

# Leveraging missing data to model simultaneous growth processes

Ethan M. McCormick, Ph.D.

Postdoctoral Fellow (Dr. Rogier Kievit)  
Donders Institute at the RadboudUMC

@McCormickNeuro  
[ethan.mccormick@radboudumc.nl](mailto:ethan.mccormick@radboudumc.nl)  
<https://mccormickneuro.github.io/>

# Acknowledgements: A Village



THE UNIVERSITY  
of NORTH CAROLINA  
at CHAPEL HILL



Universiteit  
Leiden



Erasmus  
University  
Rotterdam  
*Erasmus*



Radboudumc



European  
Research  
Council

# Growth Processes in Development

- Across the lifespan, there may be multiple developmental processes influencing behaviors/cognitions of interest
  - Age: opportunity differences, peer effects, (often a catch-all)
  - Puberty: tuning of physiology
  - Experience: practice, consolidation
- In longitudinal studies, we also worry about retest effects (e.g., practice, scanner anxiety)
- However, common models and study designs don't allow us to test these competing explanatory variables simultaneously

# Growth Models

- Will use multilevel (mixed-effects) notation
  - (But MLMs are just a specific form of latent variable models so this could all be done in an SEM framework as well)

- Unconditional model:

$$y_{ti} = \underbrace{\gamma_{00} + \gamma_{10}Time_{ti}}_{fixed\ effects} + \underbrace{u_{0i} + u_{1i}Time_{ti}}_{random\ effects} + r_{ti}$$

- Conditional Model:

$$y_{ti} = \gamma_{00} + \gamma_{10}Time_{ti} + \gamma_{20}TVC_{ti} + u_{0i} + u_{1i}Time_{ti} + r_{ti}$$

# Time as the Original TVC

- While time “defines” a growth model, time is treated as any other predictor in the model
  - Means that including multiple growth predictors is trivial in principle
  - But...
- Growth processes tend to be highly collinear
  - Monotonic increases
  - Variance inflation:  standard errors, parameter instability
  - A role for \*planned\* missing data

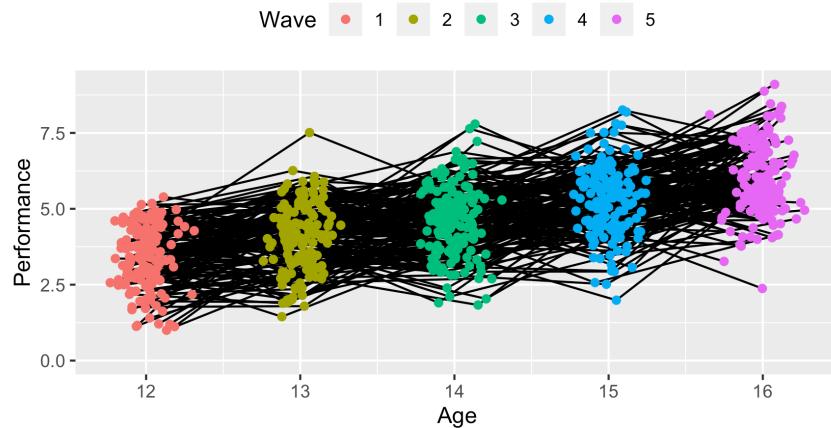
# Retest Effects: A Simulation Example

Paper Link

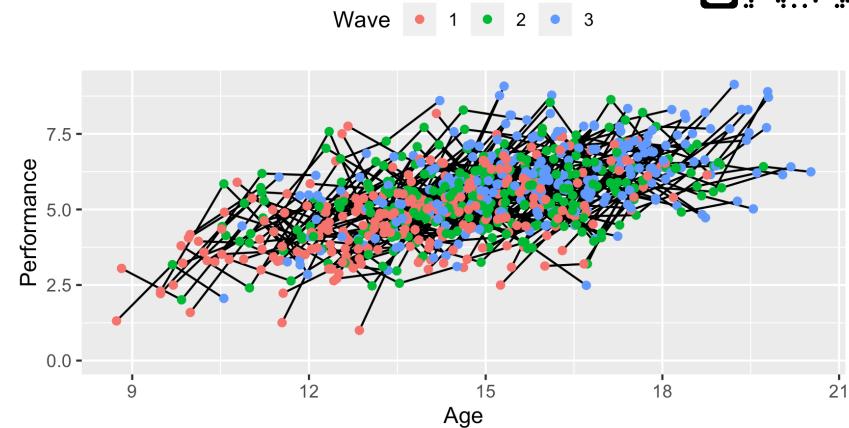


## Additive Age and Practice Effects: Comparisons Between Cohort and Accelerated Designs

A.



B.



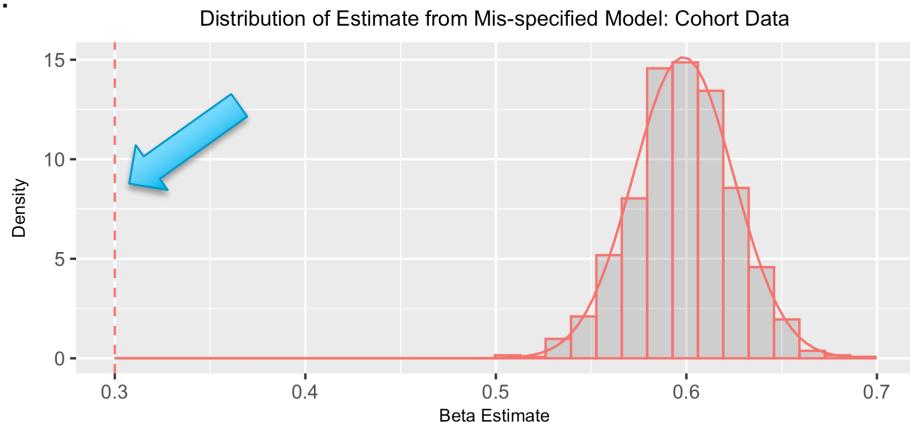
**Table 2**

Estimated Effects for Growth Models from Scenario Set 1: Additive Effects.

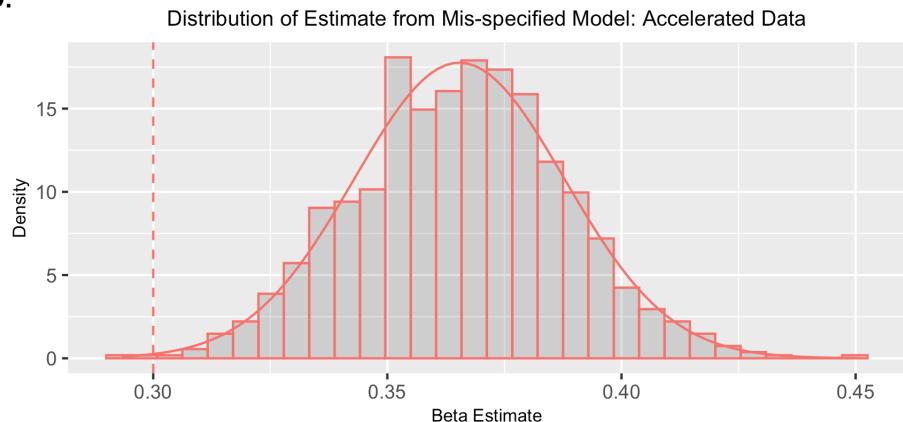
	Cohort Design						Accelerated Design			
	Mean Est.	Std. Err.	Min	Max	Std. Bias	Prop. Sig.(adj.)	Mean Est.	Std. Err.	Min	Max
<i>Additive Effects</i>										
$\rho_{growth}$	.998	.000	.997	.998			.379	.015	.338	.434
VIF	202.2	10.7	169.2	238.8			1.17	.016	1.13	1.23
14.2x larger SEs						1.03x larger SEs				

# Retest Effects: A Simulation Example

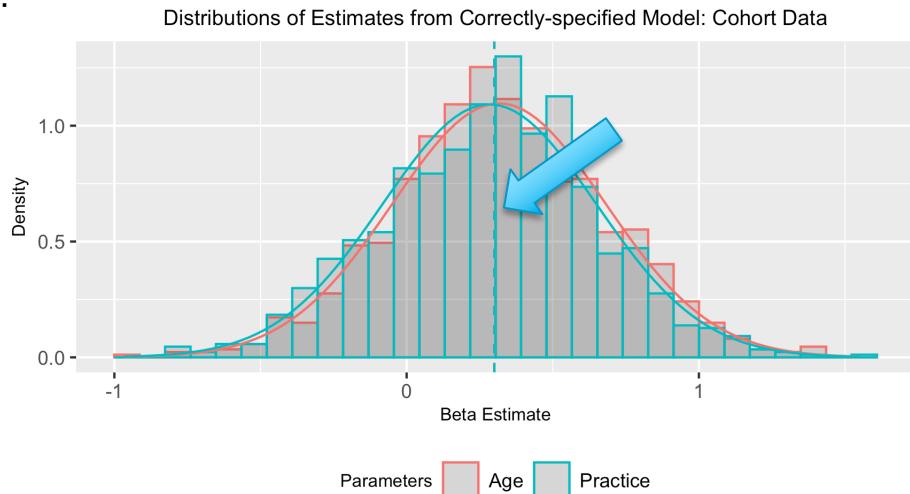
C.



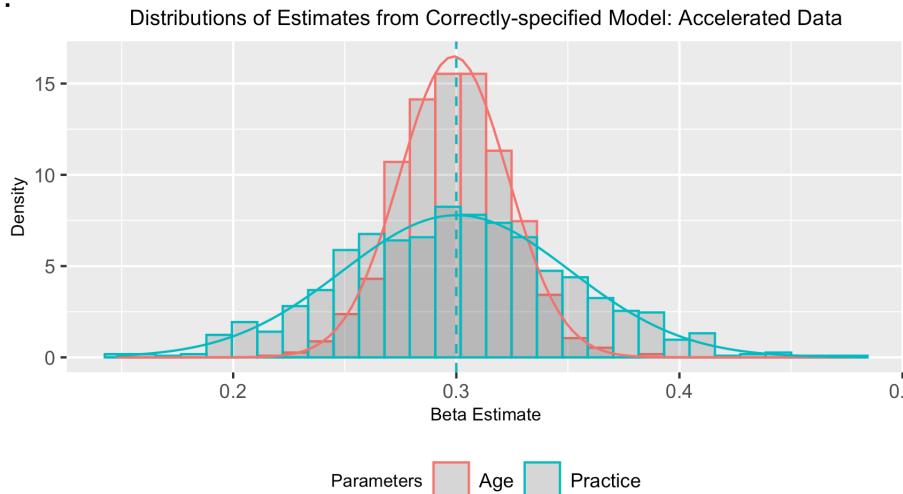
D.



E.

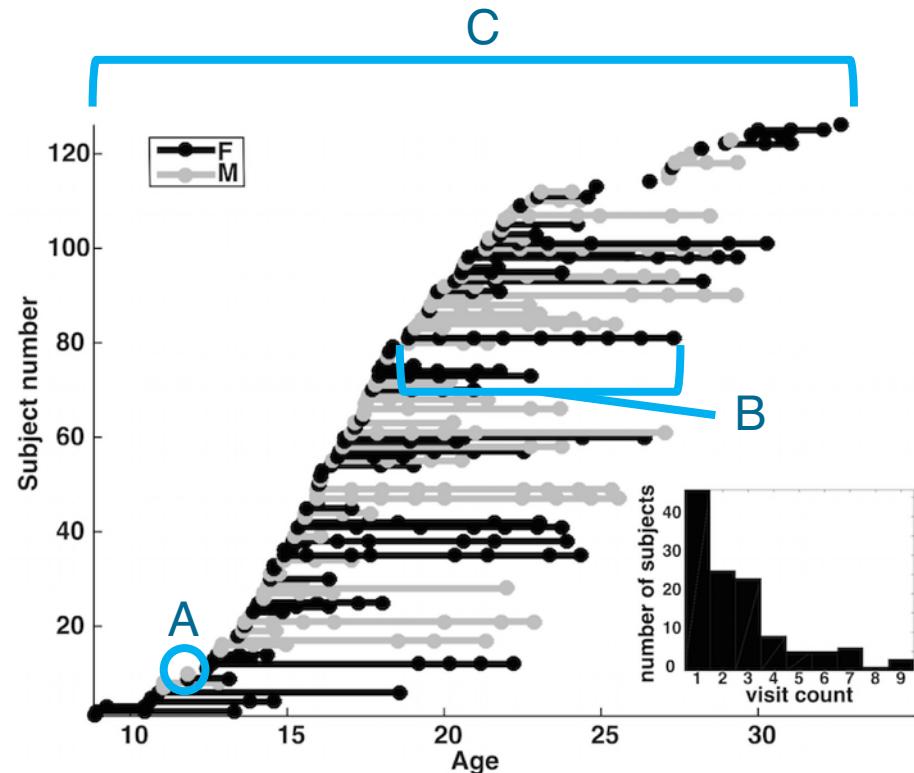


F.



# Retest Effects: An Empirical Example

- 3 levels of change possible
  - Learning within session (A)
  - Practice across sessions (B)
  - Developmental effects (C)

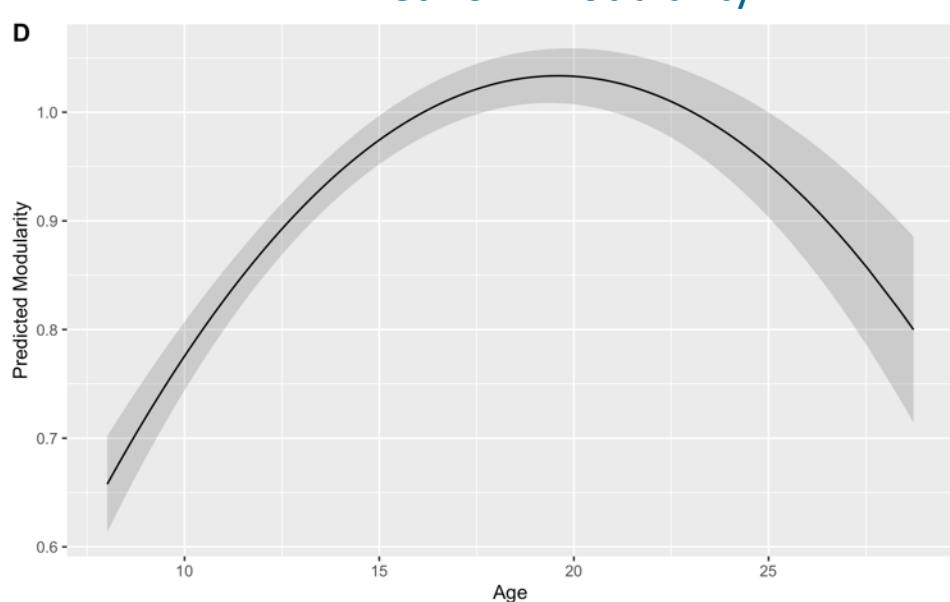
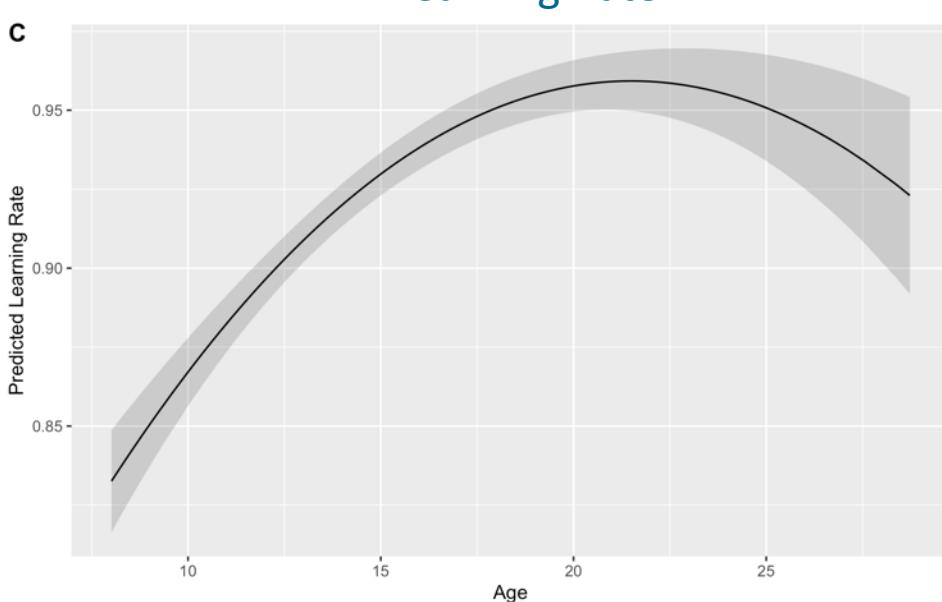


- Data
  - BrainTime: 299 participants (age 8-29 years), up to 3 waves
  - Feedback Learning Task (Peters et al., 2016; Peters & Crone, 2017)





## Age-only Models



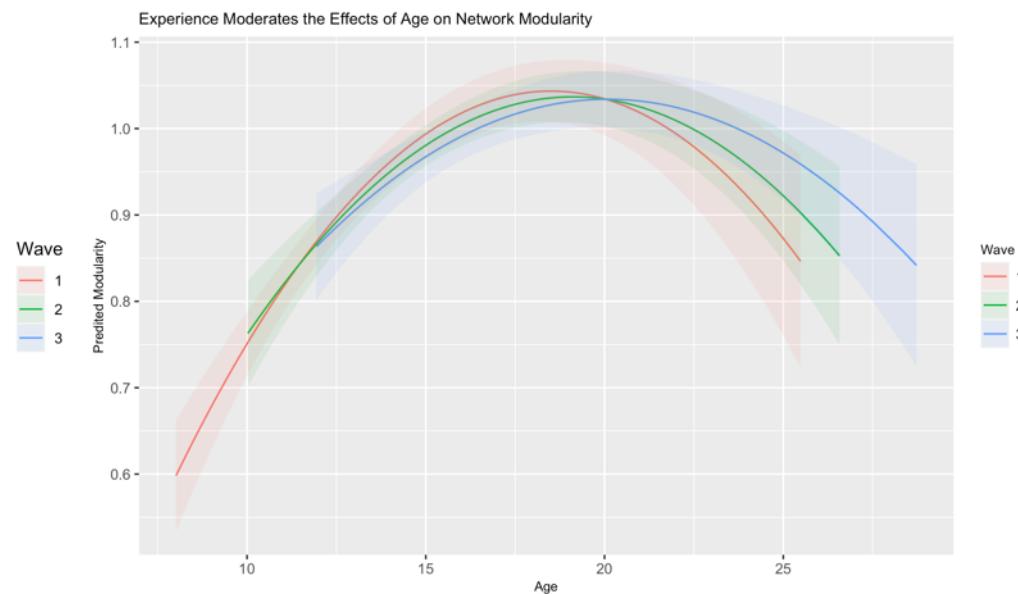
# Retest Effects: An Empirical Example

## Interaction Models

### Learning Rate



### Network Modularity

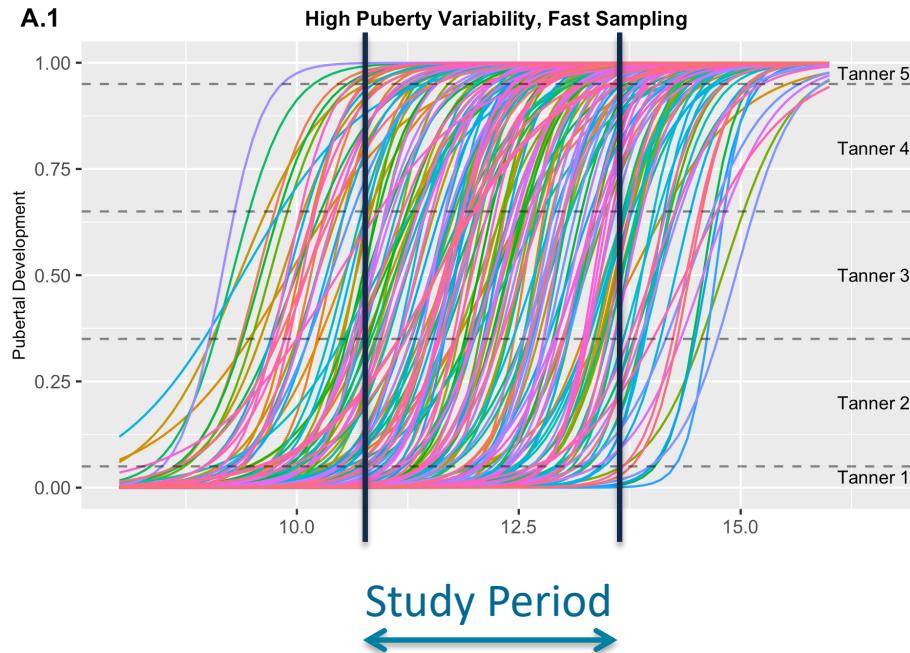


## Some Takeaways

- Highly collinear growth processes act as either unmeasured confounders or introduce variance inflation when modeled.
  - Unless controlled in design choices
- Planned missing data (e.g., using accelerated designs) reduces the linear association between these growth predictors.
- These models can be used investigate multiple growth processes and the interactions between them.

# Modeling Puberty and Age

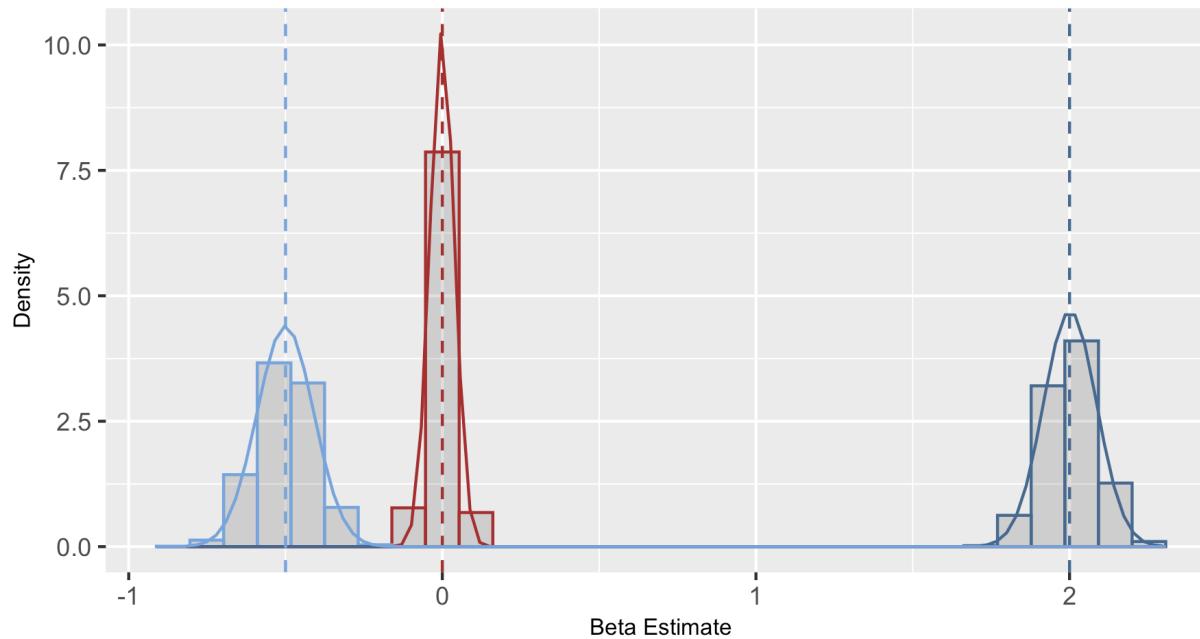
- Not all is lost for cohort designs, but we need to think of missing data on another growth process.
  - Important for there to be variation in the dataset



# Modeling Puberty and Age

In cohort data

C.1

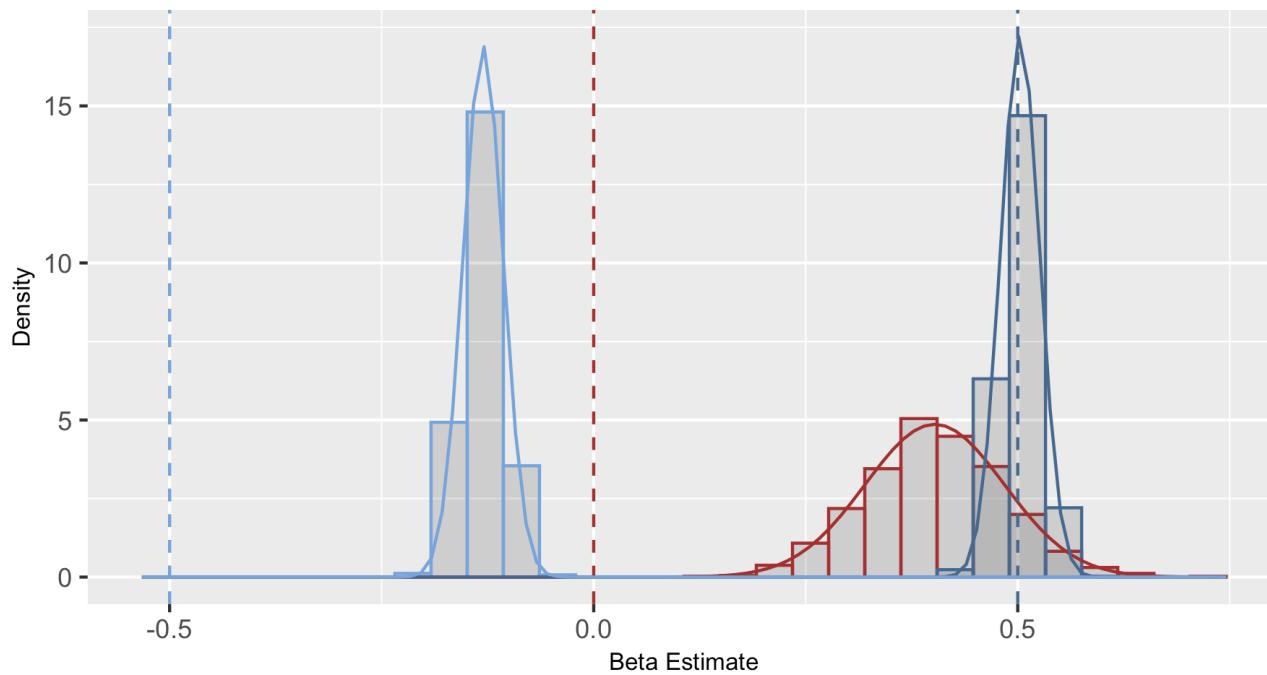


Parameters      Age      Puberty      Interaction

Note: With continuous puberty measure

# Modeling Puberty and Age

D.1



Parameters      Age      Puberty      Interaction

With course categorization of puberty measure → Tanner Stage

# The Promise of Missing Data Designs

- Missing data designs allow us to decouple correlated growth processes and model them simultaneously.
  - e.g., accelerated longitudinal, burst measurement, irregular assessment schedules
- Most of our designs are structured around age and so may need to rely on natural variation in other processes like puberty.
  - Could subset subjects from larger cohort studies like ABCD
  - Structuring assessments around pubertal timing?

# Final Takeaways

- We cannot model our way out of all of our problems.
  - Study design, measurement remain crucial to investigate the causal mechanisms of development
- Planned missing designs are a powerful tool for modeling simultaneous growth processes by decoupling temporal predictors.
- Can leverage naturally occurring variation in existing datasets and inform future data collection.

## Interested in Discussing/Collaborating Further?



@McCormickNeuro



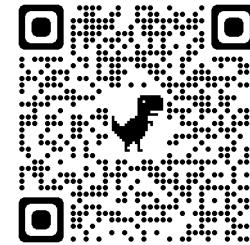
[ethan.mccormick@radboudumc.nl](mailto:ethan.mccormick@radboudumc.nl)



<https://mccormickneuro.github.io/>



McCormick (2021)  
*DCN*



McCormick, Peters, Crone, & Telzer (2021)  
*NeuroImage*