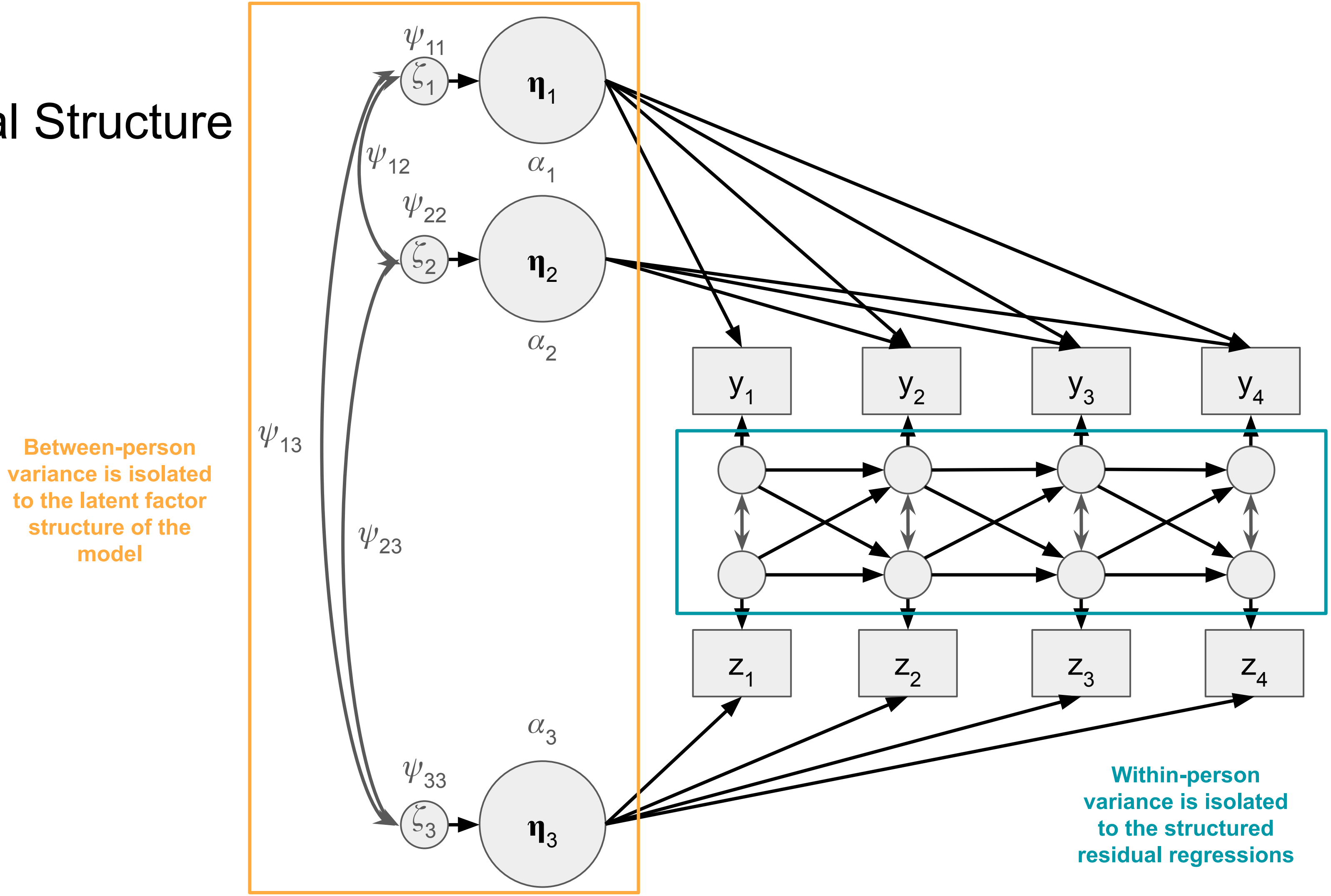
Within- vs. Between-Person Effects

- Most theories in DCN posit within-person processes
  - However often evaluated with between-person statistics
  - Even in longitudinal models (Curran & Bauer, 2011)

- TVCs that systematically vary in level confound within- and between-person effects

- Well developed procedures in mixed-effects models, but relatively rare to see in SEMs (which too be clear, is a **bad** thing)

Residual Structure

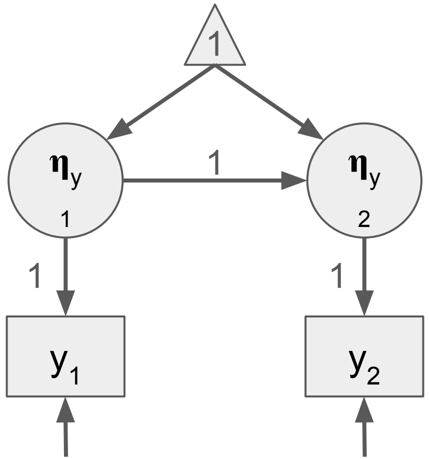


Multivariate Latent Curve Model: Readings

- Curran, P. J., & Hancock, G. R. (2021). The Challenge of Modeling Co-Developmental Processes over Time. *Child Development Perspectives*, *15*(2), 67-75.
- Hertzog, C., Lindenberger, U., Ghisletta, P., & von Oertzen, T. (2006). On the power of multivariate latent growth curve models to detect correlated change. *Psychological Methods*, *11*(3), 244.
- MacCallum, R. C., Kim, C., Malarkey, W. B., & Kiecolt-Glaser, J. K. (1997). Studying multivariate change using multilevel models and latent curve models. *Multivariate Behavioral Research*, *32*(3), 215-253.
- Curran, P.J., Howard, A.L., Bainter, S.A., Lane, S.T., & McGinley J.S. (2013). The separation of between-person and within-person components of individual change over time: A latent curve model with structured residuals. *Journal of Consulting and Clinical Psychology, 82*, 879-894.

Latent Change Score Model: 2 timepoints

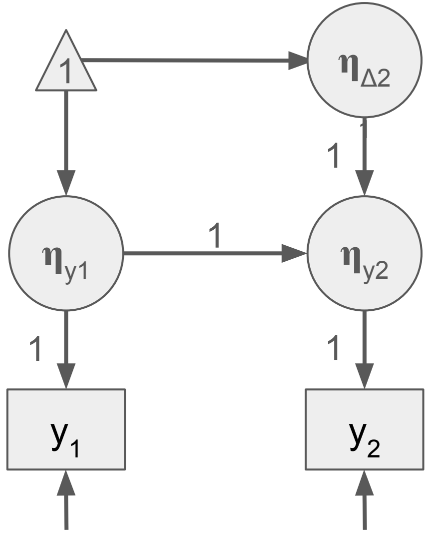
- Can re-express a raw change score as a special case of the autoregressive model where we set the autoregressive parameter to 1
  - Allows us to solve for the residual in the model: $\zeta_i = y_{2i} - (1)^* y_{1i} - \alpha_i = d_i - \alpha_i$
  - Normally don't have access to the residual in the model but using this approach, we can create a phantom variable that we can then use in the model



LCS Model: 2 timepoints

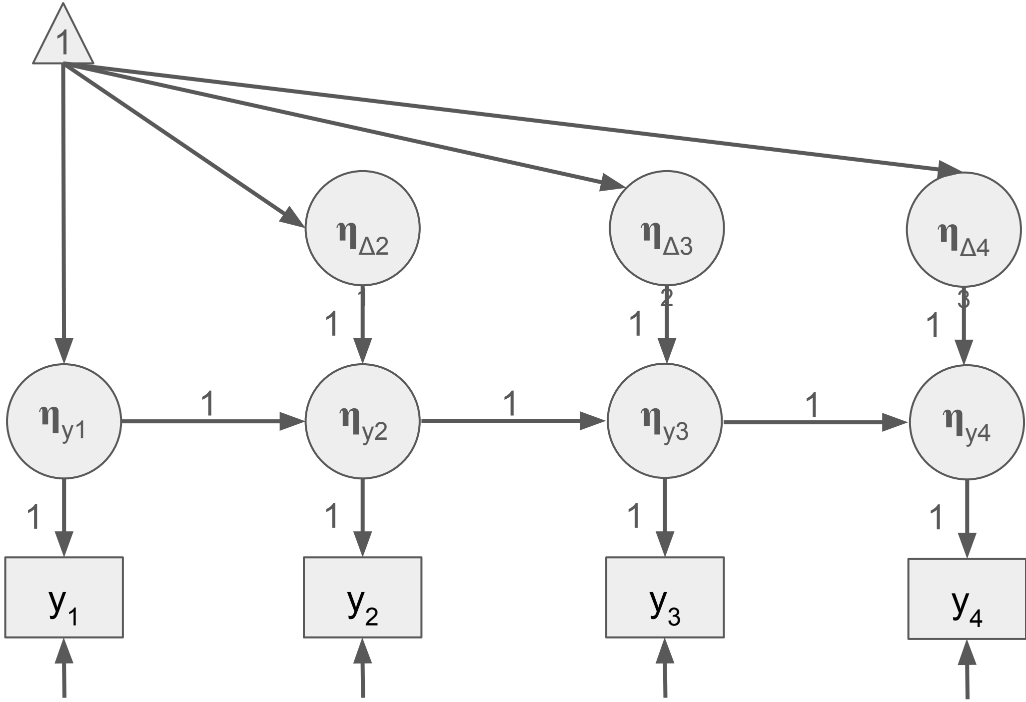
- To put raw change in the model, we need to put an additional latent factor on top of the phantom variables
  - This is not really latent in the way we normally think

- This model is an SEM expression of a paired samples t-test
  - Only difference is it allows missing data
  - Can build on this simple model to do more interesting things



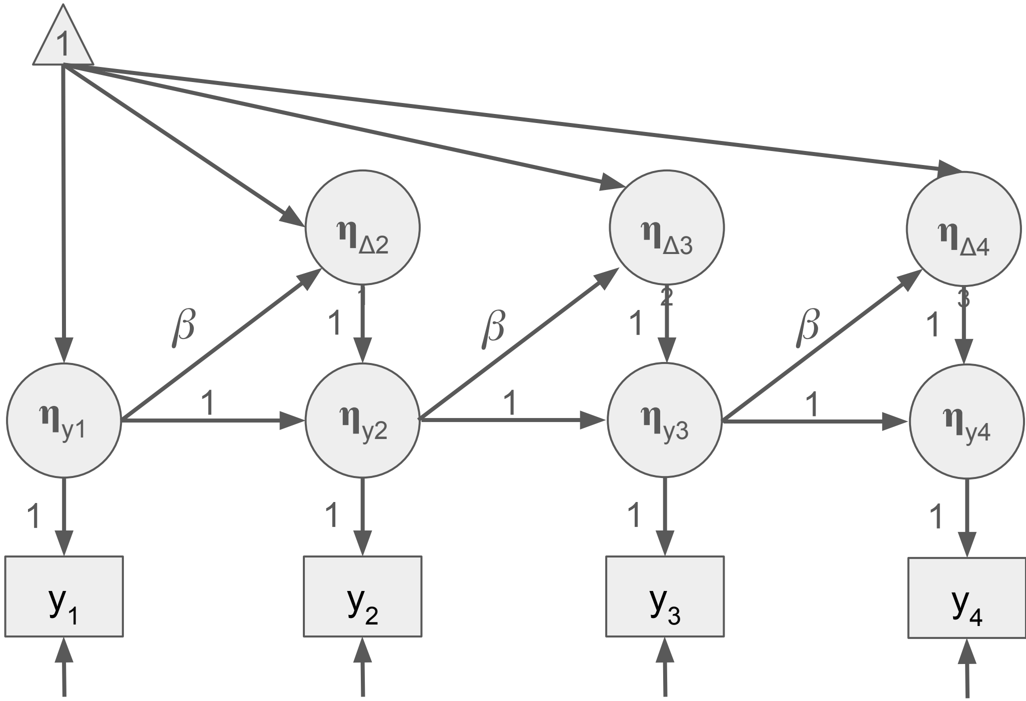
LCS Model: Univariate Latent Change

- The means of the delta latent variables can tell us something about nature of change
  - Close to zero: no change
  - Equal means: constant change
  - Unequal means: discontinuous change
- Can use these factors as predictors of other variables of interest
  - Highly flexible but can also be hard to theoretically interpret



LCS Model: Proportional Change

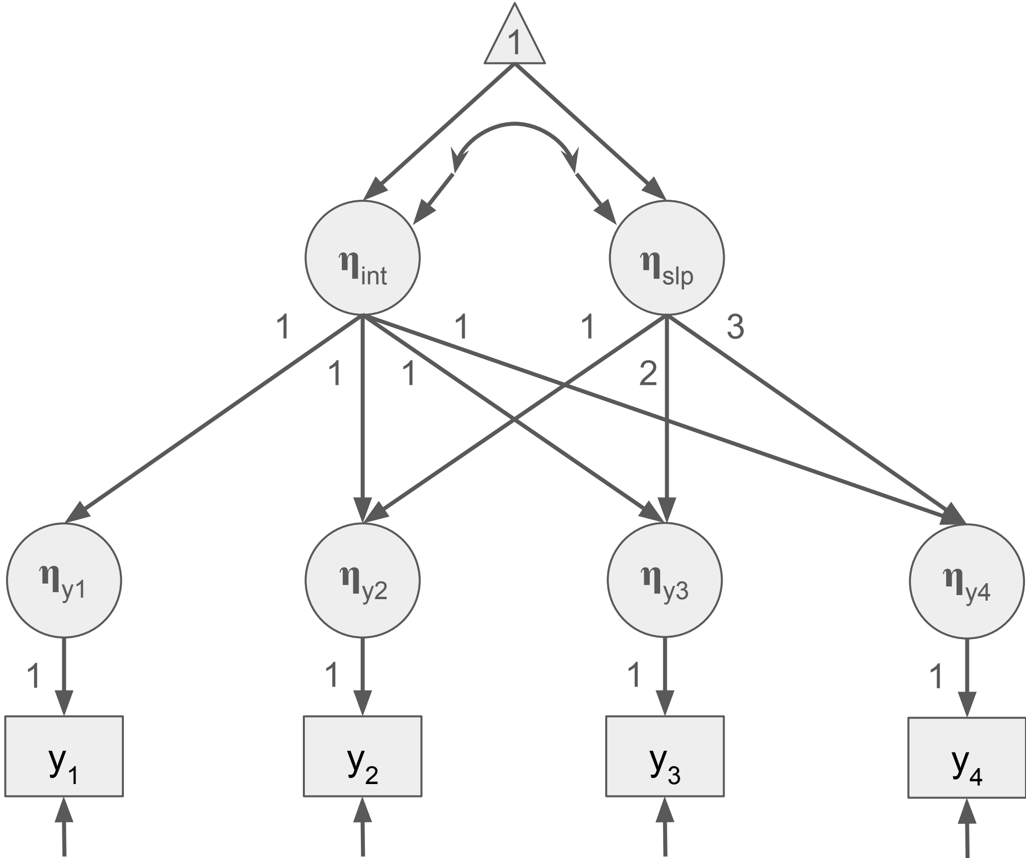
- We can regress latent change ($\boldsymbol{\eta}_{\Delta t(t-1)}$) on prior level ($\boldsymbol{\eta}_{yt}$)
    - Implies that the amount of change that occurs is in-part determined by prior level
- Typically β is set equal across time but not necessary
- Actually identical to the ARCL model with AR(1) lags
    - These initial LCS models are really just re-parameterizing models we are more familiar with



LCS Model: Trajectories

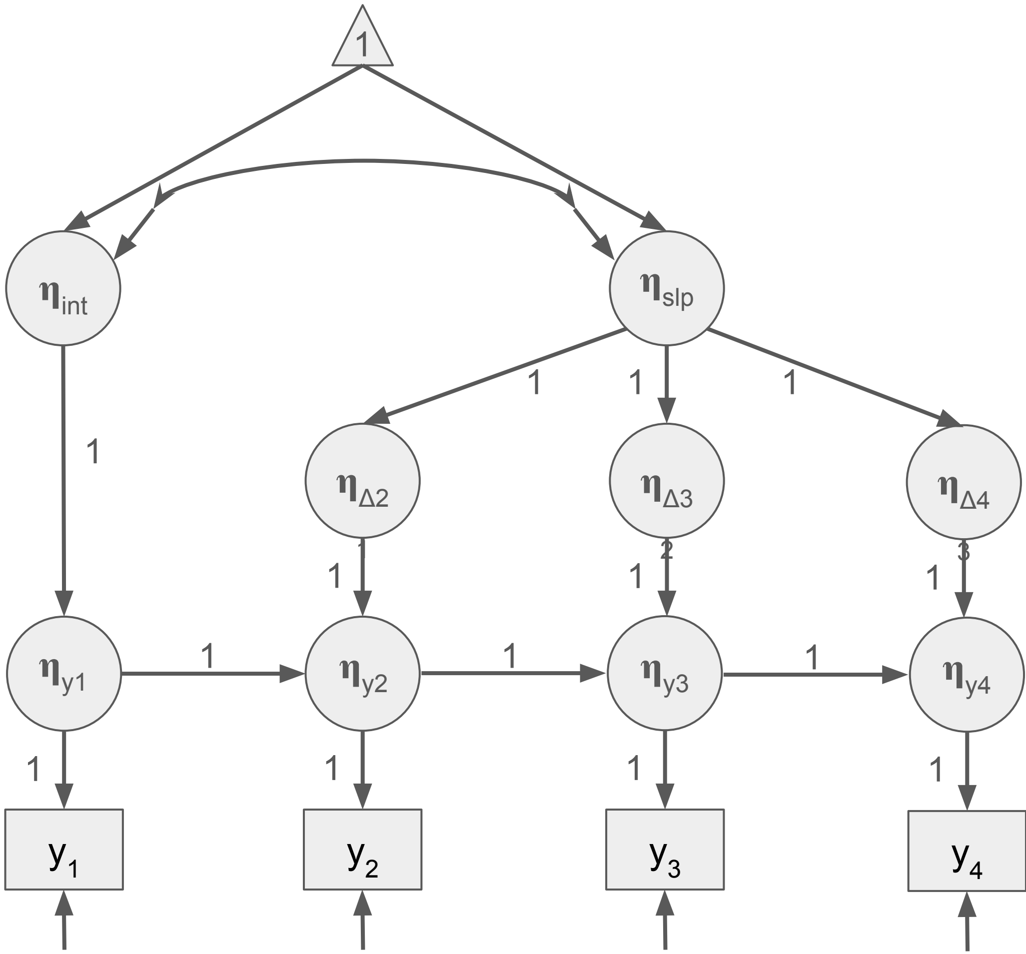
- Spoiler alert: we can also re-formulate the LCM as an LCS
  - But we'll see unique LCS features after this

- Basically just push y into phantom variables
  - Phantoms don't add model complexity
  - One more reformulation before we can get to the good stuff



LCS Model: Trajectories

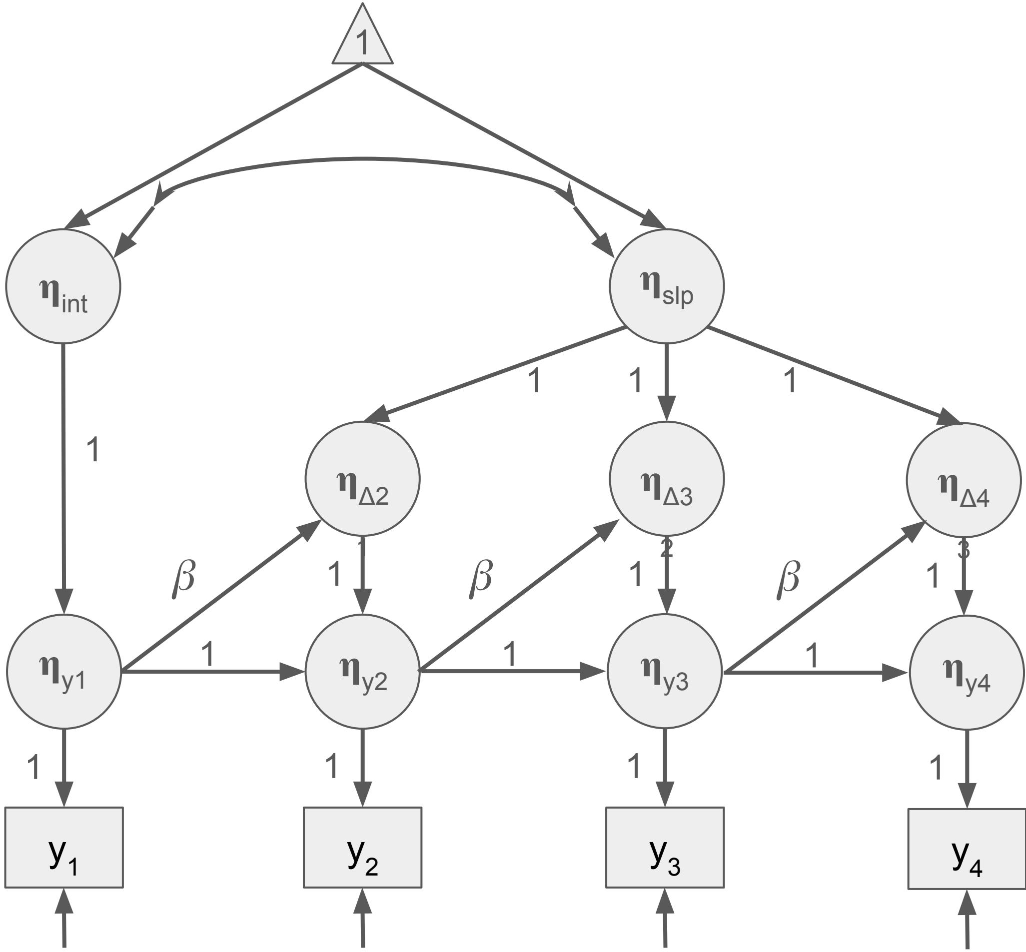
- We can also use the $\boldsymbol{\eta}_{\Delta}$ formulation
  - Intercept is now defined on only a single phantom variable
  - But note that the slope loadings are now **all** 1
    - Essentially sums across discrete change (which is what a smooth trajectory should do)
- This is the formulation that allows for something we haven't seen before



LCS Model: Trajectories

- Adding the proportional change parameter is a unique way to model non-linearities
  - Especially good for exponential trends
  - But cannot fit random effects here

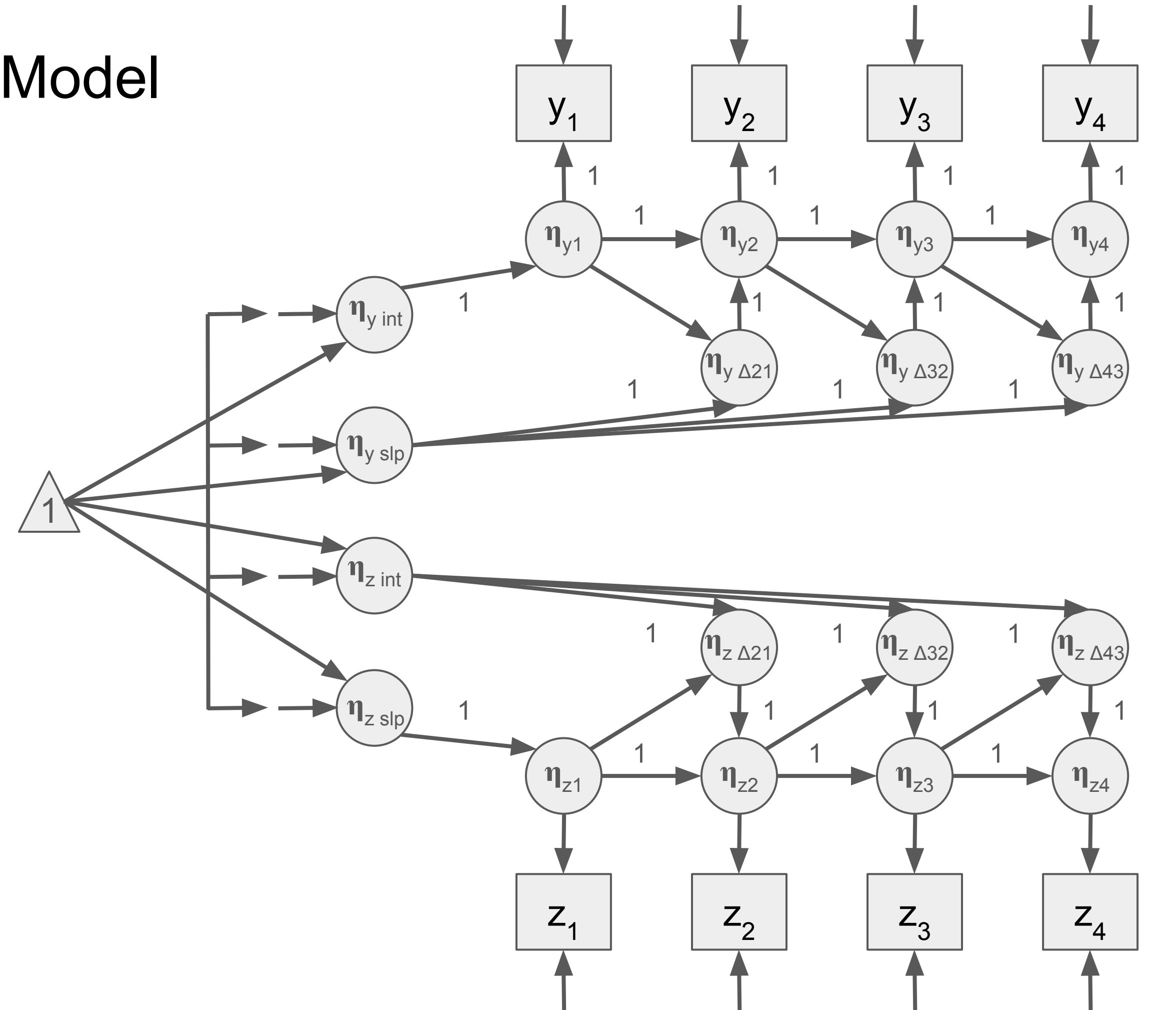
- An interesting enough model, but the really unique case is when we move to the bivariate model with proportional change



Multivariate Dual-Change Model

- Here we model non-linear changes **within** each construct separately

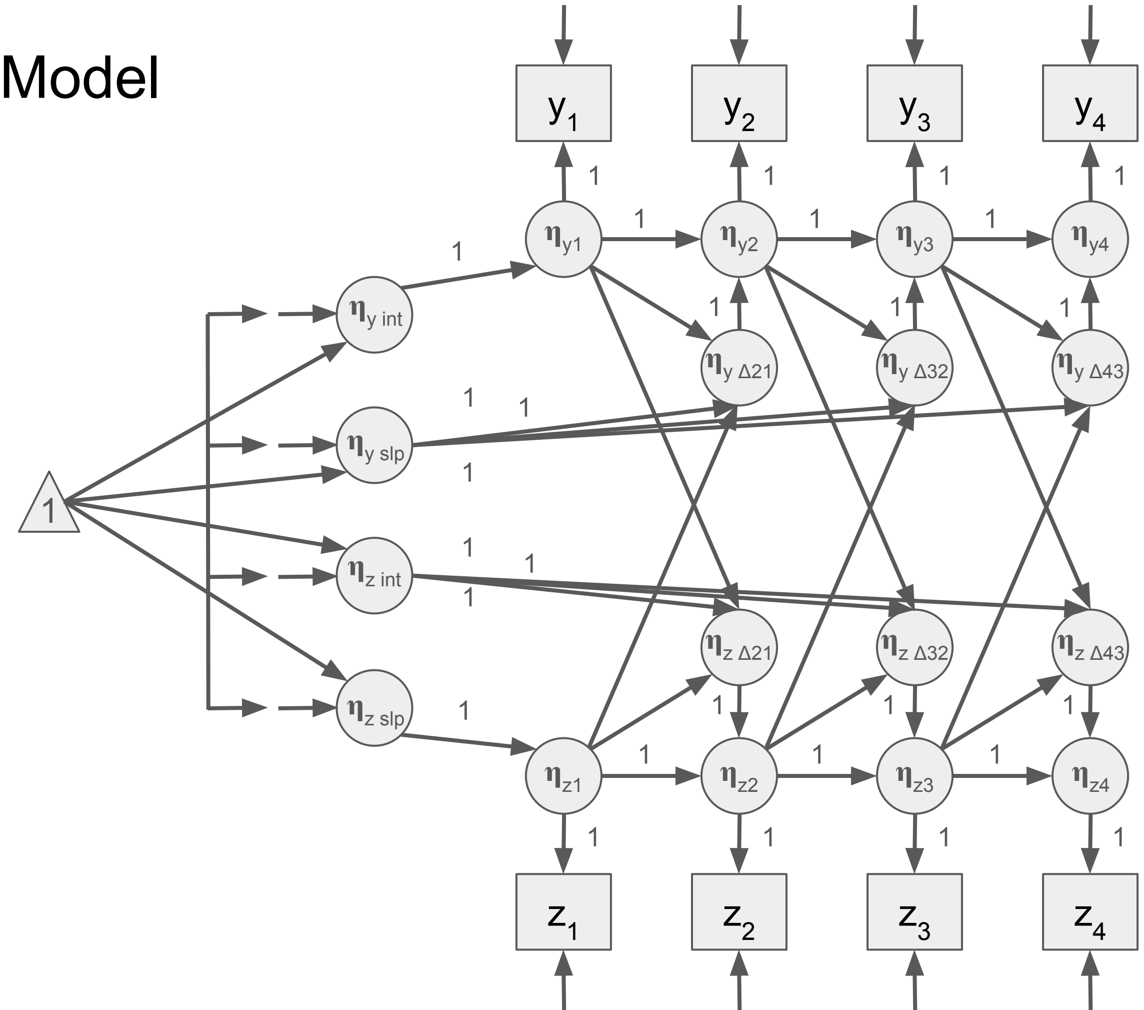
- We get cross-construct correlations at the latent level like the multivariate LCM



Multivariate Dual-Change Model

- But can use cross-construct regressions to look at proportional linkage across variables

- Incredibly flexible and powerful model
    - What theory does this model describe/test?



Latent Change Score Model: Readings

- Grimm, K. J., An, Y., McArdle, J. J., Zonderman, A. B., & Resnick, S. M. (2012). Recent changes leading to subsequent changes: Extensions of multivariate latent difference score models. *Structural Equation Modeling: A Multidisciplinary Journal*, *19*, 268-292.
- Butner, J. E., Berg, C. A., Baucom, B. R., & Wiebe, D. J. (2014). Modeling coordination in multiple simultaneous latent change scores. *Multivariate Behavioral Research*, *49*(6), 554-570.
- McArdle, J. J., & Grimm, K. J. (2010). Five steps in latent curve and latent change score modeling with longitudinal data. In *Longitudinal Research with Latent Variables* (pp. 245-273). Springer, Berlin, Heidelberg.
- Ghisletta, P., & McArdle, J. J. (2012). Latent curve models and latent change score models estimated in R. *Structural Equation Modeling: A Multidisciplinary Journal*, *19*(4), 651-682.
- Ferrer, E., & McArdle, J. J. (2010). Longitudinal modeling of developmental changes in psychological research. *Current Directions in Psychological Science*, *19*(3), 149-154.

Where do we go from here?