Deriving models of change with interpretable parameters: linear estimation with nonlinear inference

Psicostats – December 2023

Assistant Professor, Methodology and Statistics
Leiden University

@E_M_McCormick
e.m.mccormick@fsw.leidenuniv.nl
https://e-m-mccormick.github.io/

Manuscript + Package

Deriving models of change with interpretable parameters: linear estimation with nonlinear inference

Ethan M. McCormick*[1]

[1]Methodology & Statistics Department, Institute of Psychology, Leiden University, Leiden, Netherlands

December 10, 2(

Abstract

In the modeling of change over time, there is
opmental theories advanced in substantive research
longitudinal analysis. That is, theory is understood

https://osf.io/preprints/

leni: Linear Estimation with Nonlinear Inference

Ethan M. McCormick | Department of Methodology & Statistics | Leiden University

Paper Citation: McCormick (2023). Deriving models of change with interpretable parameters: linear estimation with nonlinear inference. *PsyArixv*. https://doi.org/10.31234/osf.io/r4vxb

The `leni` package provides a collection of tools for Linear Estimation with Nonlinear Inference (LENI) which involves estimating the parameters of nonlinear target functions as transformations of linear models. Includes support for regression and mixed-effects models (main function: `leni()`) which produces transformed parameter estimates, as well as structural equation models (main function: `leni_sem()`) which generates `lavaan` syntax (https://lavaan.ugent.be/) for fitting linearized SEMs.

Installation

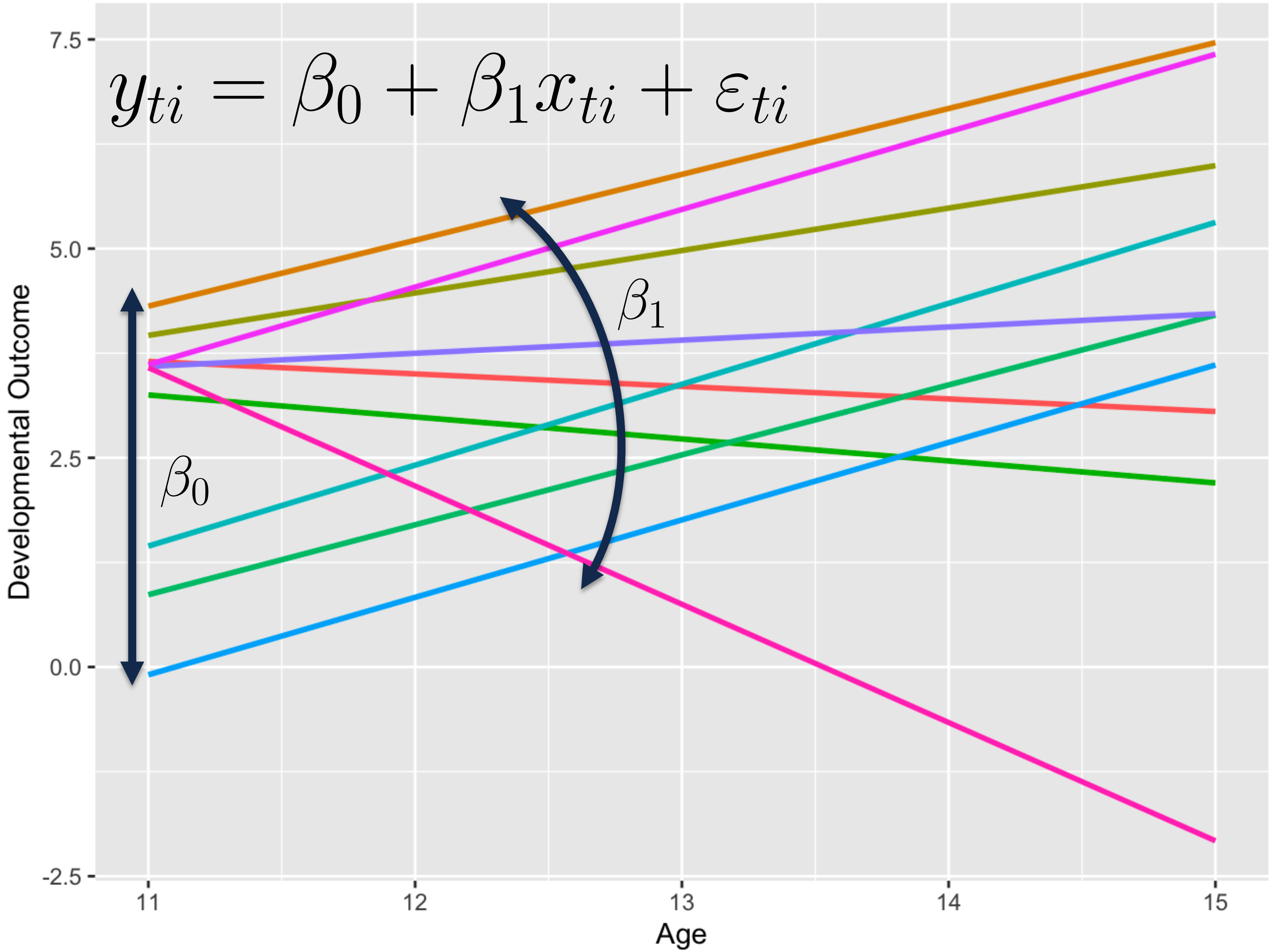
Set up for CRAN installation of `leni` is underway.

In the meantime, to install the latest development version, you can install `leni` directly from github using the `devtools` package.

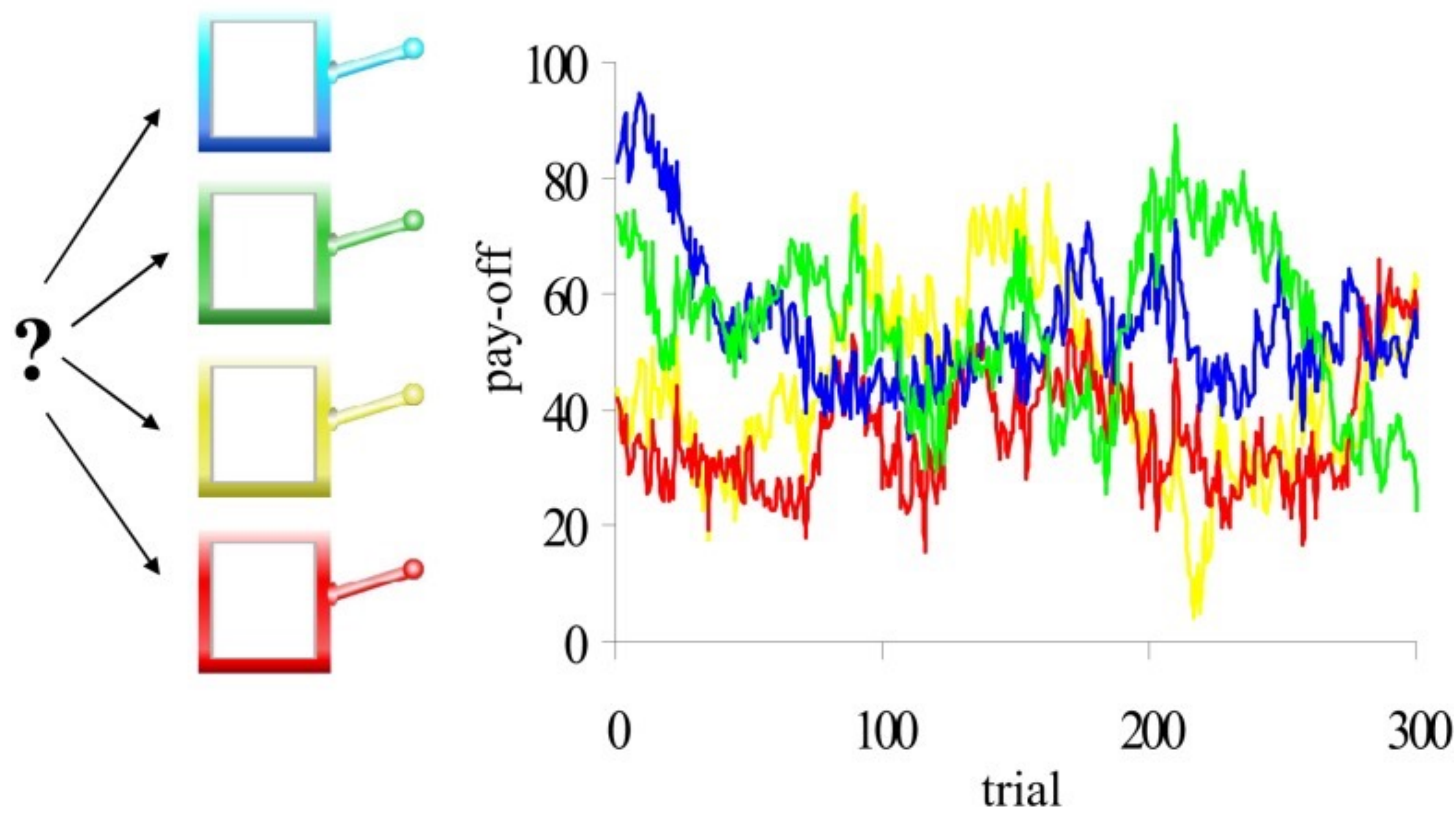
```
#install.packages("devtools")
library(devtools)

install_github("E-M-McCormick/leni")
```

https://github.com/E-M-McCormick/leni

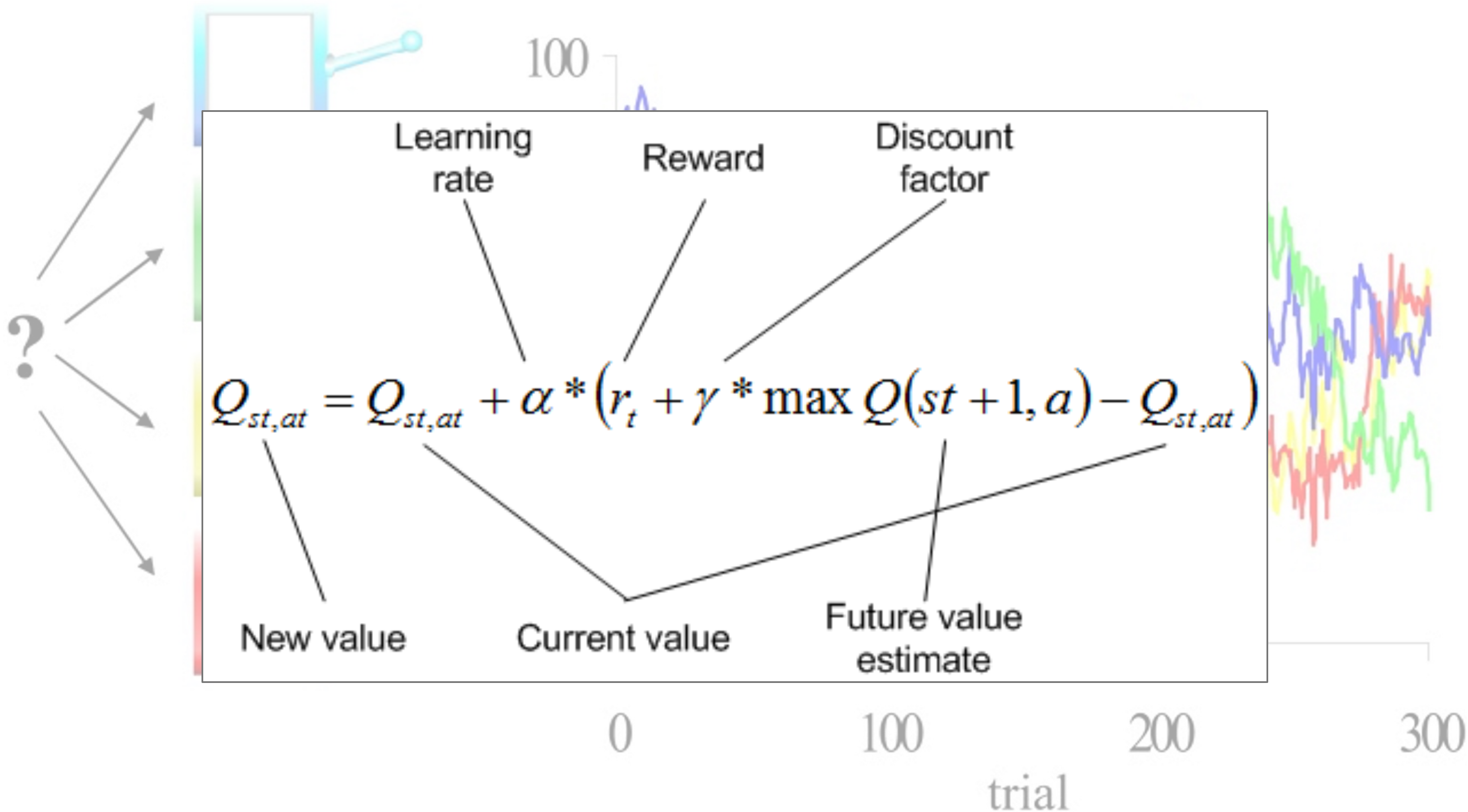
The case for meaningful parameters



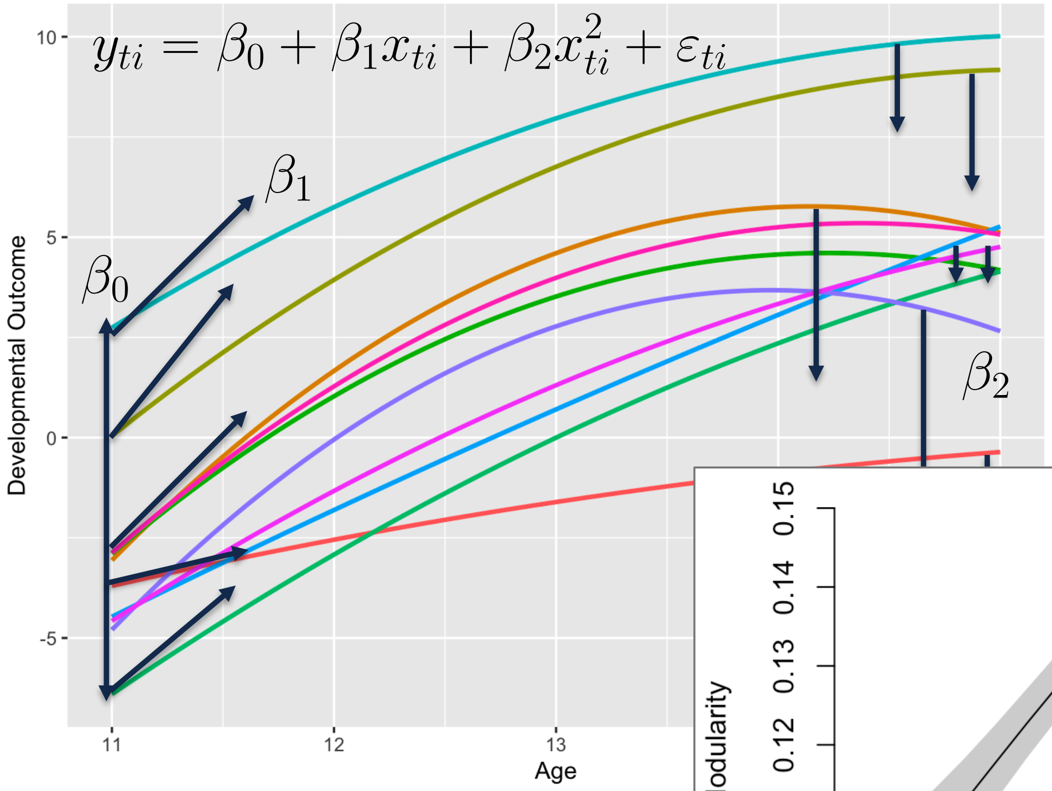
The case for meaningful parameters

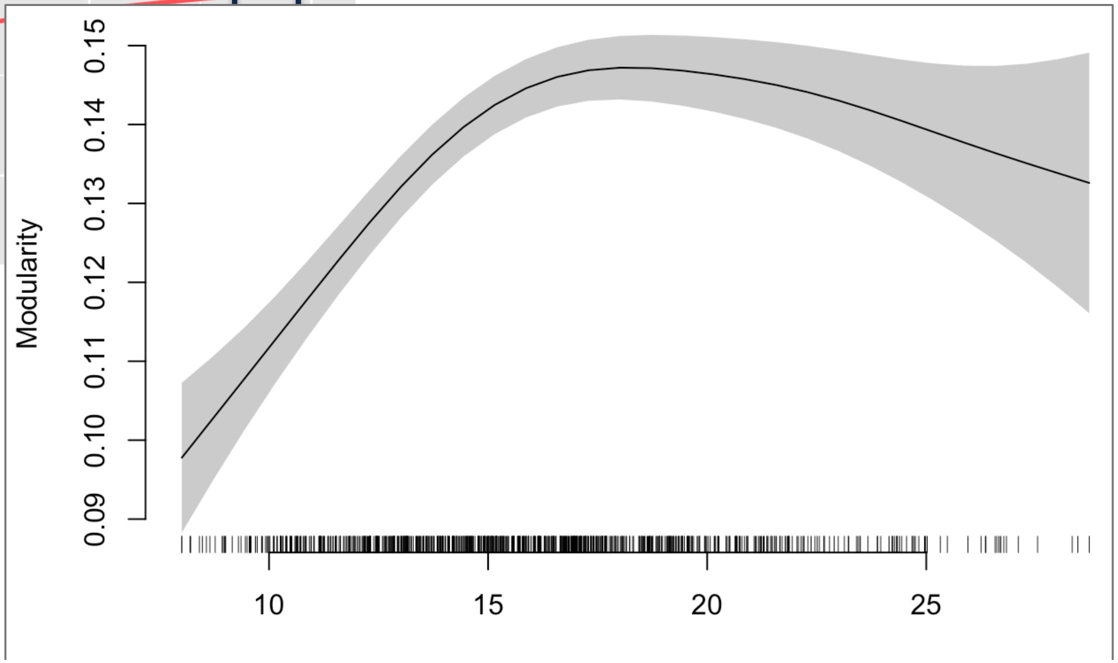


The case for meaningful parameters



The case for meaningful parameters

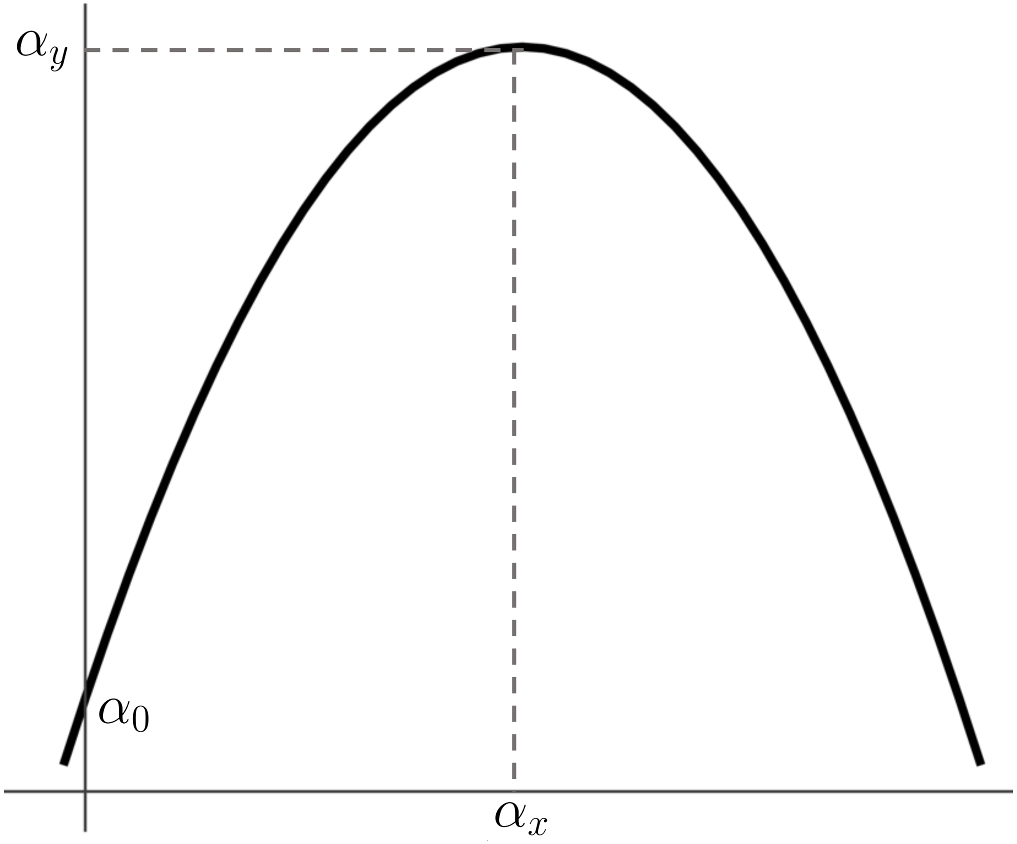




Why do we need interpretable parameters?

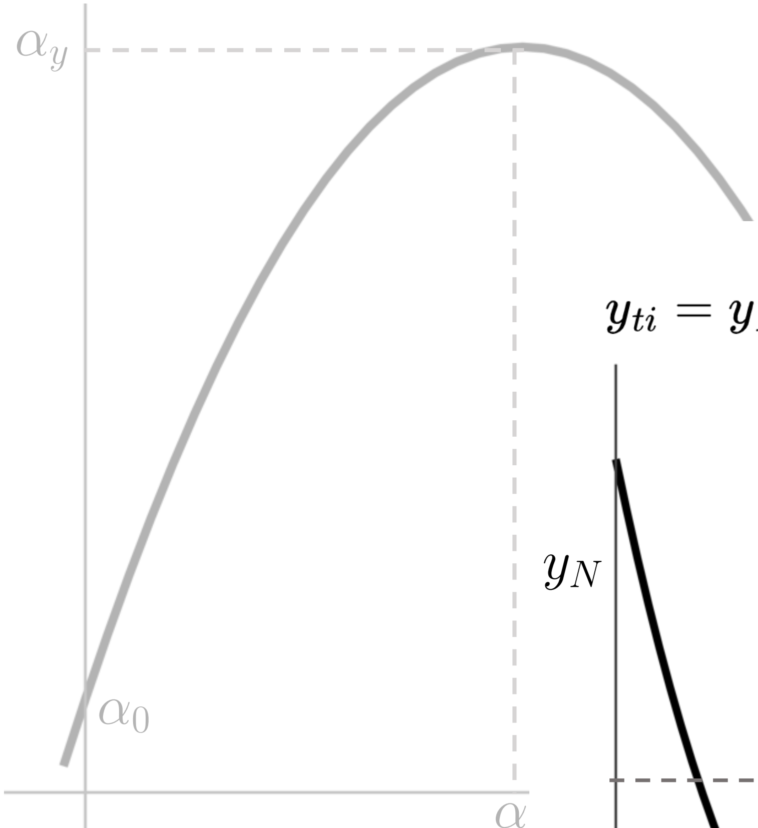
- Testing meaningful – and specifically articulated – theoretical hypotheses about change over time
  - Timing of inflections (e.g., peaks/troughs/plateaus), time-to-criterion, tempo of change
- Incorporating predictors of change
- Investigating distal outcomes associated with individual differences in change over time

Defining Nonlinear Equations



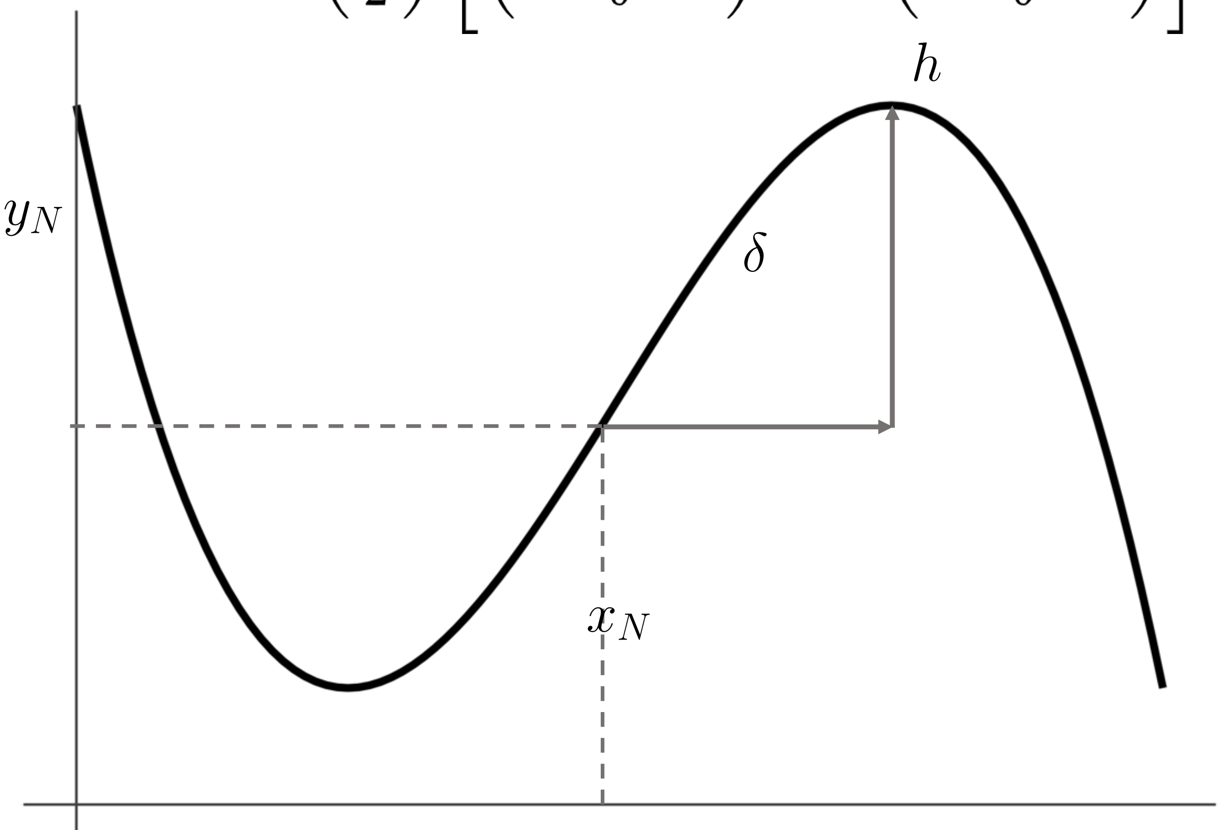
$$y_{ti} = \alpha_y - (\alpha_y - \alpha_0)\left(\frac{x_{ti}}{\alpha_x} - 1\right)^2 + \varepsilon_{ti}$$

Defining Nonlinear Equations

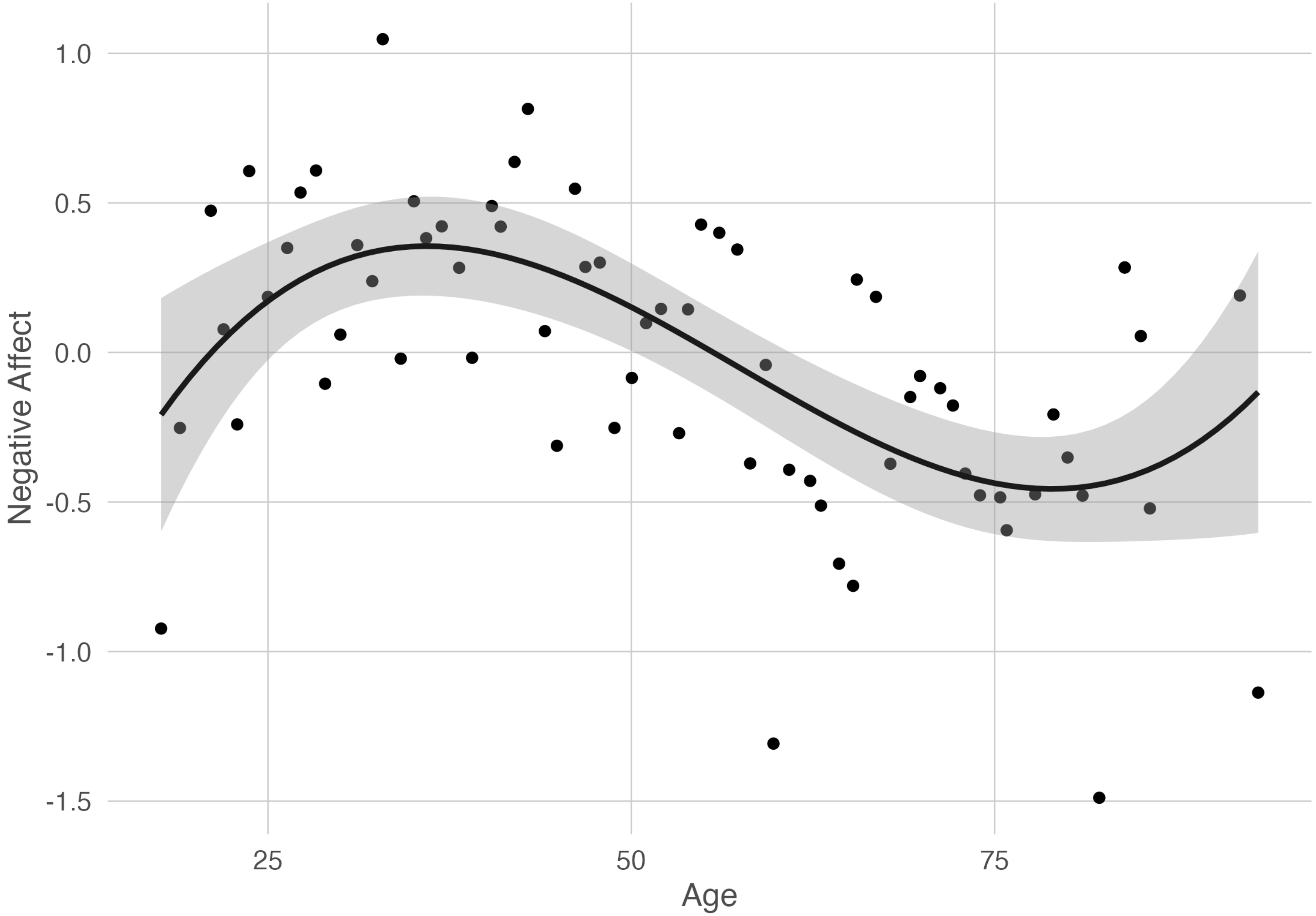


$y_{ti} = \alpha_y - (\alpha_y - \alpha$

$$y_{ti} = y_N - \left(\frac{h}{2}\right)\left[\left(\frac{x_{ti} - x_N}{\delta}\right)^3 - 3\left(\frac{x_{ti} - x_N}{\delta}\right)\right] + \varepsilon_{ti}$$



Defining Nonlinear Equations



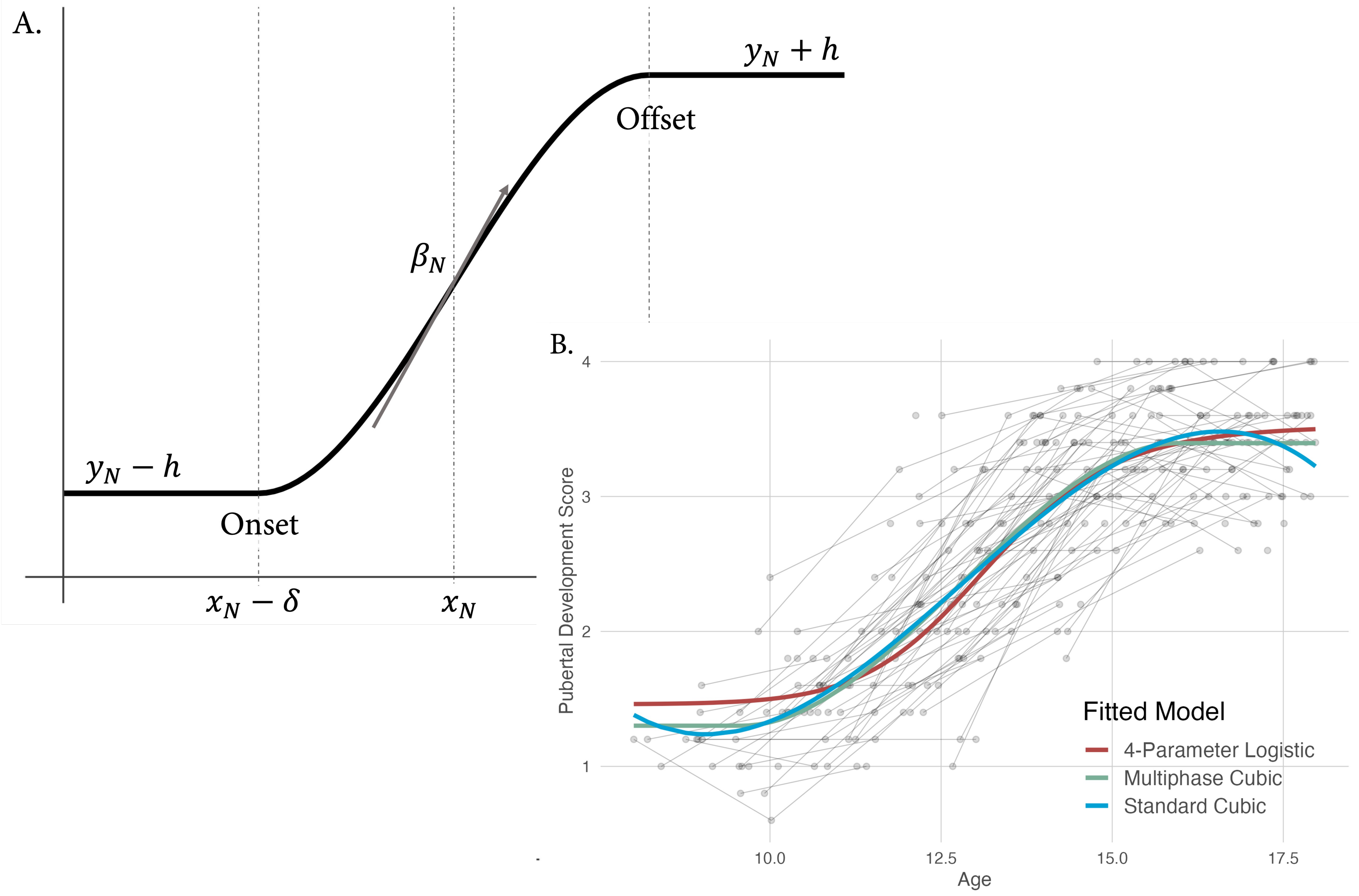
Defining Nonlinear Equations

Fitting Linear and Nonlinear Parameter Cubic Models

	Linear Parameter Model		Nonlinear Parameter Model
β_0	–2.288** (0.802)	x_N	57.428*** (2.162)
β_1	0.174** (0.052)	y_N	–0.050 (0.051)
β_2	–0.004*** (0.001)	δ	21.508*** (1.711)
β_3	2.04×10^{-5}** (6.10×10^{-6})	h	–0.406*** (0.068)
Num.Obs.	69		69
R^2	0.355		
AIC	76.6		76.6
BIC	87.7		87.7

Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Defining Nonlinear Equations



Issues with Nonlinear Models

- Nonlinear with respect the parameters
  - Not defined for all values of all parameters

$$y_{ti} = \alpha_y - (\alpha_y - \alpha_0)\left(\frac{x_{ti}}{\alpha_x} - 1\right)^2 + \varepsilon_{ti}$$

$$y_{ti} = y_N - \left(\frac{h}{2}\right)\left[\left(\frac{x_{ti} - x_N}{\delta}\right)^3 - 3\left(\frac{x_{ti} - x_N}{\delta}\right)\right] + \varepsilon_{ti}$$

Issues with Nonlinear Models

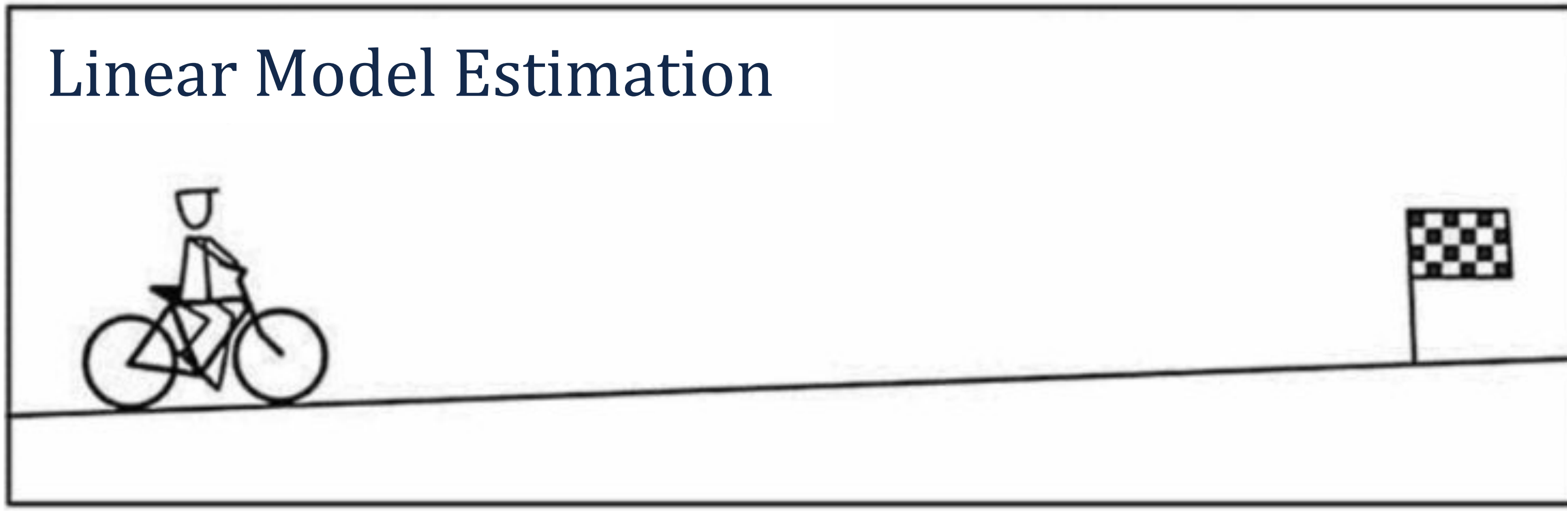
- No hierarchy of parameters for random effects

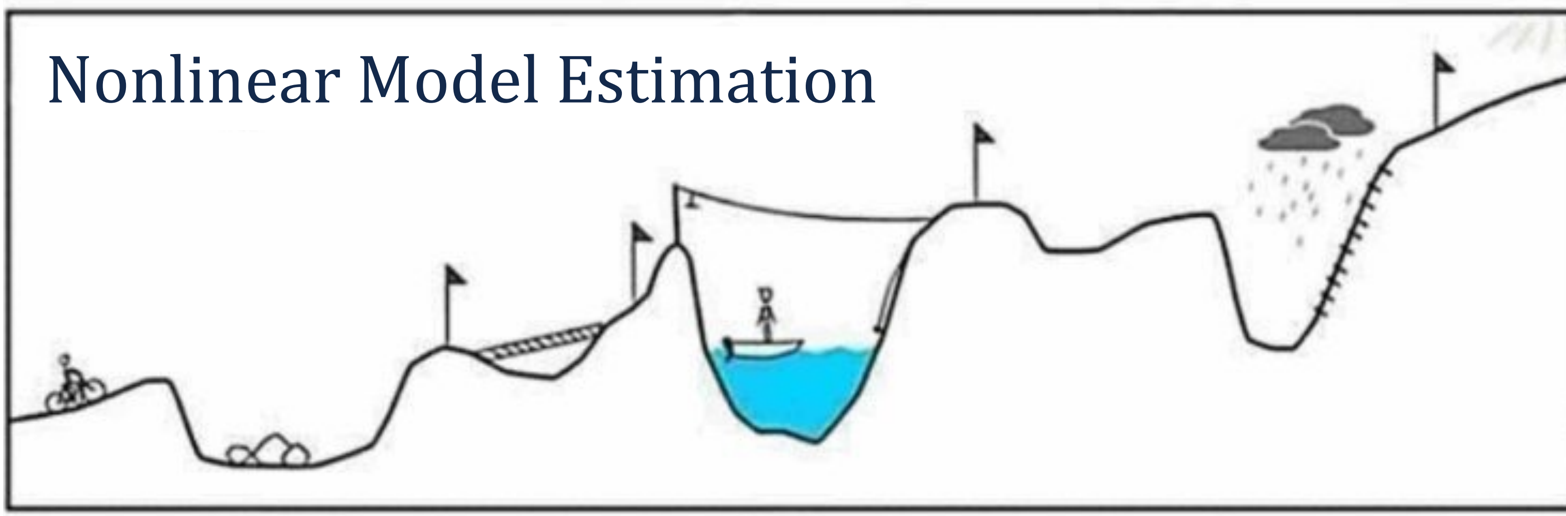
$$y_{ti} = \beta_0 + \beta_1 x_{ti} + \beta_2 x_{ti}^2 + \beta_3 x_{ti}^3 + \varepsilon_{ti}$$

$$y_{ti} = y_N - \left(\frac{h}{2}\right)\left[\left(\frac{x_{ti} - x_N}{\delta}\right)^3 - 3\left(\frac{x_{ti} - x_N}{\delta}\right)\right] + \varepsilon_{ti}$$

Issues with Nonlinear Models

- Nonlinear models are just harder to estimate





Solutions: Linear Estimation, Nonlinear Inference (LENI)

- Estimate the linear equation & define transformation functions to generate the implied nonlinear estimates

$$y_{ti} = \beta_0 + \beta_1 x_{ti} + \beta_2 x_{ti}^2 + \beta_3 x_{ti}^3 + \varepsilon_{ti}$$

Solutions: Linear Estimation, Nonlinear Inference (LENI)

- Estimate the linear equation & define transformation functions to generate the implied nonlinear estimates

$$y_{ti} = \beta_0 + \beta_1 x_{ti} + \beta_2 x_{ti}^2 + \beta_3 x_{ti}^3 + \varepsilon_{ti}$$

$$y_{ti} = y_N - \left(\frac{h}{2}\right)\left[\left(\frac{x_{ti} - x_N}{\delta}\right)^3 - 3\left(\frac{x_{ti} - x_N}{\delta}\right)\right] + \varepsilon_{ti}$$

Solutions: Linear Estimation, Nonlinear Inference (LENI)

- Estimate the linear equation & define transformation functions to generate the implied nonlinear estimates

$$y_{ti} = \beta_0 + \beta_1 x_{ti} + \beta_2 x_{ti}^2 + \beta_3 x_{ti}^3 + \varepsilon_{ti}$$

<u>Fixed Effects + Standard Errors</u>

$$x_N = \frac{-\beta_2}{3\beta_3} \qquad y_N = \beta_0 - \frac{\beta_1\beta_2}{3\beta_3} + \frac{2\beta_2^3}{27\beta_3^2} \qquad \delta = \frac{\sqrt{\beta_2^2 - 3\beta_3\beta_1}}{3\beta_3} \quad h = -2\beta_3\delta^3$$

$$\text{ACOV}(f(x_N, y_N, \delta, h)) \approx \mathbf{J}'_{f(x_N,y_N,\delta,h)}\ \text{ACOV}(\boldsymbol{\beta})\ \mathbf{J}_{f(x_N,y_N,\delta,h)}$$

$$y_{ti} = y_N - \left(\frac{h}{2}\right)\left[\left(\frac{x_{ti} - x_N}{\delta}\right)^3 - 3\left(\frac{x_{ti} - x_N}{\delta}\right)\right] + \varepsilon_{ti}$$

Solutions: Linear Estimation, Nonlinear Inference (LENI)

- Estimate the linear equation & define transformation functions to generate the implied nonlinear estimates

$$y_{ti} = \beta_0 + \beta_1 x_{ti} + \beta_2 x_{ti}^2 + \beta_3 x_{ti}^3 + \varepsilon_{ti}$$

<u>Fixed Effects + Standard Errors</u>

$$x_N = \frac{-\beta_2}{3\beta_3} \qquad y_N = \beta_0 - \frac{\beta_1\beta_2}{3\beta_3} + \frac{2\beta_2^3}{27\beta_3^2} \qquad \delta = \frac{\sqrt{\beta_2^2 - 3\beta_3\beta}}{3\beta_3} \quad h = -2\beta_3\delta^3$$

$$\mathrm{ACOV}(f(x_N, y_N, \delta, h)) \approx \mathbf{J}'_{f(x_N,y_N,\delta,h)}\ \mathrm{ACOV}(\boldsymbol{\beta})\ \mathbf{J}_{f(x_N,y_N,\delta,h)}$$

<u>Random Effects + Standard Errors</u>

$$\mathbf{T}_{f(x_N,y_N,\delta,h)} \approx \mathbf{J}'_{f(x_N,y_N,\delta,h)}\ \mathbf{T}_{\boldsymbol{\beta}}\ \mathbf{J}_{f(x_N,y_N,\delta,h)}$$

$$\mathrm{ACOV}(\mathbf{T}_{f(x_N,y_N,\delta,h)}) \approx \mathbf{J}'_{\mathbf{T}_{f(x_N,y_N,\delta,h)}}\ \mathrm{ACOV}(\mathbf{T}_{\boldsymbol{\beta}})\ \mathbf{J}_{\mathbf{T}_{f(x_N,y_N,\delta,h)}}$$

$$y_{ti} = y_N - \left(\frac{h}{2}\right)\left[\left(\frac{x_{ti} - x_N}{\delta}\right)^3 - 3\left(\frac{x_{ti} - x_N}{\delta}\right)\right] + \varepsilon_{ti}$$

Solutions: Linear Estimation, Nonlinear Inference (LENI)

- Estimate the linear equation & define transformation functions to generate the implied nonlinear estimates

LENI Approach to Fixed Effects Estimation

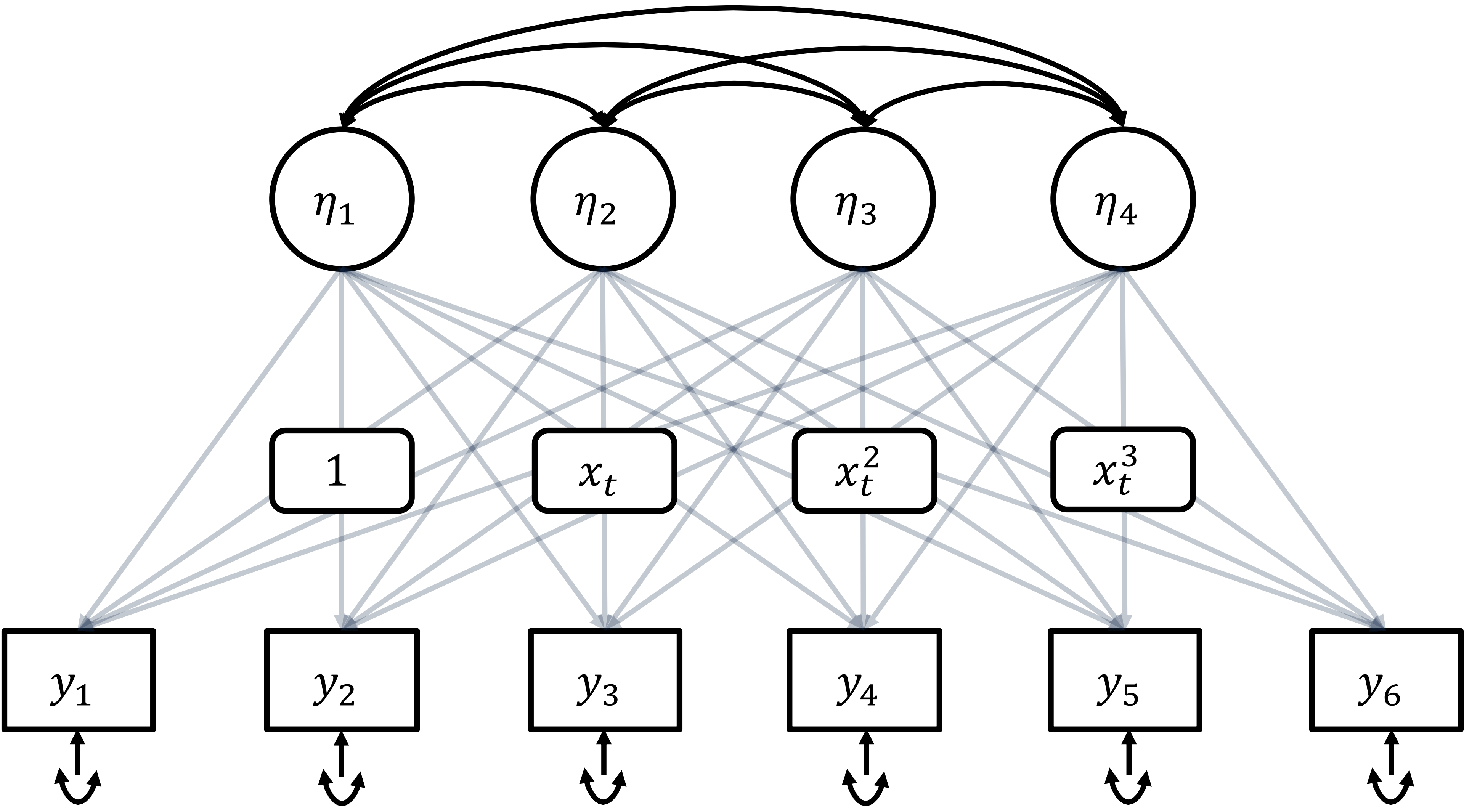
Pop. θ		Linear Estimates		LENI Estimates		Nonlinear Estimates	
Cubic Model							
x_N	0	β_0	9.994 (0.139)	x_N	−0.002 (0.116)	x_N	−0.002 (0.116)
y_N	10	β_1	−1.000 (0.080)	y_N	9.996 (0.094)	y_N	9.996 (0.094)
δ	3	β_2	2.25×10^{-4} (0.013)	δ	3.015 (0.107)	δ	3.015 (0.107)
h	−2	β_3	0.037 (0.005)	h	−2.010 (0.130)	h	−2.010 (0.130)
R^2	0.5		0.506				
BIC			921.90				921.90

Linearized SEM: Polynomials and Beyond

- LENI approach can be applied to mixed-effects polynomial models, but an linearized SEM approach can offer additional flexibility and target functions

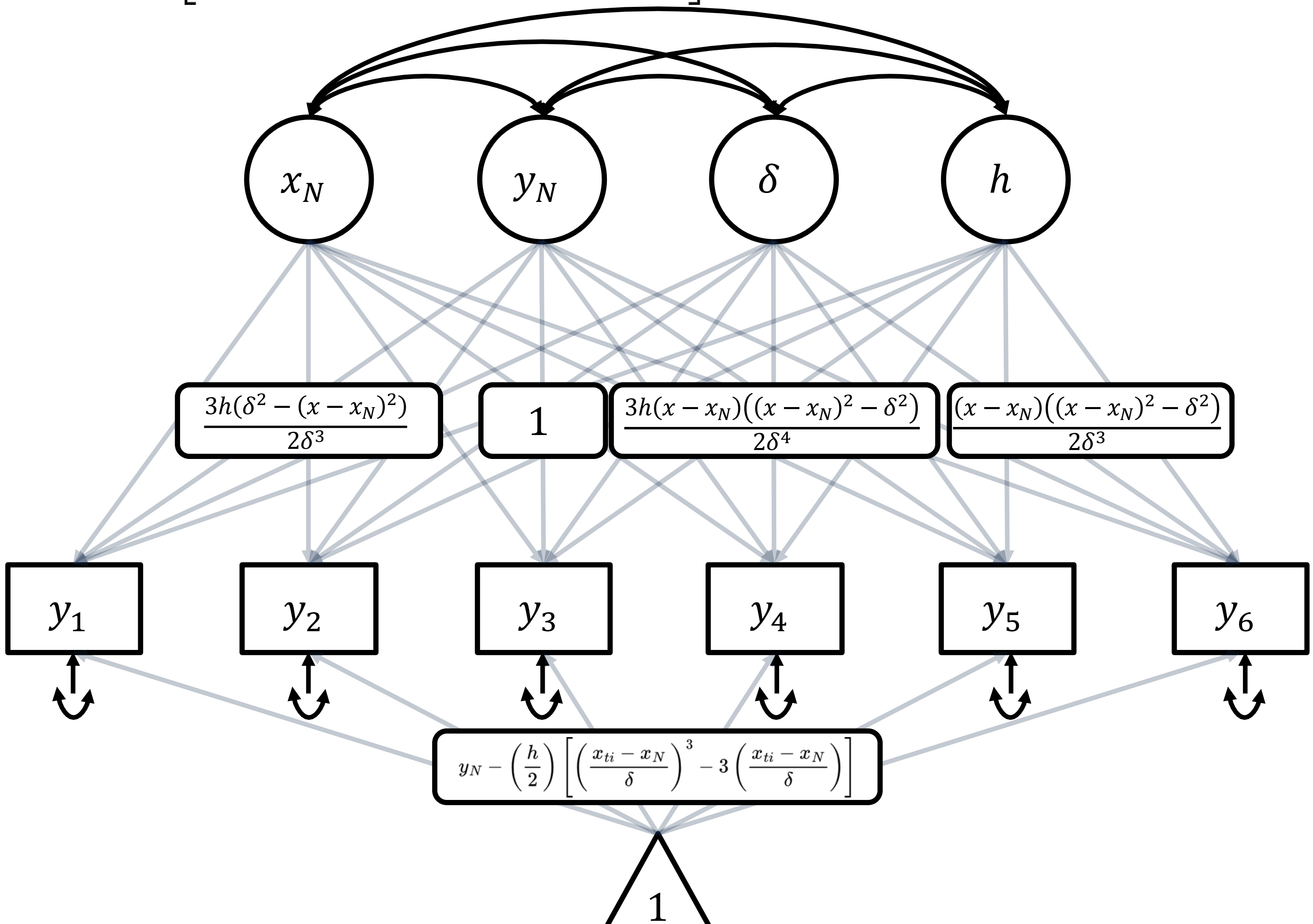
Linearized SEM: Polynomials and Beyond

$$y_{ti} = \beta_0 + \beta_1 x_{ti} + \beta_2 x_{ti}^2 + \beta_3 x_{ti}^3 + \varepsilon_{ti}$$



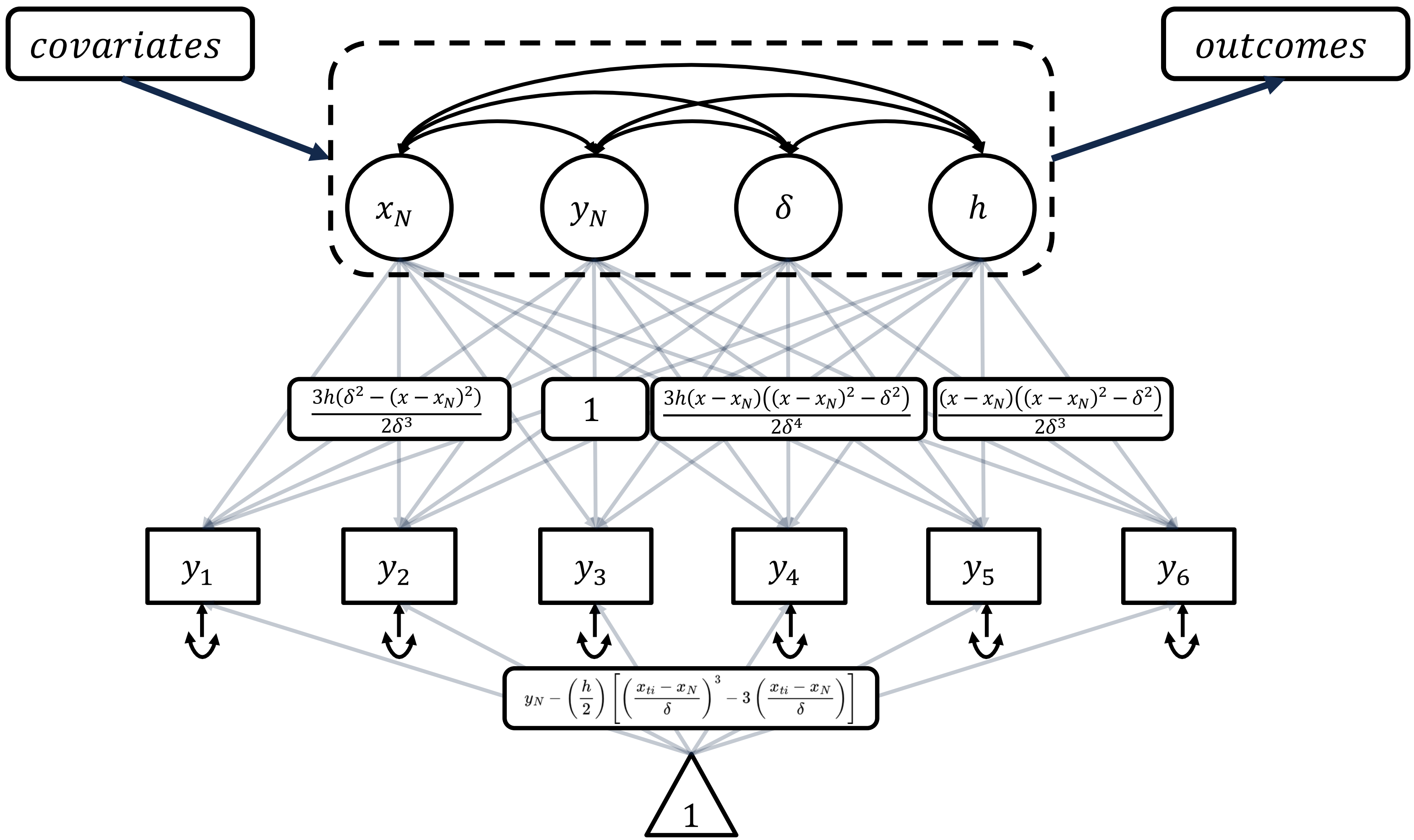
Linearized SEM: Polynomials and Beyond

$$y_{ti} = y_N - \left(\frac{h}{2}\right)\left[\left(\frac{x_{ti} - x_N}{\delta}\right)^3 - 3\left(\frac{x_{ti} - x_N}{\delta}\right)\right] + \varepsilon_{ti}$$

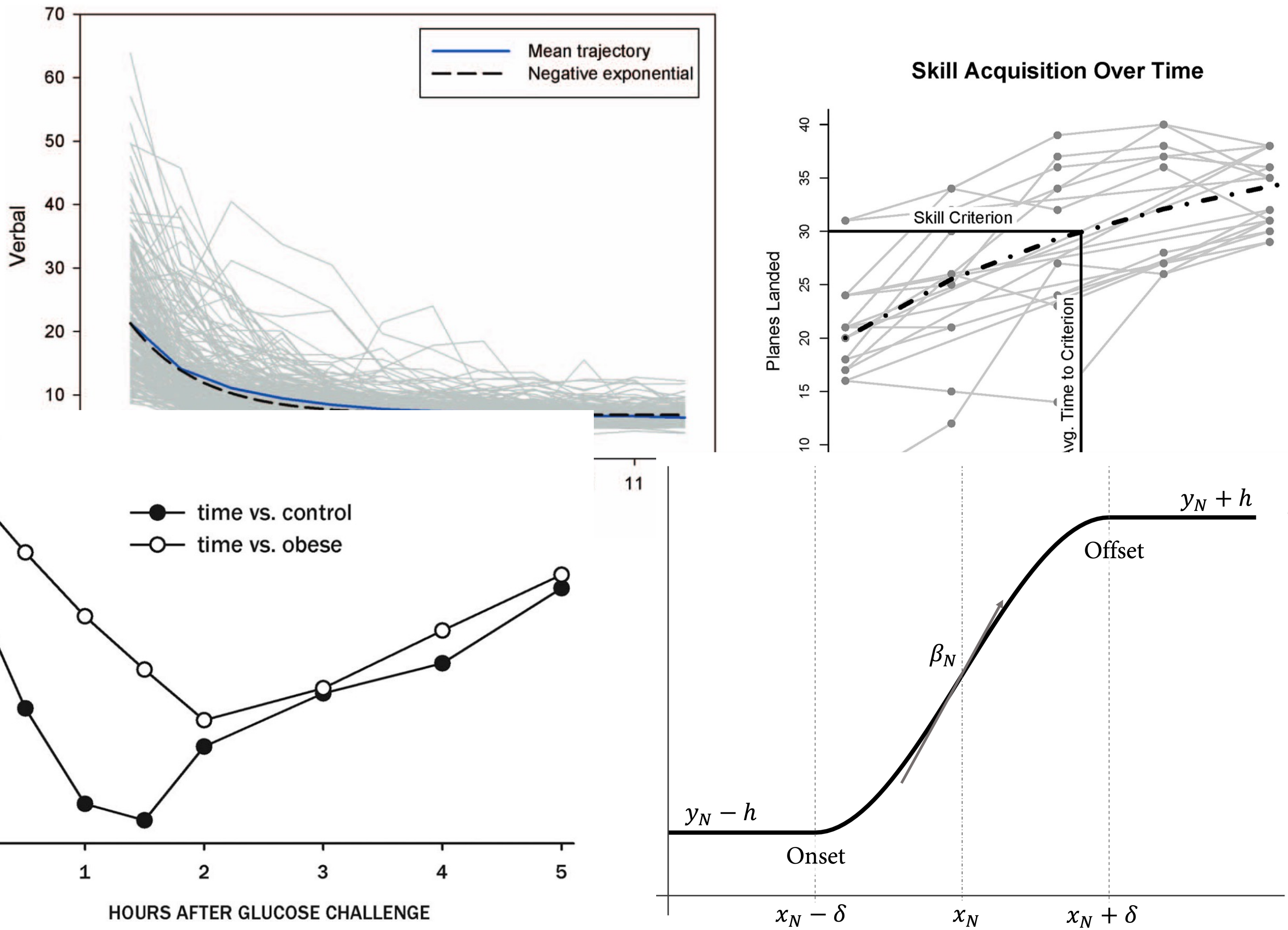


Linearized SEM: Polynomials and Beyond

- Easy inclusion of covariates and distal outcomes



Linearized SEM: Polynomials and Beyond



Moving forward with Meaningful Models

- Issues of estimation have limited the use of nonlinear models
  - Improvements in computation, LENI, linearized SEM offer new opportunities for using these models
  - Bayesian models for unstructured data

- Use models to test specific hypotheses and developmental features of interest
  - Identify new nonlinear target functions

- Meaningful parameters allow for tests of causes and consequences of individual differences in trajectories

Acknowledgements

- Collaborators
  - Patrick Curran, Ph.D.
  - Daniel Bauer, Ph.D.
  - Gregory Hancock, Ph.D.
  - Michelle Byrne, Ph.D.
  - John Flournoy, Ph.D.
  - Kathryn Mills, Ph.D.
  - Jennifer Pfeifer, Ph.D.
  - Rogier Kievit, Ph.D.
  - Sam Parsons, Ph.D.
  - Zsuzsa Bakk, Ph.D.

Universiteit Leiden

THE UNIVERSITY of NORTH CAROLINA at CHAPEL HILL

JACOBS FOUNDATION
Our Promise to Youth

NWO
Netherlands Organisation for Scientific Research

NIH
National Institute of Mental Health

NSF