An Evaluative Comparison of Random Coefficient Growth Models for Individual Development

May 26, 2007
Association for Psychological Science

Daniel J. Bauer
L.L. Thurstone Psychometric Laboratory
Department of Psychology
The University of North Carolina at Chapel Hill

H. Luz McNaughton Reyes
Department of Health Behavior and Health Education
The University of North Carolina at Chapel Hill

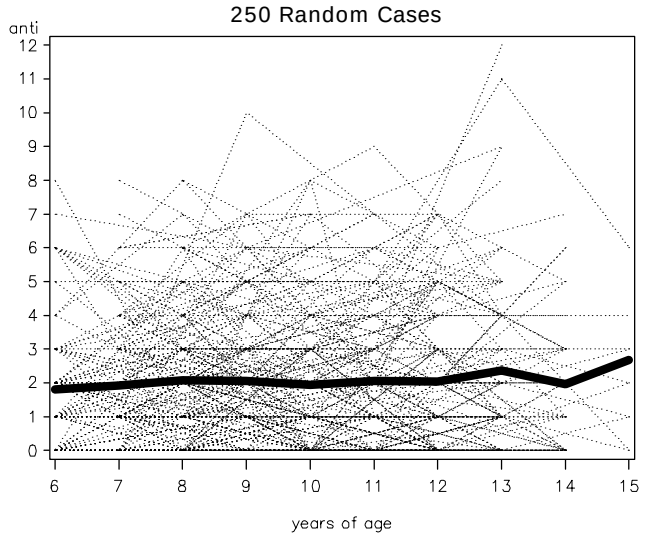
Overview of Talk

- Motivating example
- Definition of Random Coefficient Growth Model (RCGM)
- Three basic types of RCGMs
- Controversies in model selection and interpretation

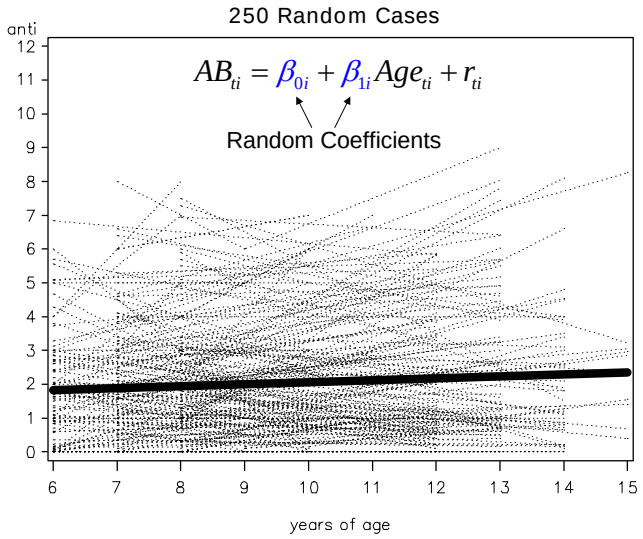
Motivating Example: Antisocial Behavior (AB)

- Concerns:
  - What do trajectories of change in AB look like from age 6 to age 15?
  - Are there sex-differences in these developmental trends?
  - Do supportive home environments protect against increases in AB?
  - Can AB be predicted by poor academic performance?
- Sample:
  - 894 children assessed biennially from 1986 to 1992 as part of the NLSY-Child Sample.
  - Between 6 and 8 years old in 1986.
- Measures:
  - Antisocial Behavior: Sum of 6 items from BPI
  - Early Home Environment: HOME-SF cognitive and emotional support scores from first assessment
  - Academic Performance: PIAT Math scores

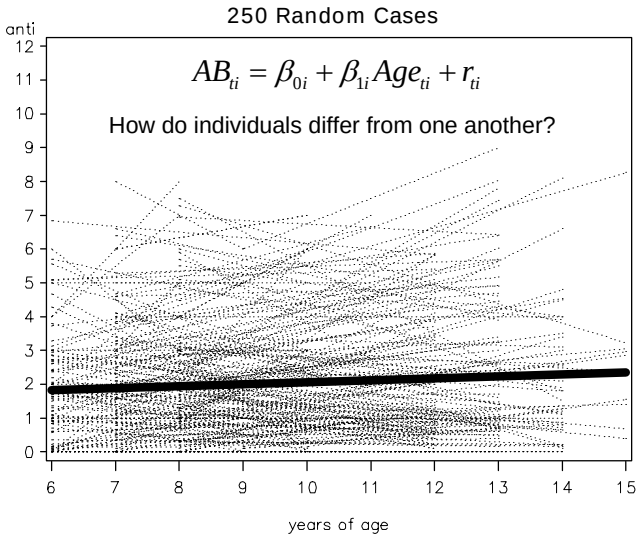
Raw Data



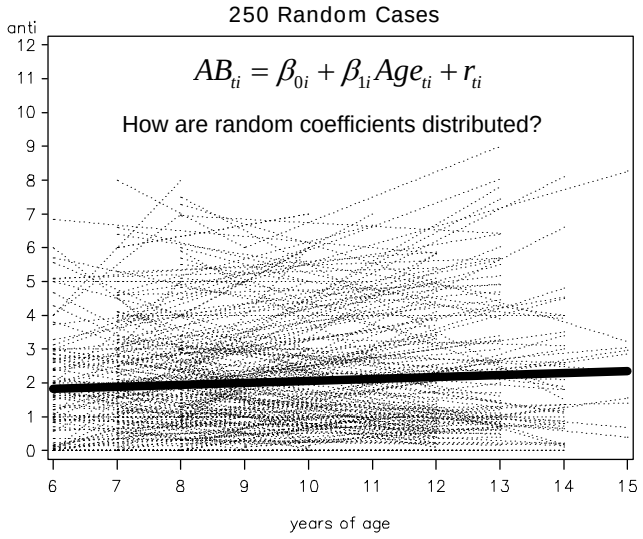
The Idea Behind RCGMs



The Idea Behind RCGMs



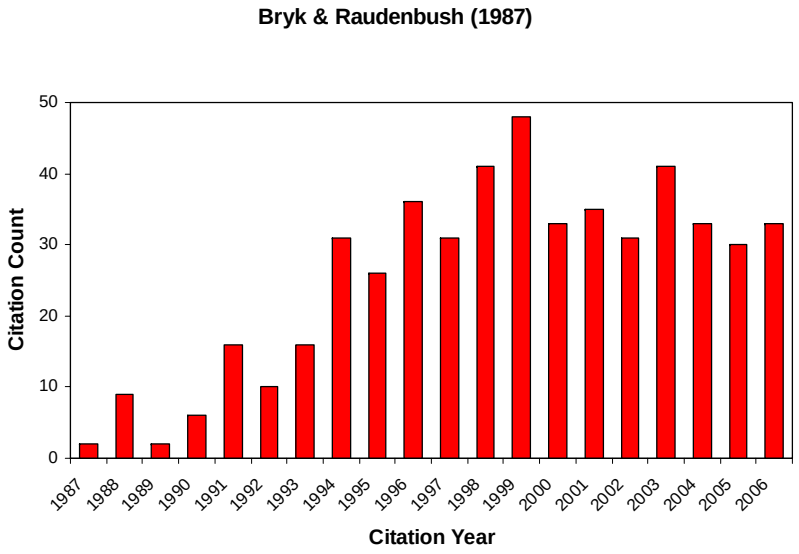
The Idea Behind RCGMs



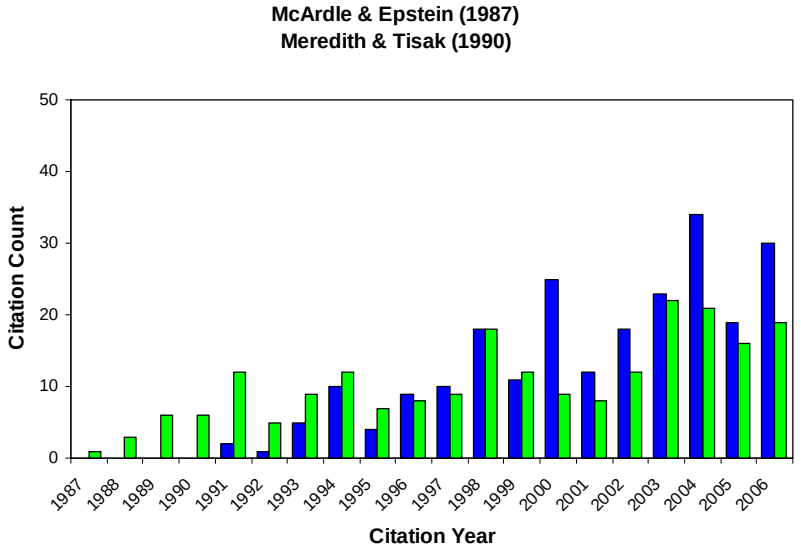
Three Types of RCGMs

1. RCN: Models that assume the random coefficients are normally distributed (conditional on predictors)
   - Latent curve/trajectory/growth models, HLM, multilevel growth models, mixed-effects models

Trends in Use of RCN Models (HLMs)



Trends in Use of RCN Models (LCMs)



RCN

- Within-Person Model:

$$AB_{ti} = \beta_{0i} + \beta_{1i}\left(Age_{ti} - 6\right) + r_{ti} \qquad r_{ti} \sim N\left(0, \sigma^2\right)$$

- Between-Person Model:

$$\beta_{0i} = \gamma_{00} + u_{0i} \qquad \begin{pmatrix} u_{0i} \\ u_{1i} \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \tau_{00} & \\ \tau_{10} & \tau_{11} \end{pmatrix}\right]$$
$$\beta_{1i} = \gamma_{10} + u_{1i}$$

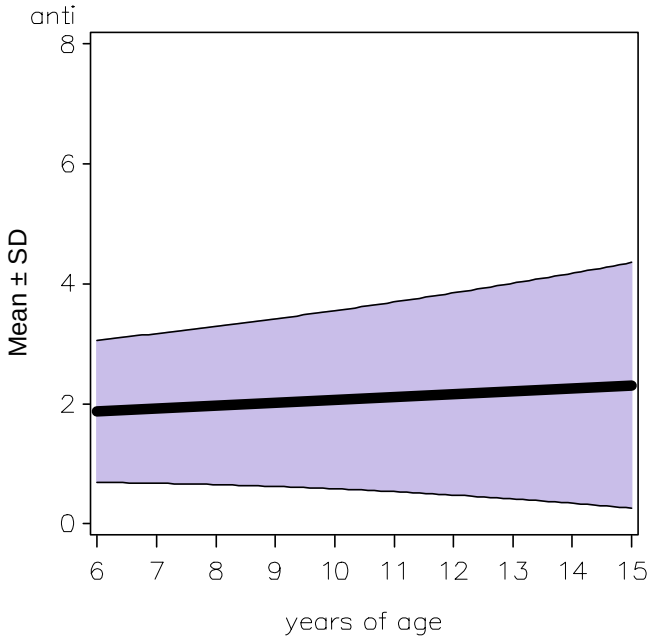
- Assumes continuous individual differences in change over time.

RCN Fit to Antisocial Data

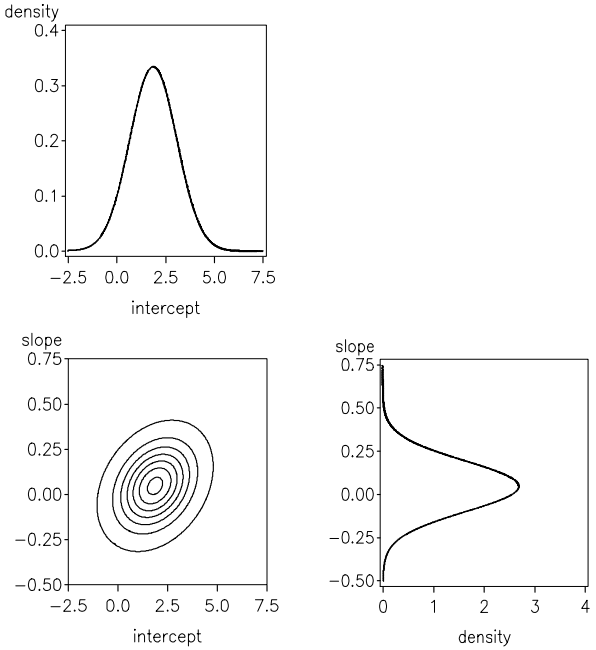
Parameter	Estimate (SE)
Fixed effects	
Intercept (γ_{00})	1.87 (0.07)**
Age (γ_{10})	0.05 (0.01)**
Variance / Covariance Parameters	
Intercept (τ_{00})	1.43 (0.25)**
Age (τ_{11})	0.02 (0.01)*
Covariance (τ_{10})	0.05 (0.04)
Residual (σ^2)	2.09 (0.14)**

Note. Robust standard errors reported.
* p < .05 ; ** p < .01

Graphical Depiction of RCN



Graphical Depiction of RCN



Adding Predictors to RCN

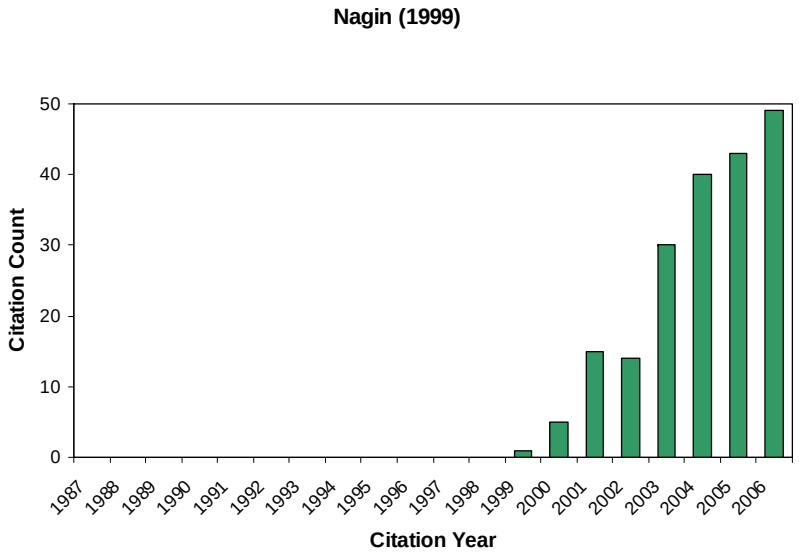
Parameter	Estimate (SE)
Prediction of Intercepts (β_{0i})	
Intercept	1.18 (0.14)**
Sex	**0.85** (0.13)**
Home	**-0.22** (0.04)**
Prediction of Slopes (β_{1i})	
Intercept	0.12 (0.03)**
Sex	0.01 (0.03)
Home	-0.01 (0.01)
Time Varying Covariates	
Math	**-0.17** (0.04)**
Variance / Covariance Parameters	
Intercept	1.05 (0.16)**
Age	0.02 (0.01)*
Covariance	0.06 (0.03)*
Residual	2.11 (0.06)**

* p<.05; ** p<.01

Three Types of RCGMs

1. RCN: Models that assume the random coefficients are normally distributed (conditional on predictors)
   - Latent curve/trajectory/growth models, HLM, multilevel growth models, mixed-effects models
2. RCD: Models that assume the random coefficients are discretely distributed across K groups
   - Semi-parametric group-based trajectory method, non-parametric random coefficient model, latent class growth analysis, latent class regression

Trends in Use of RCD Models (SPGBM)



RCD

- Within-Person Model:

$$AB_{ti} = \beta_{0i} + \beta_{1i}\left(Age_{ti} - 6\right) + r_{ti} \qquad r_{ti} \sim N\left(0, \sigma^2\right)$$

- Between-Person Model:

$$\beta_{0i} = \gamma_{00(k)} \text{ if } C_i = k \qquad C_i = 1, 2, \ldots, K$$
$$\beta_{1i} = \gamma_{10(k)} \text{ if } C_i = k \qquad P(C_i = k) = \pi_{(k)}$$

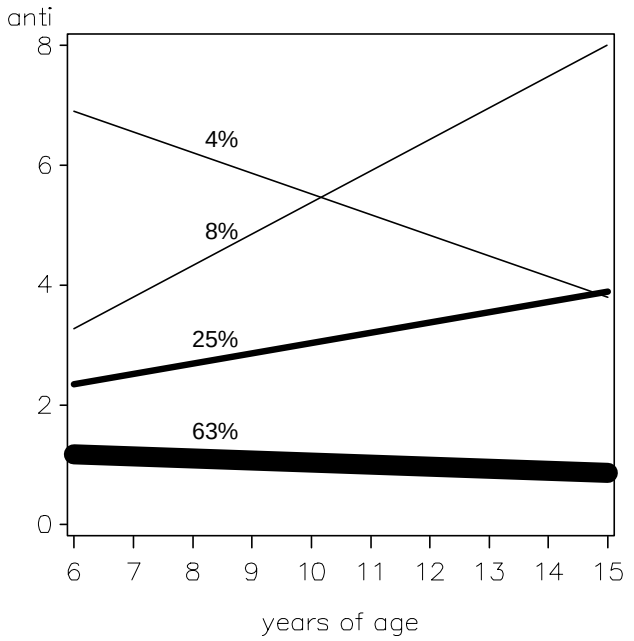
- Assumes individual differences in change over time are discretely distributed – K types of trajectories.

RCD Fit to Antisocial Data

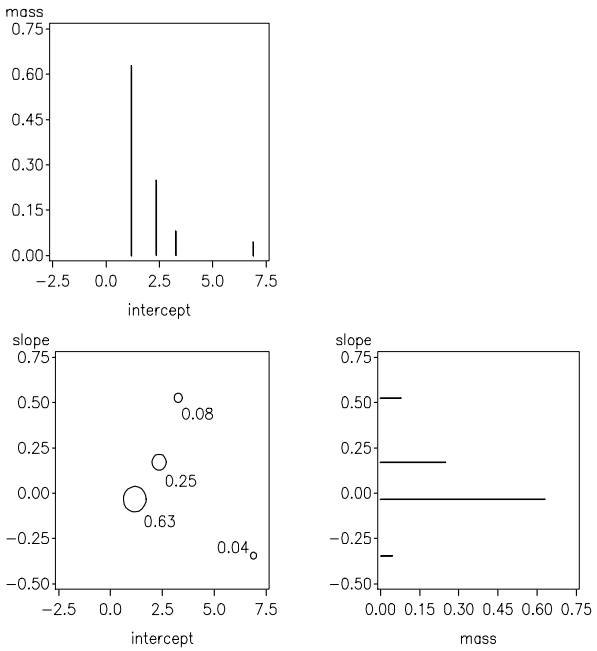
	Latent Class			
Parameter	"High Increasing"	"Low Increasing"	"High Declining"	"Abstaining"
Class Size				
Class probability ($\pi_{(k)}$)	.08	.25	.04	.63
Sample *N*	70	222	39	562
Fixed effects				
Intercept ($\gamma_{00(k)}$)	3.27 (0.91)**	2.35 (0.22)**	6.90 (1.95)**	1.17 (0.07)**
Age ($\gamma_{10(k)}$)	0.53 (0.20)**	0.17 (0.04)**	-0.34 (0.37)	-0.03 (0.01)*
Variance / Covariance Parameters				
Residual (σ^2)	2.02 (0.10)**	2.02 (0.10)**	2.02 (0.10)**	2.02 (0.10)**

Note. Robust standard errors reported.
* p < .05 ; ** p < .01

Graphical Depiction of RCD



Graphical Depiction of RCD



Adding Predictors to RCD

	Latent Class			
Parameter	"High Increasing"	"Low Increasing"	"High Declining"	"Abstaining"
Class Size[a]				
Class probability	.08	.28	.05	.59
Sample *N*	71	236	39	498
Between-Class Effects[b]				
Sex	**4.25** (2.15,8.43)	**3.70** (2.30,5.94)	**9.70** (3.15,29.88)	N/A
Home	**0.59** (0.47,0.72)	**0.65** (0.56,0.75)	**0.68** (0.48,0.96)	N/A
Within-Class Effects				
Intercept	3.08 (0.22)**	2.23 (0.14)**	6.77 (0.33)**	-0.91 (0.11)**
Age	0.56 (0.05)**	0.17 (0.03)**	-0.32 (0.07)*	0.03 (0.03)
Math	**-0.11** (0.04)**	**-0.11** (0.04)**	**-0.11** (0.04)**	**-0.11** (0.04)**
Variance / Covariance Parameters				
Residual	1.96 (0.04)**	1.96 (0.04)**	1.96 (0.04)**	1.96 (0.04)**

* p < .05 ; ** p < .01

[a] Based on estimated posterior probabilities

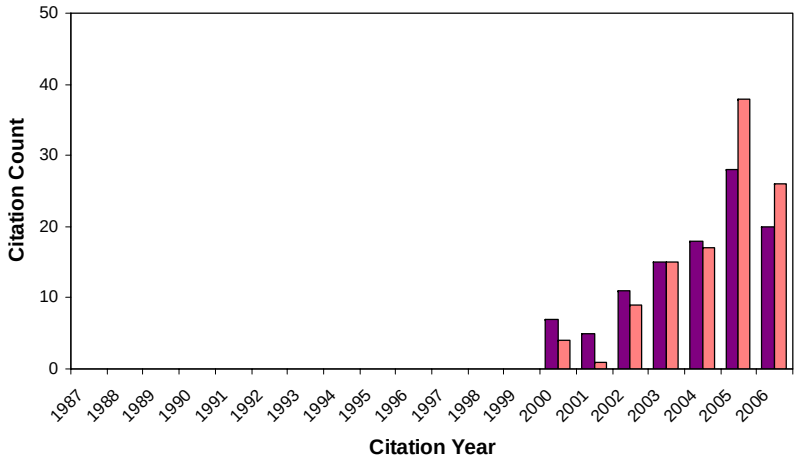
[b] Odds Ratio (95% CI); Reference class is Abstaining

Three Types of RCGMs

1. RCN: Models that assume the random coefficients are (conditionally) normally distributed
   - Latent curve/trajectory/growth models, HLM, multilevel growth models, mixed-effects models
2. RCD: Models that assume the random coefficients are discretely distributed across K latent groups
   - Semi-parametric group-based trajectory method, non-parametric random coefficient model, latent class growth analysis, latent class regression
3. RCNM: Models that assume the random coefficients are from a (conditional) normal mixture distribution with K latent groups
   - General growth mixture models

Trends in Use of RCNM Models (GMMs)

Muthen & Shedden (1999)
Muthen & Muthen (2000)



RCNM

- Within-Person Model:

$$AB_{ti} = \beta_{0i} + \beta_{1i}\left(Age_{ti} - 6\right) + r_{ti} \qquad r_{ti} \sim N\left(0, \sigma^2_{(k)}\right)$$

- Between-Person Model:

$$\beta_{0i} = \gamma_{00(k)} + u_{0i(k)} \text{ if } C_i = k$$
$$\beta_{1i} = \gamma_{10(k)} + u_{1i(k)} \text{ if } C_i = k$$

$$\begin{pmatrix} u_{0i(k)} \\ u_{1i(k)} \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \tau_{00(k)} & \\ \tau_{10(k)} & \tau_{11(k)} \end{pmatrix}\right]$$

$$C_i = 1, 2, \ldots, K$$
$$P(C_i = k) = \pi_{(k)}$$

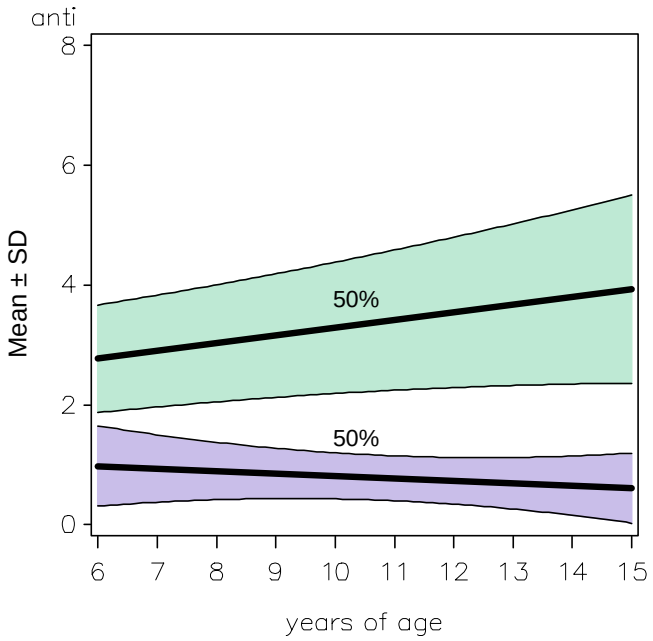
- Assumes K discrete trajectory groups, within which individual differences in change over time are continuously distributed.

RCNM Fit to Antisocial Data

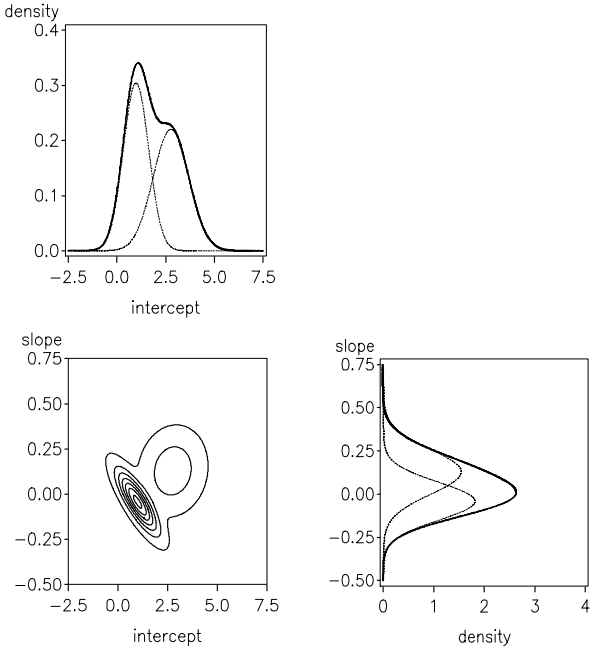
Parameter	Latent Class "Increasing"	Latent Class "Decreasing"
Class Size		
Class probability ($\pi_{(k)}$)	.50	.50
Sample N	445	449
Fixed effects		
Intercept ($\gamma_{00(k)}$)	2.78 (0.13)**	0.98 (0.09)**
Age ($\gamma_{10(k)}$)	0.13 (0.02)**	-0.04 (0.02)**
Variance / Covariance Parameters		
Intercept ($\tau_{00(k)}$)	0.83 (0.42)*	0.43 (0.10)**
Age ($\tau_{11(k)}$)	0.02 (0.02)	0.01 (0.004)*
Covariance ($\tau_{10(k)}$)	0.01 (0.07)	-0.06 (0.02)**
Residual ($\sigma^2_{(k)}$)	3.55 (0.32)**	0.62 (0.07)**

Note. Robust standard errors reported.
* p < .05 ; ** p < .01

Graphical Depiction of RCNM



Graphical Depiction of RCNM



Adding Predictors to RCNM

Parameter	Latent Class "Increasing"	Latent Class "Decreasing"
Class Size[a]		
Class probability	.51	.49
Sample *N*	426	417
Between-Class Effects[b]		
Sex	**2.82** (1.93,4.11)	N/A
Home	**0.66** (0.58,0.75)	N/A
Within-Class Prediction of Intercepts (β_{0i})		
Intercept	1.89 (0.24)**	0.76 (0.10)
Sex	**0.62** (0.24)**	**0.31** (0.12)**
Home	-0.06 (0.04)	-0.06 (0.04)
Within-Class Prediction of Slopes (β_{1i})		
Intercept	0.26 (0.05)**	0.01 (0.02)
Sex	-0.01 (0.05)	**-0.05** (0.03)*
Home	0.01 (0.01)	0.01 (0.01)
Time Varying Covariates		
Math	**-0.30** (0.07)**	-0.06 (0.03)
Variance / Covariance Parameters		
Intercept	0.45 (0.34)	0.35 (0.11)**
Age	0.01 (0.02)	0.01 (0.004)**
Covariance	0.05 (0.06)	-0.05 (0.02)**
Residual	3.60 (0.17)**	0.61 (0.05)**

* p < .05 ; ** p < .01
[a] Based on estimated posterior probabilities
[b] Odds Ratio (95% CI); Reference class is Decreasing

Comparison of Three Models

- Overall trends are similar
  - Most children had low levels of AB at age 6.
  - Some children showed increases in AB from age 6 to 15, others showed stable or decreasing AB.
  - Those with highest initial levels of AB tended to show greatest increases over time.
  - Male children, and children from less supportive home environments displayed higher levels of AB.
  - Children scoring poorly on the PIAT-M displayed higher levels of AB.
- Despite general consistency, much controversy over appropriate model

Controversy in the Selection of RCGMs

- 5 recent papers on RCD and RCNM models have been followed by commentaries and rejoinders
  - Bauer & Curran (2003, *Psychological Methods*), commentaries by Cudeck & Henly, Muthen, & Rindskopf.
  - Eggleston, Laub & Sampson (2004, *J. Quant. Criminology*), commentary by Nagin.
  - Nagin & Tremblay (2005, *Annals AAPSS*), commentaries by Maughan and Raudenbush.
  - Nagin & Tremblay (2005, *Criminology*), commentary by Sampson & Laub.
  - Connell & Frye (2006, *Infant & Child Dev*), commentaries by Hoeksma & Kelderman, Muthen, and Stanger.
- Reflects disagreements, misunderstandings about relative merits of different RCGMs.

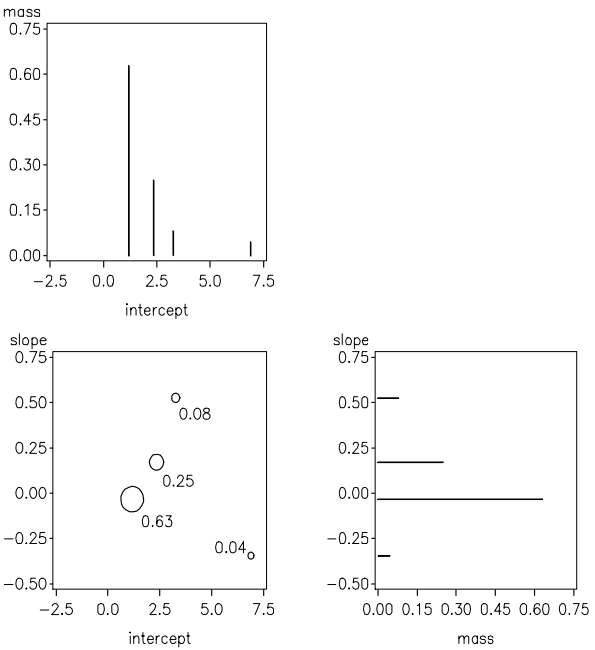
Key Issue: The Interpretation of Groups

- Often, the latent classes estimated from RCD or RCNM models are interpreted as true taxa (i.e., "real" groups).
- This can be problematic
  - Are there 4 groups (RCD), 2 groups (RCNM), or no groups (RCN)?
  - Should groups be strictly homogeneous (RCD) or do we permit within-group variability (RCNM)?
  - Spurious groups can compensate for errors in model specification, e.g., lack of normality of residuals (Bauer & Curran, 2003, 2004).
  - Number/nature of groups can change with minor alterations in model specification, covariates, measurement, or design (Eggleston et al, 2004; Jackson & Sher, 2005, 2006).
  - RCN does permit groups as a function of observed covariates (when strong etiological theory for taxa).

Key Issue: The Interpretation of Groups

- Nagin (2005) advocates use of groups as an approximating device even when groups are artificial.

Graphical Depiction of RCD



Key Issue: The Interpretation of Groups

- Nagin (2005) advocates use of groups as an approximating device even when groups are artificial.
- Drawbacks to Groups:
  - Artificial groups prone to reification.
  - Use of RCD and RCNM models may reduce power.
- Alternatives to Groups:
  - RCN models relatively robust to violation of normality assumption for random effects.
  - Other semiparametric RCGMs (Chen, Zhang & Davidian, 2002; Zhang & Davidian, 2001)
- Perhaps better to focus less on groups and more on overall trends.

Conclusions

- RCGMs offer many conceptual and statistical advantages for modeling individual change
- Many different RCGMs (HLM, LCM, GMM, LCGA, SPGBA, etc) can be organized into three categories
  - RCN assumes continuously distributed individual differences.
  - RCD assumes discretely distributed individual differences.
  - RCNM assumes continuously distributed individual differences within a small number of discrete groups.
- Key issue in choosing between models is utility of groups
  - Groups correspond nicely to taxonomic theory
  - Groups often improve model flexibility and fit.
  - Interpretation of latent groups can be risky.