# A Visual Tour of Extended Structural Equations & State Space Models

Michael D. Hunter

Developmental and Behavioral Pediatrics
University of Oklahoma Health Sciences Center

Statistical Computing User Group
June 3, 2014

### Cross-Cutting Theme Program

[Big Data: Understanding Patterns of Human Behavior]

**Michael N. Jones**, *Indiana University Bloomington*
**Tanzeem Choudhury**, *Cornell University*
**Brian M. D'Onofrio**, *Indiana University Bloomington*
**Susan T. Dumais**, *Microsoft Research*
**Tal Yarkoni**, *University of Texas*


### Invited Symposia

[Data Integration in a Big Data Environment to Understand Human Behavior]

**Allison Ottenbacher** (Chair), *National Cancer Institute*,
**Keith Widaman** (Discussant), *University of California, Davis*
**Damon Centola**, *University of Pennsylvania*
**Mark Cummings**, *San Diego State University and Orchestral Networks*
**John Ayers**, *San Diego State University*

[Using Big Data to Advance Psychological Theory: Report from the NCI "Big Data for Theory Advancement"]

**William Klein** (Discussant), *National Cancer Institute*
**Donna Spruijt-Metz**, *University of Southern California*
**Genevieve Dunton**, *University of Southern California*
**Noshir Contractor**, *Northwestern University*



### Workshop

[Big Data: Promises and Challenges]

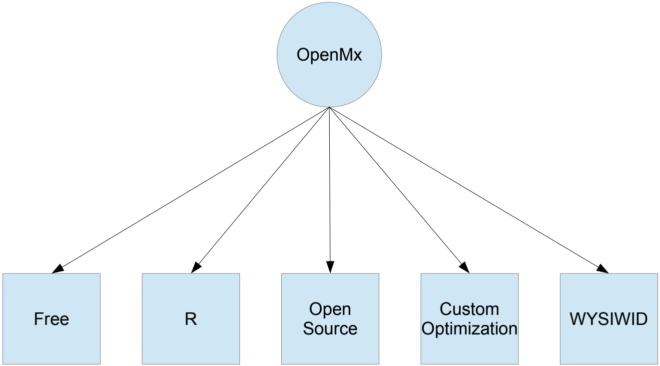
**Richard D. Gonzalez**, *University of Michigan, Ann Arbor*

What are you going to do with Big Data?

# Outline

- Extended Structural Equation Models
- State Space Models
- Visual Tour of Extensions of State Space Models

# OpenMx



# SEM

- Model specifies a distribution
- Distribution is often Gaussian/Normal
  - Mean
  - Covariance
- Maximum Likelihood

# Extended SEM

## Sold Separately

- Ordinal Variables
- Alternative (User Defined) Fit Functions
- Other Distributions (Non-Normal, Mixtures)
  - Non-Normal (Poisson)
  - Mixtures
  - Categorical Latent Variables
- Different modeling specifications
- Multiple Groups
- SEM Trees
- Alternative Optimization Routines

# Extended SEM

## Sold Separately

- Ordinal Variables
- Alternative (User Defined) Fit Functions
- Other Distributions (Non-Normal, Mixtures)
  - Non-Normal (Poisson)
  - Mixtures
  - Categorical Latent Variables
- Different modeling specifications
- Multiple Groups
- SEM Trees
- Alternative Optimization Routines
- How?

# Expectations

## Some Assembly Required

- Four Steps
  - Expectation: Create a model that has implications for your data
  - Fit: Compare model expectations to your actual data
  - Optimize: Adjust movable parts of model expectation to make fit as good as possible
  - Profit
- Models expect the data to match
- Usually expected covariance and means
- Fit: ML, FIML, WLS

# State Space Models

- Dynamic Factor Analysis
- Latent Time Series
- No Stationarity Required!
- Modular implementation (Uses same code for FIML and missing data as RAM/LISREL. Modularity makes life easier.)
- Understand the unknown by the known . . .

# Mplus Model

- Structural Model

$$\eta_i = B\eta_i + \Gamma x_i + \zeta_i \qquad \zeta_i \sim \mathcal{N}\left(\vec{0}, \Psi\right) \tag{1}$$

- Measurement Model

$$y_i = \Lambda\eta_i + K x_i + \varepsilon_i \qquad \varepsilon_i \sim \mathcal{N}\left(\vec{0}, \Theta\right) \tag{2}$$

# State Space Model

- Structural Model

$$\eta_{i+1} = B\eta_i + \Gamma x_i + \zeta_i \qquad \zeta_i \sim \mathcal{N}\left(\vec{0}, \Psi\right) \tag{3}$$

- Measurement Model

$$y_i = \Lambda\eta_i + K x_i + \varepsilon_i \qquad \varepsilon_i \sim \mathcal{N}\left(\vec{0}, \Theta\right) \tag{4}$$

# State Space Model

## No Greek

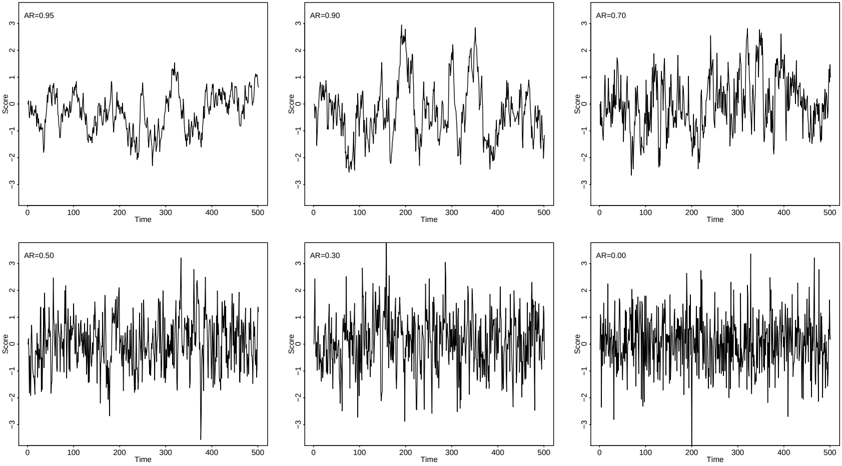
- Structural Model

$$x_{i+1} = Ax_i + Bu_i + q_i \qquad q_i \sim \mathcal{N}\left(\vec{0}, Q\right) \tag{5}$$

- Measurement Model

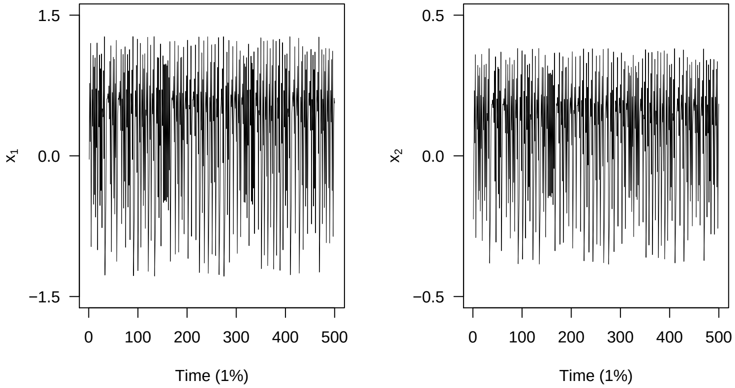
$$y_i = Cx_i + Du_i + r_i \qquad r_i \sim \mathcal{N}\left(\vec{0}, R\right) \tag{6}$$

# From Autoregression to White Noise



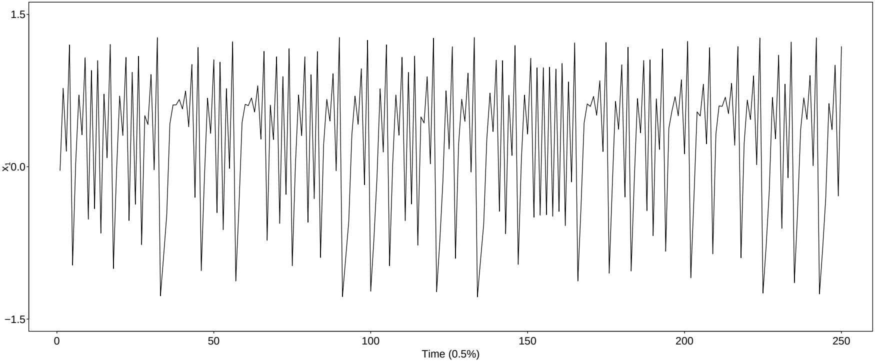
# Variable/Time Space

## Hénon Map



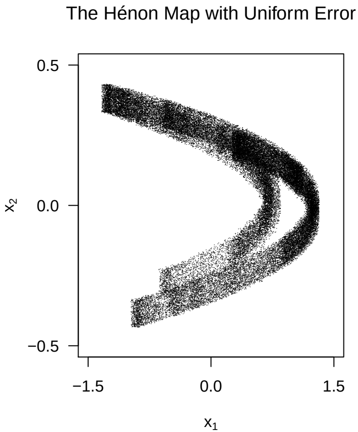
# Variable/Time Space

## Hénon Map



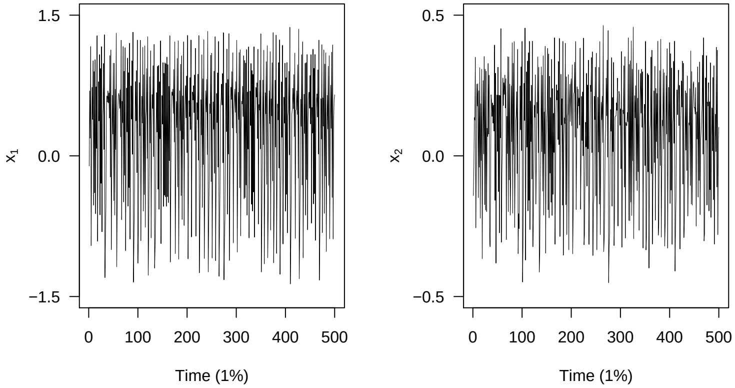
# State Space

## Hénon Map with Uniform Error



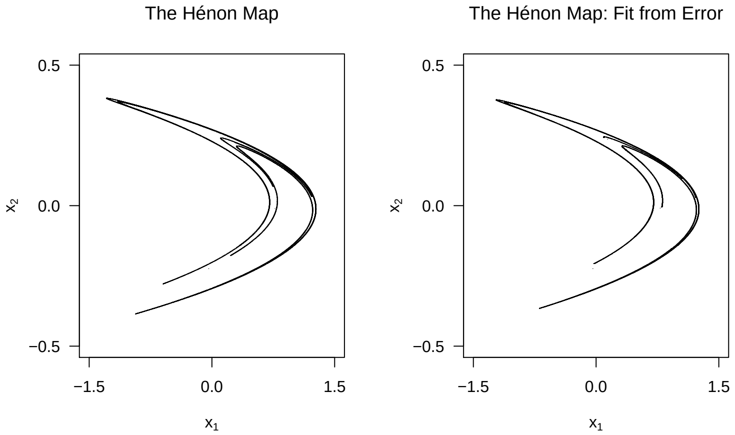
# Variable/Time Space

## Hénon Map with Uniform Error

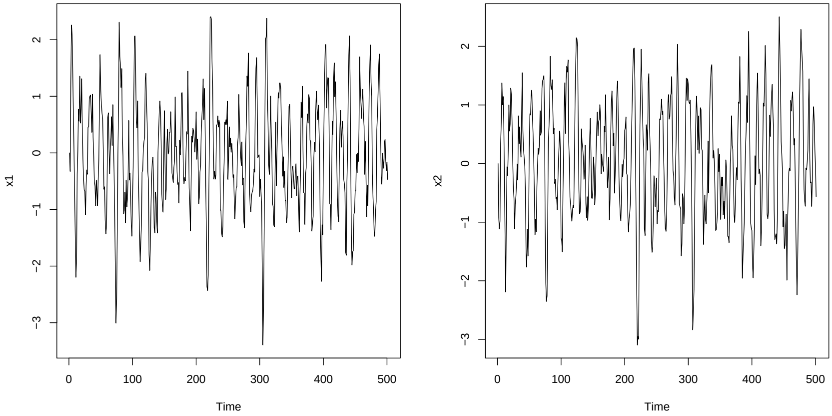


# State Space

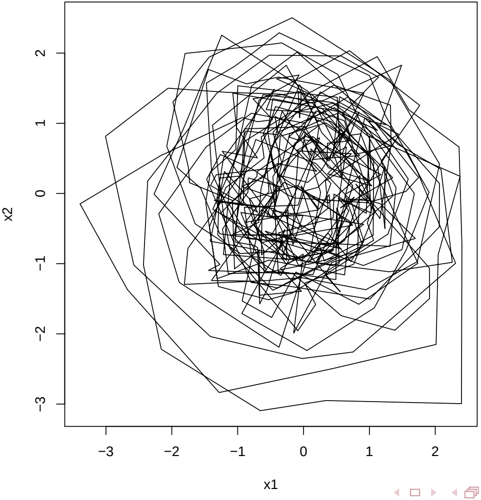
## Hénon Map



# Variable/Time Space

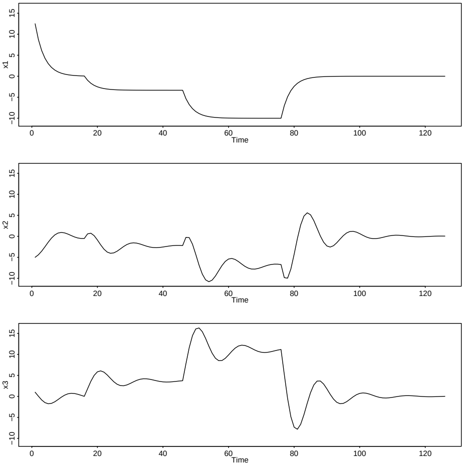


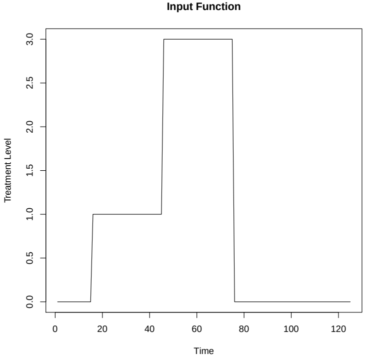
# State Space



# Inputs, Covariates, or Known Shocks

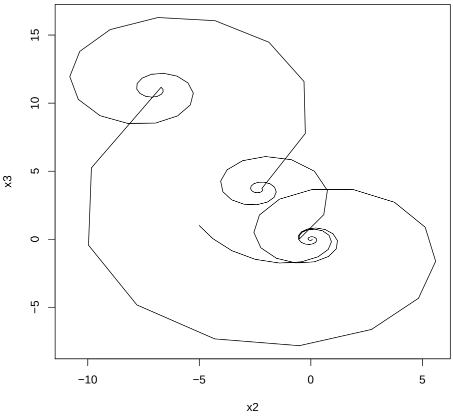
## Experimental



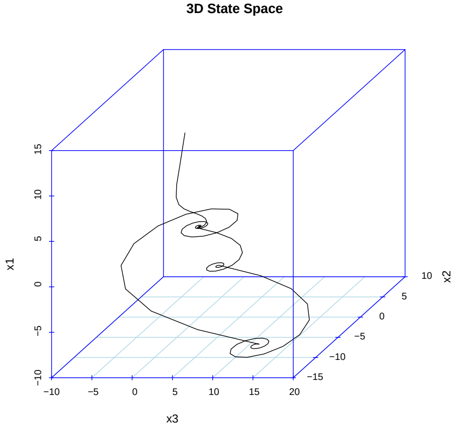
# Inputs, Covariates, or Known Shocks

## Experimental



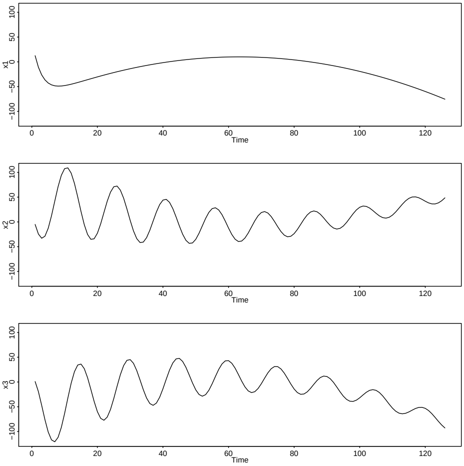
# Inputs, Covariates, or Known Shocks

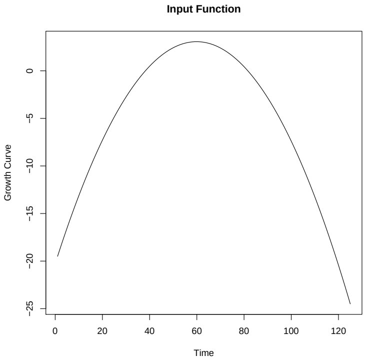
## Experimental



# Inputs, Covariates, or Known Shocks

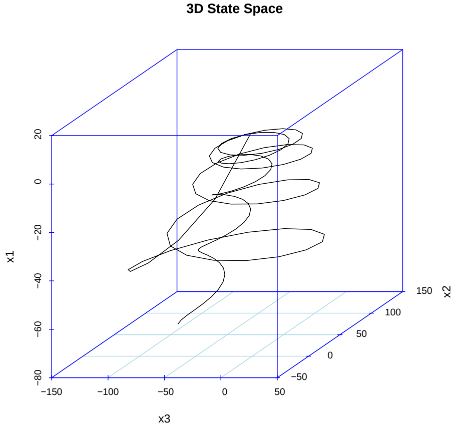
## Growth



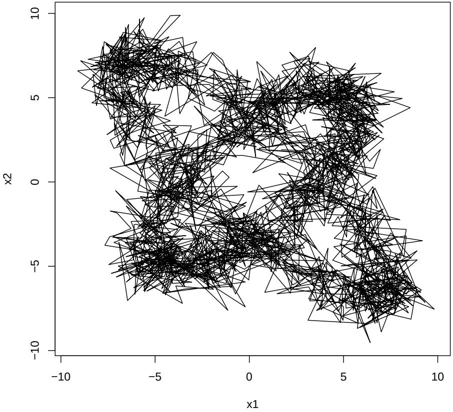



# Inputs, Covariates, or Known Shocks

## Growth

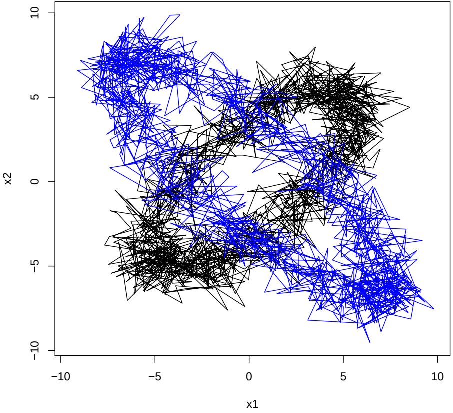


# Multigroup State Space



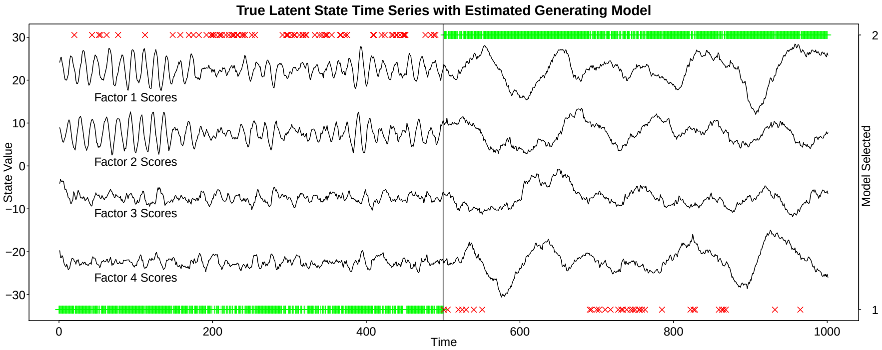
# Multigroup State Space

## $N \neq 1 \ldots$ Novel!

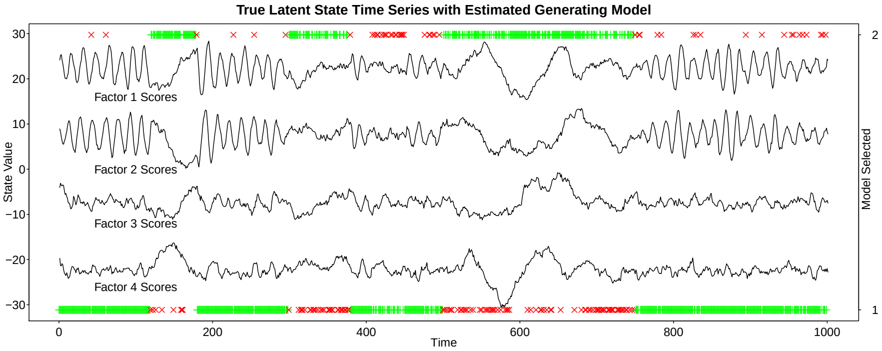


# State Space Mixture with Single Switch

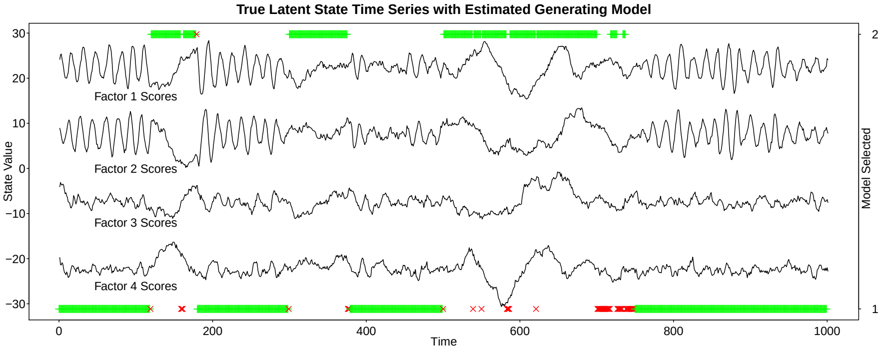
## Hunter (2014) SMEP/MBR



# State Space Mixture with Several Switches



# State Space Reqime Switching



# Summary

- What will you do with big data?
- Expectations and Fits
- State Space models are not hugely different, but are importantly so.
- Visualization/modeling in state space can make simple patterns evident.
- SEM, Covariates, Multigroup, Mixtures, and State Space all in the same program, same interface, allow combinations.

# Future Work

- More fit functions: WLS
- Item Factor Analysis (Joshua Pritikin)
- CRAN

# Future Work

## Multilevel

*Multivariate Behavioral Research*, 49:119–129, 2014

ISSN: 0027-3171 print / 1532-7906 online
DOI: 10.1080/00273171.2013.866537

Routledge
Taylor & Francis Group

### A Computationally Efficient State Space Approach to Estimating Multilevel Regression Models and Multilevel Confirmatory Factor Models

Fei Gu
*Department of Psychology, McGill University*

Kristopher J. Preacher
*Department of Psychology and Human Development, Vanderbilt University*

Wei Wu
*Department of Psychology, University of Kansas*

Yiu-Fai Yung
*SAS Institute Inc.*

# Multilevel Random Intercepts Factor Analysis

128 GU, PREACHER, WU, YUNG

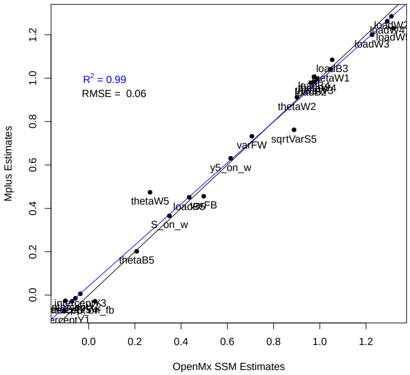
TABLE 5
Results of the Random Intercepts Confirmatory Factor Model

| | Mplus | | SSM-IML | |
|---|---|---|---|---|
| Parameter | Point | *SE* | Point | *SE* |
| $\gamma 1_{00}$ | 3.677 | .073 | 3.677 | .073 |
| $\gamma 2_{00}$ | 3.731 | .082 | 3.731 | .082 |
| $\gamma 3_{00}$ | 3.706 | .080 | 3.706 | .080 |
| $\lambda_{W2}$ | .978 | .040 | .979 | .040 |
| $\lambda_{W3}$ | 1.070 | .044 | 1.071 | .044 |
| $\lambda_{B2}$ | 1.039 | .234 | 1.037 | .233 |
| $\lambda_{B3}$ | 1.091 | .228 | 1.081 | .225 |
| $\psi_W$ | 1.248 | .074 | 1.247 | .074 |
| $\psi_B$ | .107 | .042 | .108 | .042 |
| $\theta_{W1}$ | 1.769 | .063 | 1.769 | .063 |
| $\theta_{W2}$ | 1.868 | .063 | 1.868 | .063 |
| $\theta_{W3}$ | 1.510 | .065 | 1.510 | .065 |
| $\theta_{B1}$ | .025 | .019 | .025 | .019 |
| $\theta_{B2}$ | .058 | .026 | .058 | .026 |
| $\theta_{B3}$ | .037 | .023 | .038 | .023 |
| –2*loglik | 42566.158 | | 42566.158 | |

```
free parameters:
   name       matrix row col   Estimate  Std.Error
1  g100     indiv1.D   1   1 3.67667459 0.07294239
2  g200     indiv1.D   2   1 3.73057530 0.08202516
3  g300     indiv1.D   3   1 3.70614681 0.07990545
4   lw2 indiv1.LamW   2   1 0.97810518 0.03970694
5   lw3 indiv1.LamW   3   1 1.07046380 0.04419464
6   psw indiv1.PsiW   1   1 1.24770679 0.07383896
7  thw1 indiv1.ThdW   1   1 1.76909329 0.06267951
8  thw2 indiv1.ThdW   2   2 1.86846555 0.06271462
9  thw3 indiv1.ThdW   3   3 1.51013218 0.06461675
10  lb2 indiv1.LamB   2   1 1.03845610 0.23261330
11  lb3 indiv1.LamB   3   1 1.09057000 0.22704118
12  psb indiv1.PsiB   1   1 0.10671548 0.04197646
13 thb1 indiv1.ThdB   1   1 0.02494997 0.01895396
14 thb2 indiv1.ThdB   2   2 0.05801682 0.02577111
15 thb3 indiv1.ThdB   3   3 0.03706296 0.02279143

observed statistics:  0
estimated parameters:  15
degrees of freedom:  -15
-2 log likelihood:  42566.16
saturated -2 log likelihood:  -3
number of observations:  3750
```

# Multilevel RandomSlopes & Intercepts Factor Analysis



# Acknowledgments


- Advisors: Joseph L. Rodgers & Hairong Song
- OpenMx Development Team: Steven Boker, Michael Neale, Hermine Maes, Timothy Brick, Ryne Estabrook, Timo von Oertzen, Joshua Pritikin, Timothy Bates, and many more!

# Thank You

mhunter@ou.edu

# Linear Growth

## Explicit and Recursive

- Explicit

- Recursive

# Linear Growth

## Explicit and Recursive

- Explicit
- Linear

$$y_i = b_0 + b_1 t_i + e_i \tag{7}$$

- Recursive
- Linear

$$y_{t+1} = y_t + b_1 + e_t \quad \text{with} \quad y_0 = b_0 \tag{9}$$

# Linear Growth

## Explicit and Recursive

- Explicit
- Linear

$$y_i = b_0 + b_1 t_i + e_i \tag{7}$$

- Quadratic

$$y_i = b_0 + b_1 t_i + b_2 t_i^2 + e_i \tag{8}$$

- Recursive
- Linear

$$y_{t+1} = y_t + b_1 + e_t \quad \text{with} \quad y_0 = b_0 \tag{9}$$

- Quadratic

$$y_{t+1} = y_t + b_1 + b_2 + 2b_2 t + e_t \quad \text{with} \quad y_0 = b_0 \tag{10}$$

# Polynomial Growth

## Explicit and Recursive

- Explicit
- 

$$y_i = b_0 + b_1 t_i + e_i \tag{11}$$

$$y_i = b_0 + b_1 t_i + b_2 t_i^2 + e_i \tag{12}$$

$$y_i = b_0 + b_1 t_i + b_2 t_i^2 + b_3 t_i^3 + e_i \tag{13}$$

- Recursive
- 

$$y_{t+1} = y_t + b_1 + e_t \quad \text{with} \quad y_0 = b_0 \tag{14}$$

$$y_{t+1} = y_t + b_1 + b_2 + 2b_2 t + e_t \quad \text{with} \quad y_0 = b_0 \tag{15}$$

$$y_{t+1} = y_t + b_1 + b_2 + b_3 + 2b_2 t + 2b_3 t + 3b_3 t^2 + e_t \quad \text{with} \quad y_0 = b_0 \tag{16}$$