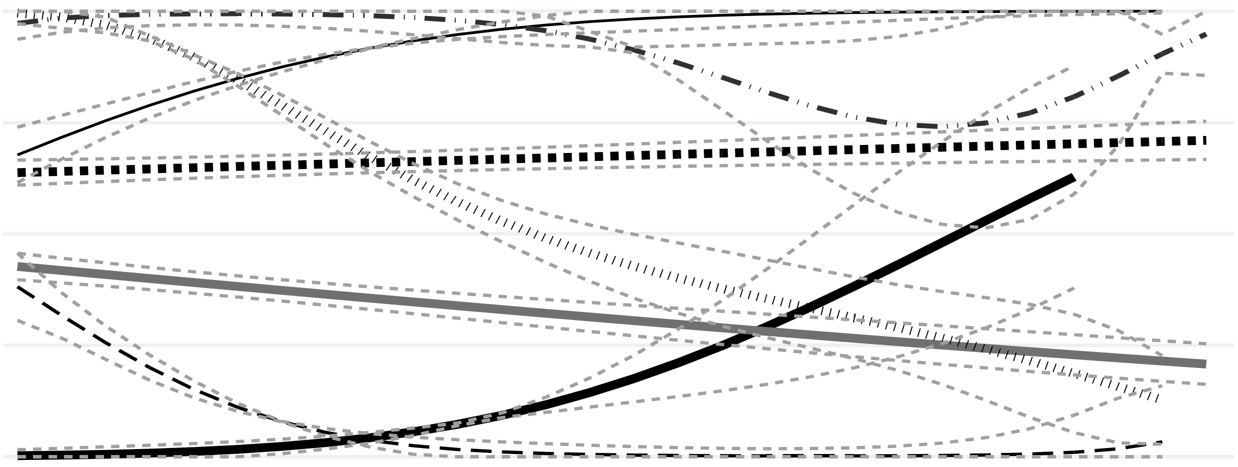
Tornblad/LUPOP seminar series
March 9, 2023; 13:00-14:00



Group-based trajectory modeling

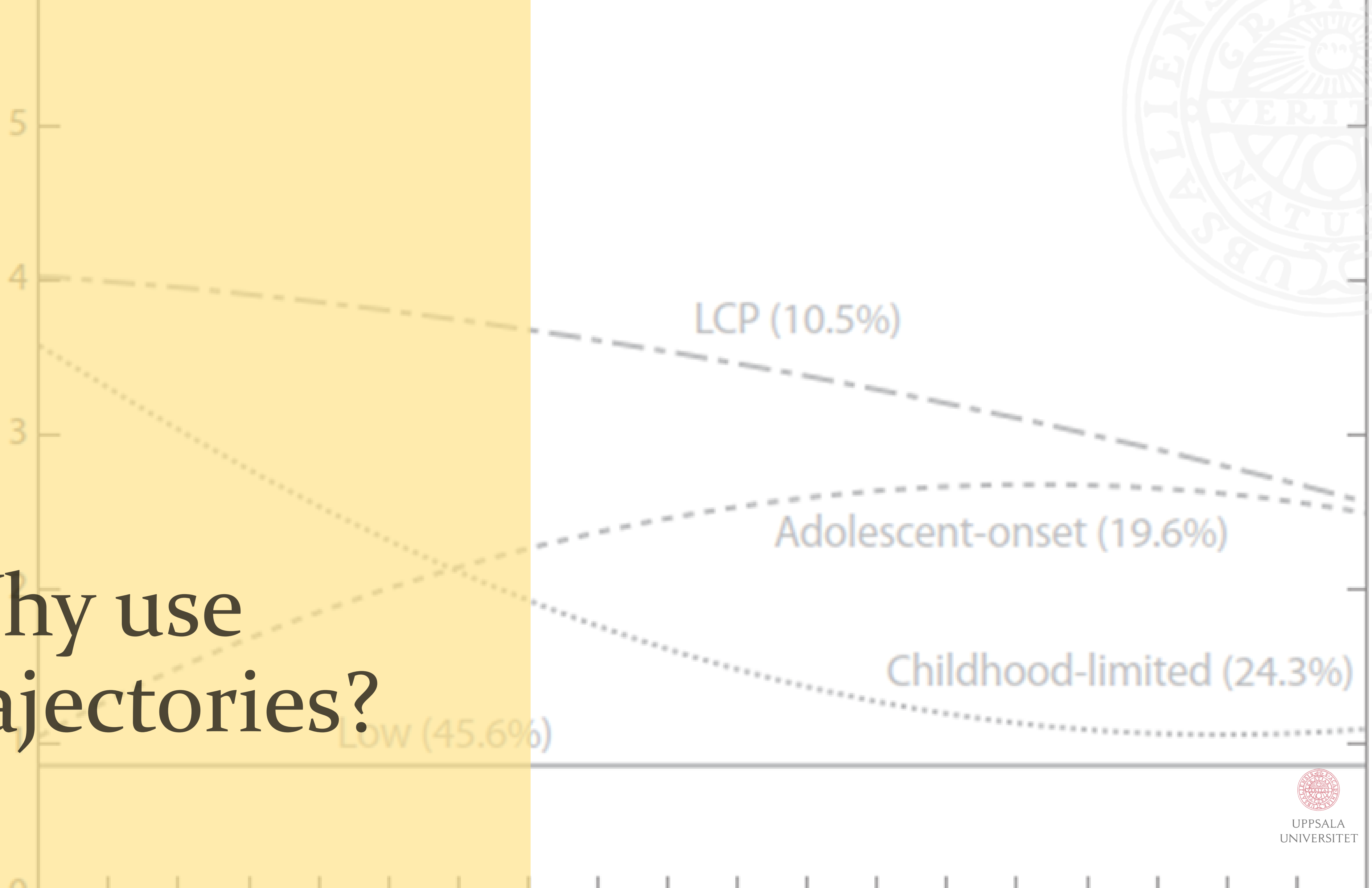
Emerald G. Heiland, postdoc

Medical Epidemiology, Dept Surgical Sciences, Uppsala University

Dept Physical Activity & Health, The Swedish School of Sport and Health Sciences (GIH)

UPPSALA UNIVERSITET

Why use trajectories?



Estimating trajectories

- Standard statistical approaches
  - Hierarchical modeling
  - Latent curve analysis
- Accounts for individual variability about a mean population trend
- Doesn’t take into account the qualitative dimension of longitudinal data

UPPSALA UNIVERSITET

Trajectory modeling

Growth curve modeling (GCM)

Hierarchical modeling

Latent curve analysis

Growth mixture modeling (GMM)

Group-based trajectory modeling (GBTM)

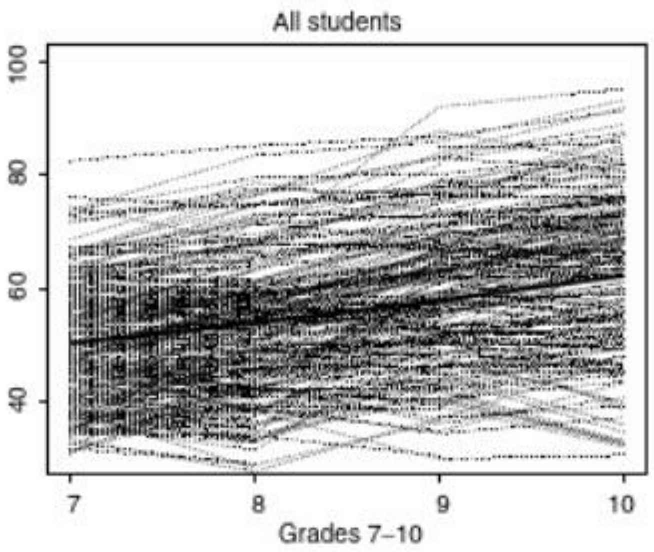
Assumptions

Growth curve modelling (GCM)

Does one type fit all?

Growth mixture modelling (GMM)

Group-based trajectory modelling (GBTM)

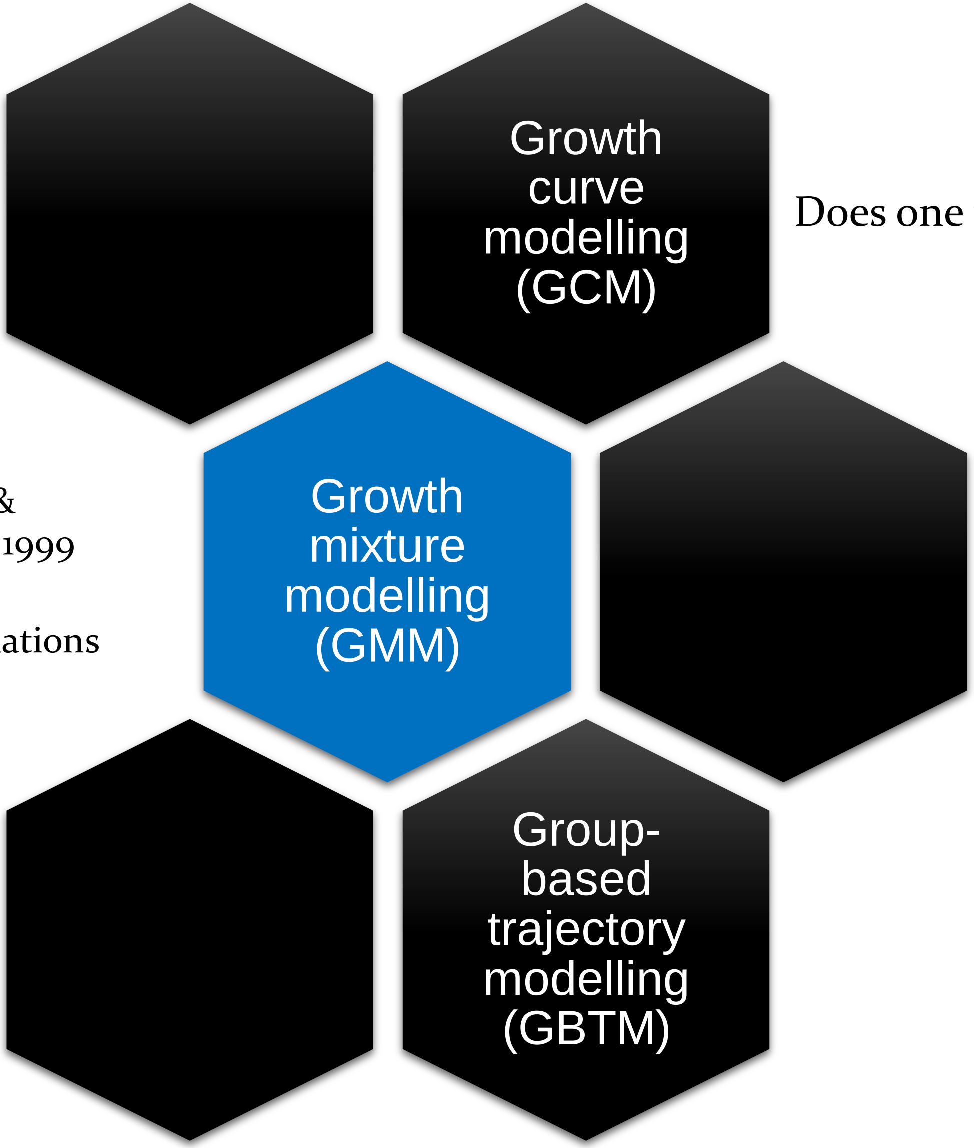


UPPSALA UNIVERSITET

Assumptions

Muthén & Shedden 1999

Distinct subpopulations



Does one type fit all?

UPPSALA UNIVERSITET

Assumptions

Muthén & Shedden 1999

Distinct subpopulations

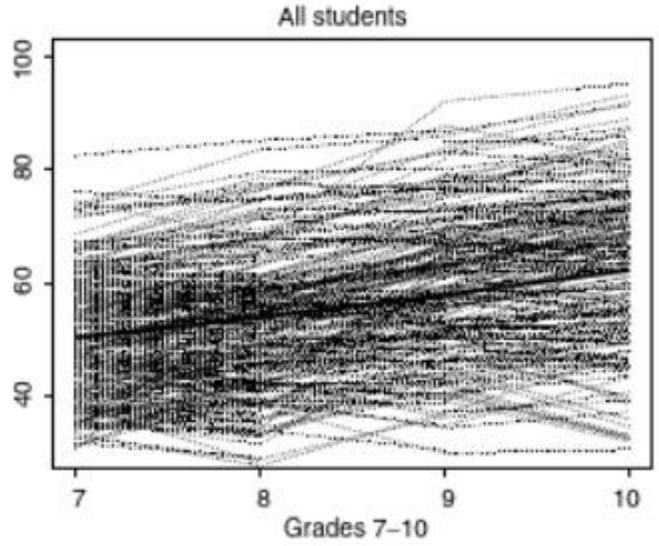
Growth curve modelling (GCM)

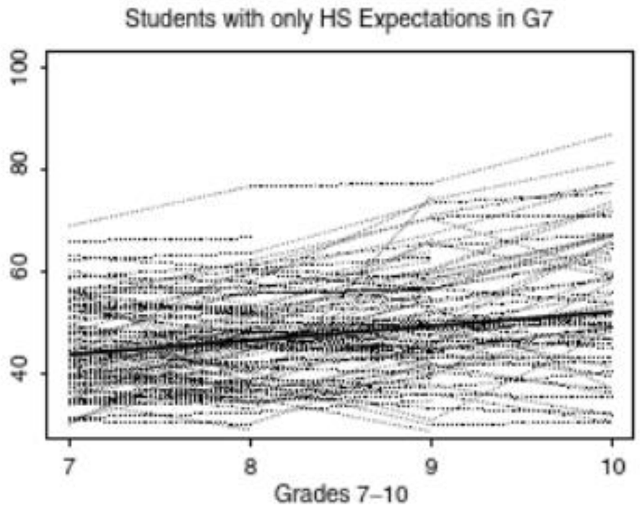
Does one type fit all?

Growth mixture modelling (GMM)

FINITE MIXTURE MODELS

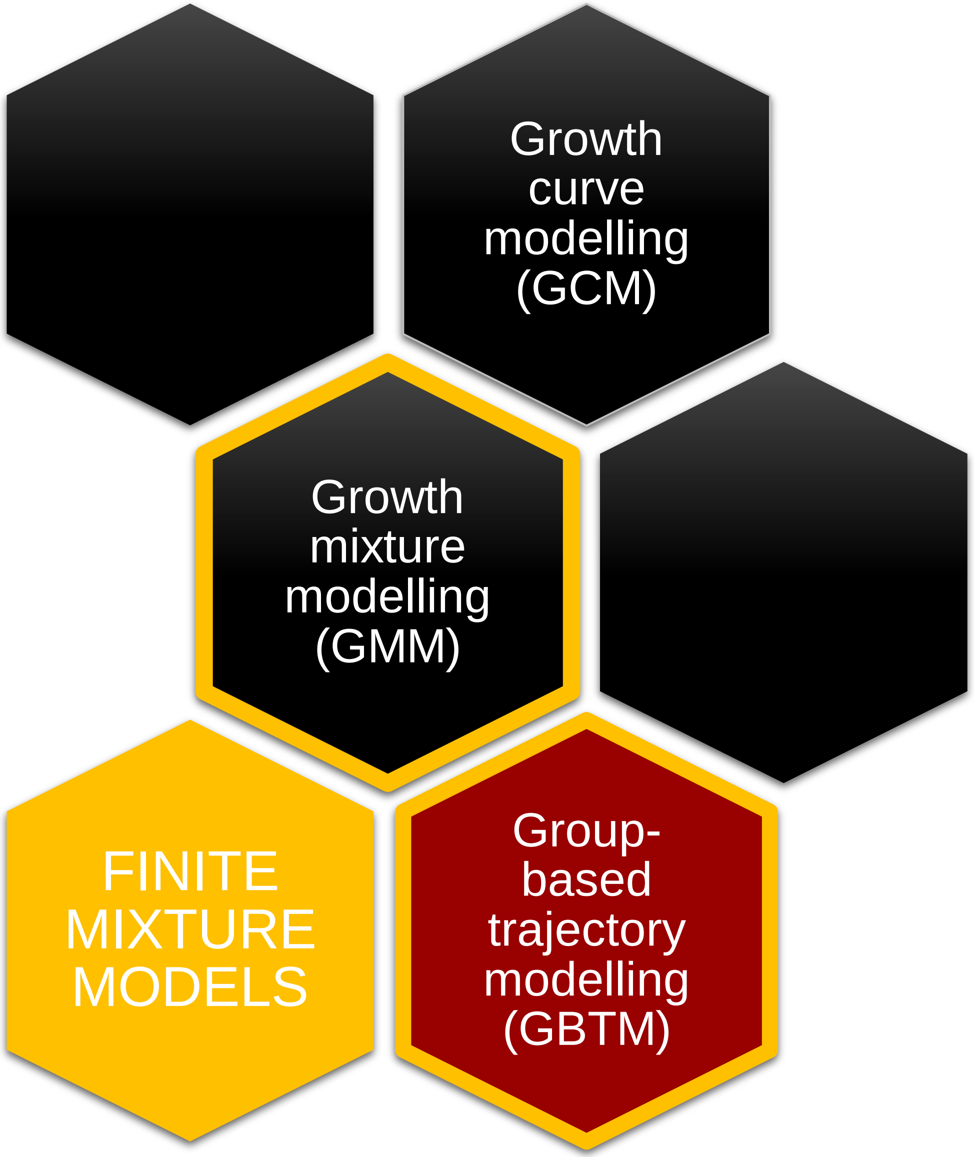
Group-based trajectory modelling (GBTM)





UPPSALA UNIVERSITET

Assumptions



Group-based trajectory modeling

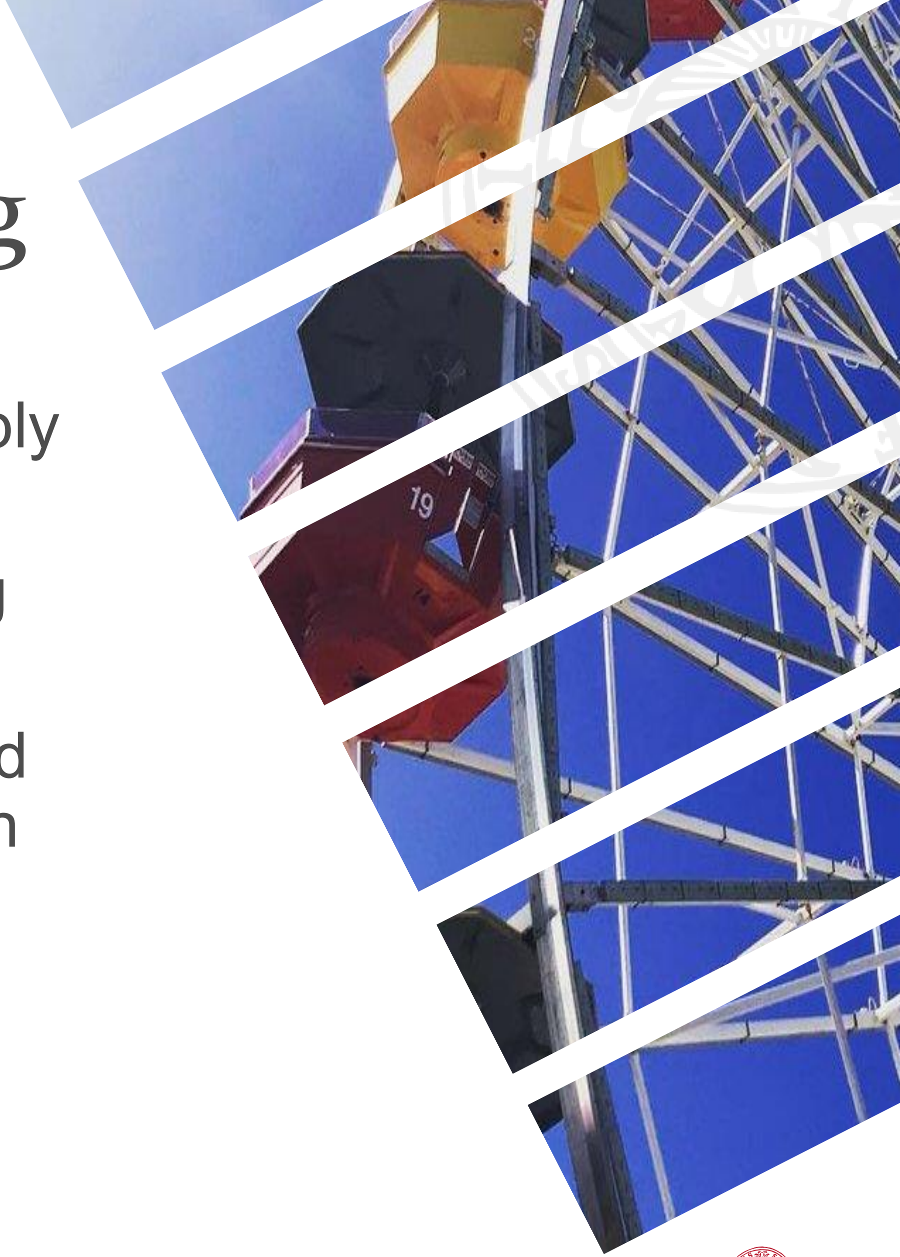
A SAS Procedure Based on Mixture Models
for Estimating Developmental Trajectories

BOBBY L. JONES
DANIEL S. NAGIN
KATHRYN ROEDER
Carnegie Mellon University

2001

Group-based trajectory modeling

- A mixture of probability distributions that are suitably specificed to describe the data to be analyzed
- It is intended to complement hierarchical modeling and latent growth modeling
- Conceptually, group-based trajectory modeling and growth mixture modeling (GMM) are the same with some technical differences



UPPSALA UNIVERSITET

Example of GBTM

Trajectories of physical aggression from age 6 to 15 for males in the Montreal-based longitudinal study sample. (Data from Nagin & Tremblay 1999.)

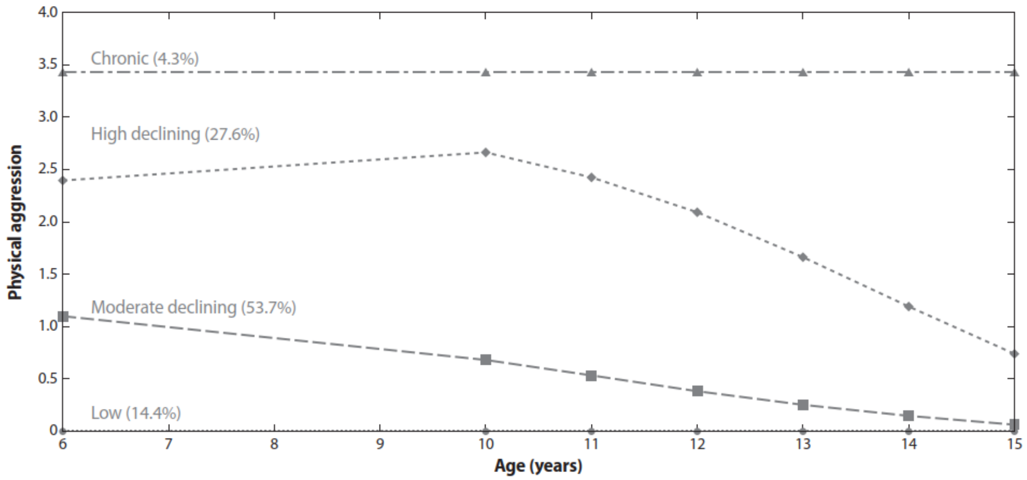


Table 1 Physical aggression group profiles in the Montreal-based longitudinal study. (Data from Nagin & Tramblay 1999)

	Group			
Variable	**Low**	**Moderate declining**	**High declining**	**Chronic**
Years of school: mother	11.1	10.8	9.8	8.4
Years of school: father	11.5	10.7	9.8	9.1
Low IQ (% in lowest quartile)	21.6	26.8	44.5	46.4
Completed eighth grade on time (%)	80.3	64.6	31.8	6.5
Juvenile record (%)	0.0	2.0	6.0	13.3
# of sexual partners age 17[a]	1.2	1.7	2.2	3.5

[a]Number of sexual partners at age 17 within the past year.

UPPSALA
UNIVERSITET

Model selection

Model selection

- Type of distribution
- Number of trajectories
- Shape
- Size

Distribution types

Type of distribution	Type of data	Example
The censored normal distribution	**Continuous**	Longitudinal data on a scale of depression symptoms
The zero-inflated Poisson distribution (ZIP)	**Count**	Arrests by age
Binary logistic distribution	**Dichotomous**	Whether hospitalized in year t or not

The number of groups

- **Bayesian information criteria (BIC)** – most common
- Akaike information criterion (AIC)
- Lo-Mendell-Rubin likelihood ratio test (LMR-LRT)
- Entropy
  - Indexes classification accuracy by averaging the posterior probabilities after individuals have been assigned to their most likely class (range 0 to 1; closer to 1 is greater precision)

The objective of the model selection is not the maximization of some statistic of model fit; rather, it is to summarize the distinctive features of the data in the most parsimonious—and useful—fashion possible

Comparison of goodness of fit in pre-diagnosis physical activity trajectories.			
	Pre-cancer diagnosis		
No. of groups	BIC	Entropy	Polynomial Order
5	-172287.76	0.551	3 1 2 3 3
5	-172274.77	0.570	3 3 2 3 3
6	-171692.67	0.586	2 3 4 2 3 3
6	-171850.70	0.599	2 2 3 1 2 2
7	**-171234.26**	**0.579**	**3 2 3 1 2 2 2**
7	-171251.82	0.579	3 2 3 1 2 1 2
8	-171022.32	0.544	3 2 3 1 2 1 2 2
8	-171005.27	0.541	3 2 3 1 2 1 2 3
8	-171036.61	0.542	3 2 3 1 2 1 2 1

BIC = Bayesian information Criteria. Only show the models where there was significance (P<0.05) in all the groups.

Polynomial order (shape): linear, quadratic, cubic

UPPSALA UNIVERSITET

Suggestions for model selection

Statistical criteria

A) Proportion assigned to the group ≥5%

B) Average of the posterior probabilites ≥0.7

C) Odds of correct classification >5

D) Observing confidence intervals

Important to clearly communicate the decision points and justifications employed to select the best trajectory model

Application Example

Data on self-reported delinquent group membership from age 11 to 17 in a large Montreal-based longitudinal study of over 1,000 males

Outcome

- Self-reported delinquent group membership (yes = 1 / no =0)

Time scale

- Age 11 to 17

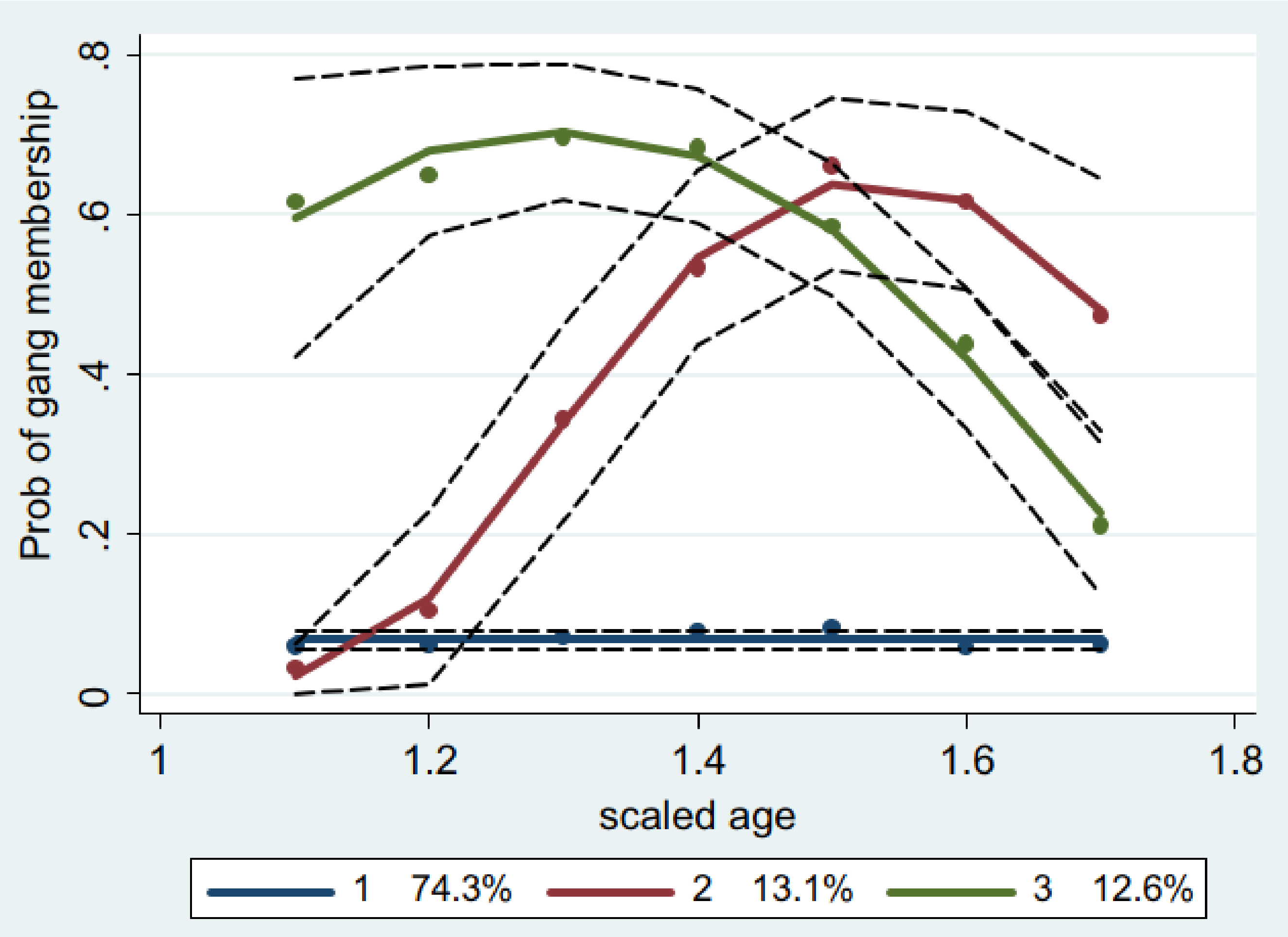
Logistic specification

Best fit

- Number of trajectory groups: 3 based on BIC

UPPSALA
UNIVERSITET

Application Example



UPPSALA
UNIVERSITET

Model Extensions

- **Time-stable covariates**
- **Time-dependent covariates**
- **Dual-trajectory modeling** (Nagin & Tremblay 2001; Nagin 2005 chp 8)
  - designed to analyze the developmental course of 2 distinct but related outcomes/time periods
- GBTM and **propensity score matching** (Haviland 2007, 2008, Haviland & Nagin 2005)- for causal inference
- Group-based **multi-trajectory modeling** (Nagin et al. 2018)
  - For multiple indicators
  - Example: Rod et al. *The Lancet* 2020

Example: **Dual-trajectory modeling**

Analyze 2 distinct but related outcomes

A. Probability of delinquency group conditional on opposition group			
	Opposition trajectory group		
Property delinquency group	Low	Moderate	High
Low 1	0.54	0.29	0.23
Low 2	0.30	0.41	0.34
Rising	0.15	0.19	0.26
Chronic	0.01	0.11	0.17
	1.00	**1.00**	**1.00**

Example: **Multi-trajectory modeling**

Identifies latent clusters of individuals following similar trajectories across multiple indicators of an outcome of interest

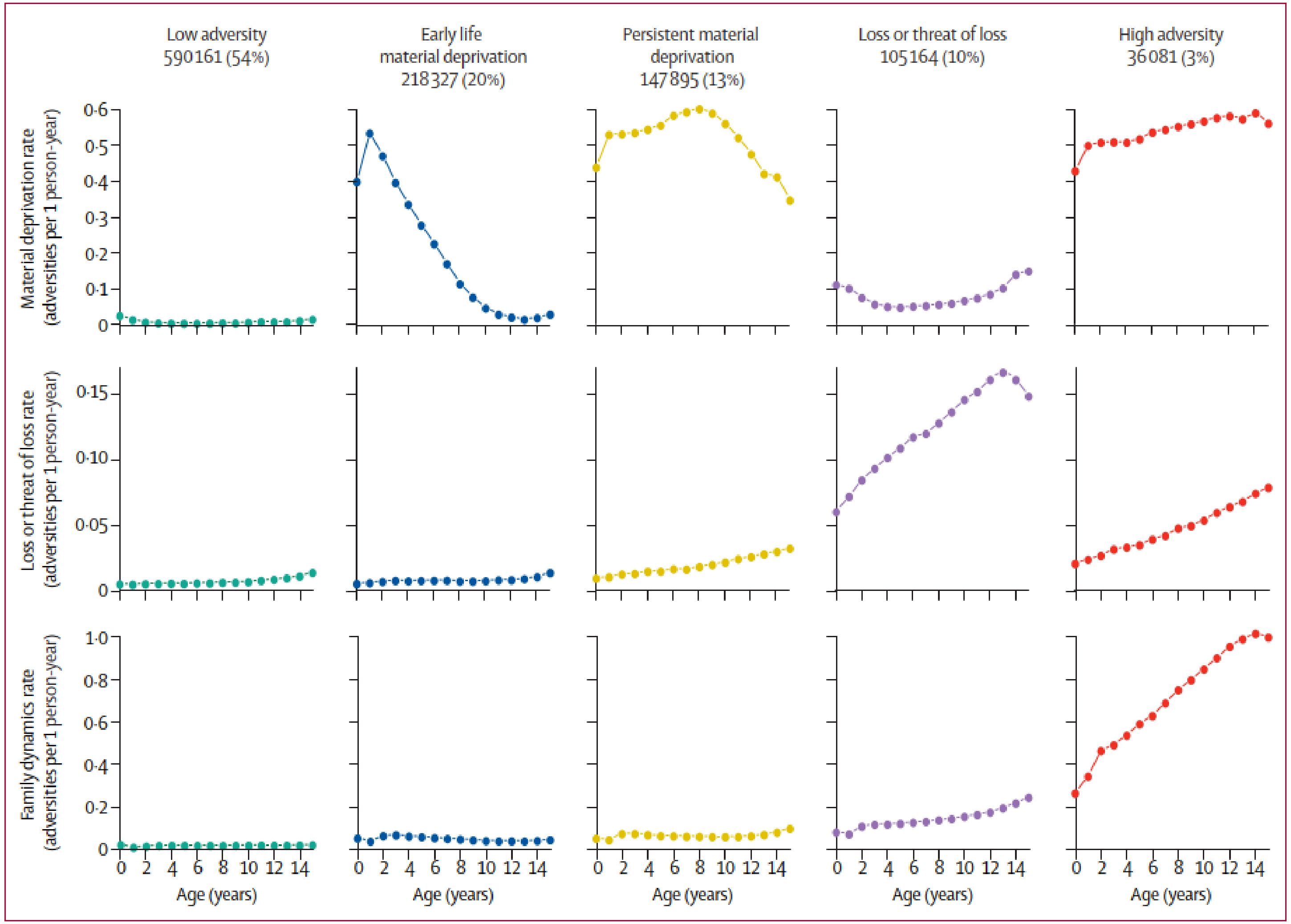


Figure 2: Estimated trajectory groups of childhood adversities among Danish children
1097628 Danish children were divided into the five estimated trajectory groups of childhood adversities.

UPPSALA
UNIVERSITET

Software packages

Stata

- *traj*
- Jones & Nagin. 2013, *Soc Meths & Resch*

SAS

- *Traj*
- Jones, Nagin, Roeder. 2001, *Soc Meths & Resch*

R

- lcmm

Article

Sociological Methods & Research
00(0) 1-6

DOI: 10.1177/0049124113503141
smr.sagepub.com
SAGE

A Note on a Stata Plugin for Estimating Group-based Trajectory Models

Bobby L. Jones[1] and Daniel S. Nagin[2]

UPPSALA
UNIVERSITET

Viewer - help traj

File Edit History Help

help traj

help traj

Dialog Also see

{*26Sep2021}**help traj**

---

traj: Trajectory modeling
trajplot: Plot results
multtrajplot: Plot multi-trajectory model results
trajstart: Generate random start values

Description

traj uses a discrete mixture model to model longitudinal data. This model accommodates data groups with different parameter values for each group distribution. Groupings may identify distinct subpopulations. Alternatively, groupings may represent components of an approximation to an unknown and possibly complex data distribution.

Examples

1. Censored normal (cnorm) model

2. Variability (sigma) by group option - cnorm model

3. Zero-inflated Poisson (zip) model

4. Logistic (logit) model

In my research

UPPSALA
UNIVERSITET

References

1. Jones BL, Nagin DS, Roeder K. A SAS procedure based on mixture models for estimating developmental trajectories. *Sociological Methods & Research* 2001; 29(3); 374.393.
2. Muthén B: Latent variable analysis: growth mixture modeling and related techniques for longitudinal data; in Kaplan D (ed): The Sage Handbook of Quantitative Methodology for the Social Sciences. Newbury Park, Sage, 2004, pp 345–368.
3. Nagin DS: Group-Based Modeling of Development. Cambridge, Harvard University Press, 2005.
4. Nagin DS, Odgers CL: Group-based trajectory modeling in clinical research. *Annu Rev Clin Psychol* 2010;6:109–138.
5. Nagin DS. Group-based trajectory modeling: An overview. *Ann Nutr Metab* 2015;65:205-210.
6. Nagin DS et al. Group-based multi-trajectory modeling. *Statistical Methods in Medical Research* 2018;27(7):2015-2023.
7. Haviland A et al. Combining Group-Based Trajectory Modeling and Propensity Score Matching for Causal Inferences in Nonexperimental Longitudinal Data. *Developmental Psychology*. 2008.
8. https://www.andrew.cmu.edu/user/bjones/