Modeling between and within-subject variances using mixed effects location scale models for intensive longitudinal data

Donald Hedeker
University of Chicago

hedeker@uchicago.edu

Supported by National Cancer Institute grant P01 CA98262 (Mermelstein, PI) and National Heart Lung and Blood Institute grant R01 HL121330 (Hedeker & Dunton)

Ecological Momentary Assessment (EMA) data

experience sampling and diary methods, intensive longitudinal data

- Subjects provide frequent reports on events and experiences of their daily lives (*e.g.*, 30-40 responses per subject collected over the course of a week or so)
  - electronic diaries: palm pilots, personal digital assistants (PDAs), smart phones

- Capture particulars of experience in a way not possible with more traditional designs
  e.g., allow investigation of phenomena as they happen over time

- Reports could be time-based, following a fixed-schedule, randomly triggered, event-triggered

Data are rich and offer many modeling possibilities!

- person-level and occasion-level determinants of occasion-level responses $\Rightarrow$ potential influence of context and/or environment
  e.g., subject response might vary when alone vs with others

- allows examination of why subjects differ in variability rather than just mean level
  - between-subjects variance
    e.g., subject heterogeneity could vary by gender or age
  - within-subjects variance
    e.g., subject degree of stability could vary by gender or age

Carroll (2003) Variances are not always nuisance parameters, *Biometrics*.

Multilevel (mixed-effects regression) model for measurement y of subject i ($i = 1, 2, \ldots, N$) on occasion j ($j = 1, 2, \ldots, n_i$)

$$y_{ij} \;=\; \boldsymbol{x}'_{ij}\boldsymbol{\beta} + \upsilon_i + \epsilon_{ij}$$

$\boldsymbol{x}_{ij} = p \times 1$ vector of regressors (including a column of ones)

$\boldsymbol{\beta} = p \times 1$ vector of regression coefficients

$\upsilon_i \sim N(0, \sigma^2_\upsilon)$ BS variance

$\epsilon_{ij} \sim N(0, \sigma^2_\epsilon)$ WS variance

Log-linear models for variances

$$\text{BS variance} \quad \sigma^2_{\upsilon_{ij}} = \exp(\boldsymbol{u}'_{ij}\boldsymbol{\alpha}) \quad \text{or} \quad \log(\sigma^2_{\upsilon_{ij}}) = \boldsymbol{u}'_{ij}\boldsymbol{\alpha}$$

$$\text{WS variance} \quad \sigma^2_{\epsilon_{ij}} = \exp(\boldsymbol{w}'_{ij}\boldsymbol{\tau}) \quad \text{or} \quad \log(\sigma^2_{\epsilon_{ij}}) = \boldsymbol{w}'_{ij}\boldsymbol{\tau}$$

- $\boldsymbol{u}_{ij}$ and $\boldsymbol{w}_{ij}$ include covariates (and $\mathbf{1}$)
- subscripts i and j on variances indicate that these change depending on covariates $\boldsymbol{u}_{ij}$ and $\boldsymbol{w}_{ij}$ (and their coefficients)
- exp function ensures a positive multiplicative factor, and so resulting variances are positive

How can WS variables influence BS variance?

$$\sigma^2_{\upsilon_{ij}} = \exp(\boldsymbol{u}'_{ij}\boldsymbol{\alpha})$$

- Do rainy days and Mondays get everyone down?
- Is Tuesday just as bad as Stormy Monday for all?
- Are all kids happy on the last day of school?

Example: strong positive effect of being alone on BS variance of positive and negative mood

$\Rightarrow$ being alone increases subject heterogeneity (or, subjects report more similar mood when with others)

WS variance varies across subjects

$$\sigma^2_{\epsilon_{ij}} = \exp(\boldsymbol{w}'_{ij}\boldsymbol{\tau} + \omega_i) \quad \text{where} \quad \omega_i \sim N(0, \sigma^2_\omega)$$

$$\log(\sigma^2_{\epsilon_{ij}}) = \boldsymbol{w}'_{ij}\boldsymbol{\tau} + \omega_i$$

- ω_i are log-normal subject-specific perturbations of WS variance
- ω_i are "scale" random effects - how does a subject differ in terms of the variation in their data
- υ_i are "location" random effects - how does a subject differ in terms of the mean of their data

Multilevel model of WS variance

$$\log(\sigma^2_{\epsilon_{ij}}) \;=\; \boldsymbol{w}'_{ij}\boldsymbol{\tau} + \omega_i$$

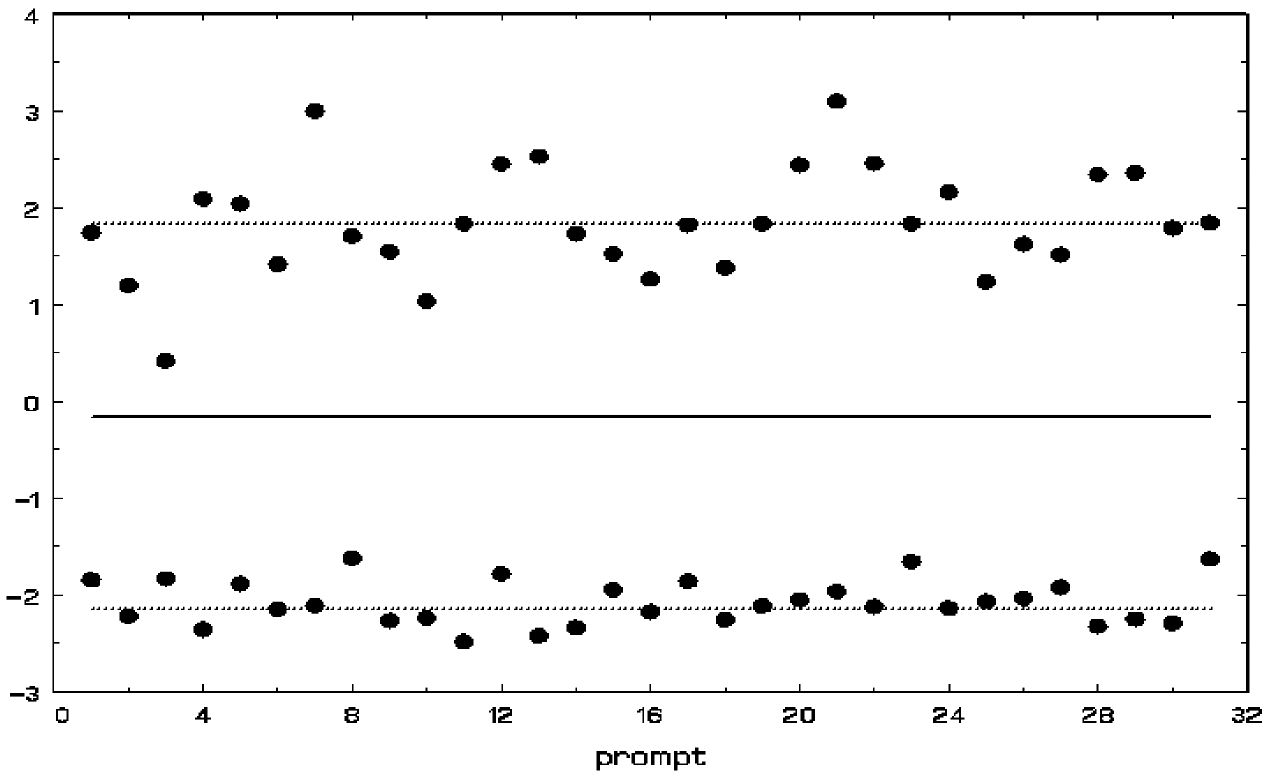
Why not use some summary statistic per subject (say, calculated subject standard deviation S_{y_i}) in a second-stage model?

$$S_{y_i} = \boldsymbol{x}'_i\boldsymbol{\beta} + \epsilon_i$$

latter approach

- treats all standard deviations as if they are equally precise (but some might be based on 2 prompts or 40 prompts)
- does not recognize that these are estimated quantities (underestimation of sources of variation)
- does not allow occasion-varying predictors

$\Rightarrow$ We use multilevel models for mean response, why not for variance?



Model allows covariates to influence

- mean: level of solid line
- BS variance: dispersion of dotted lines
- WS variance: dispersion of points

additional random subject effects on: mean and WS variance

Estimation

- SAS PROC NLMIXED (slow and must provide starting values)

  Hedeker, D., Mermelstein, R.J., & Demirtas, H. (2008). An application of a mixed-effects location scale model for analysis of Ecological Momentary Assessment (EMA) data. *Biometrics, 64*, 627-634, *Supplemental Materials*.

- MIXREGLS freeware (faster and no starting values); also DLL is accessible via R

  Hedeker, D. & Nordgren, R. (2013). MIXREGLS: A program for mixed-effects location scale analysis. *Journal of Statistical Software, 52(12)*, 1-38.

- MIXREGLS via STATA

  Leckie, G. runmixregls - A Program to Run the MIXREGLS Mixed-effects Location Scale Software from within Stata. *Journal of Statistical Software, Code Snippet*, 1-41. Forthcoming.

- Bayesian approach using WinBUGS or JAGS

  Rast, P., Hofer, S. M., & Sparks, C. (2012). Modeling individual differences in within-person variation of negative and positive affect in a mixed effects location scale model using BUGS/JAGS. *Multivariate Behavioral Research, 47*, 177-200.

Ecological Momentary Assessment (EMA) Study of Adolescent Smokers (Mermelstein)

- 461 adolescents (9th and 10th graders); former and current smoking experimenters, and regular smokers

- Carry PDA for a week, answer questions when prompted

  average = 30 answered prompts (range = 7 to 71)

- $\Sigma_i^N n_i = 14,105$ total number of observations

Outcomes: positive and negative affect

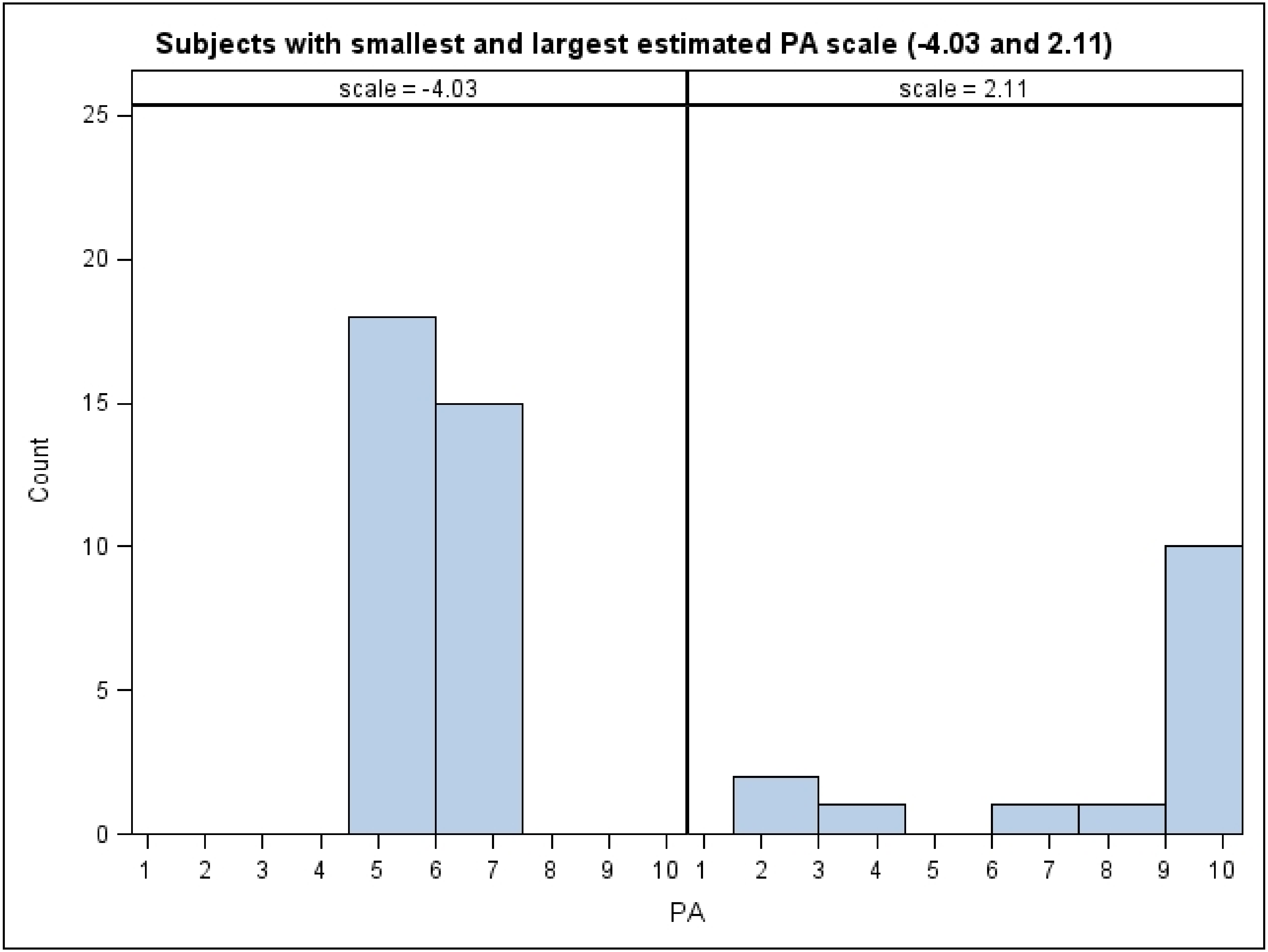
Interest: characterizing determinants of affect level, as well as BS and WS affect heterogeneity

Dependent Variables

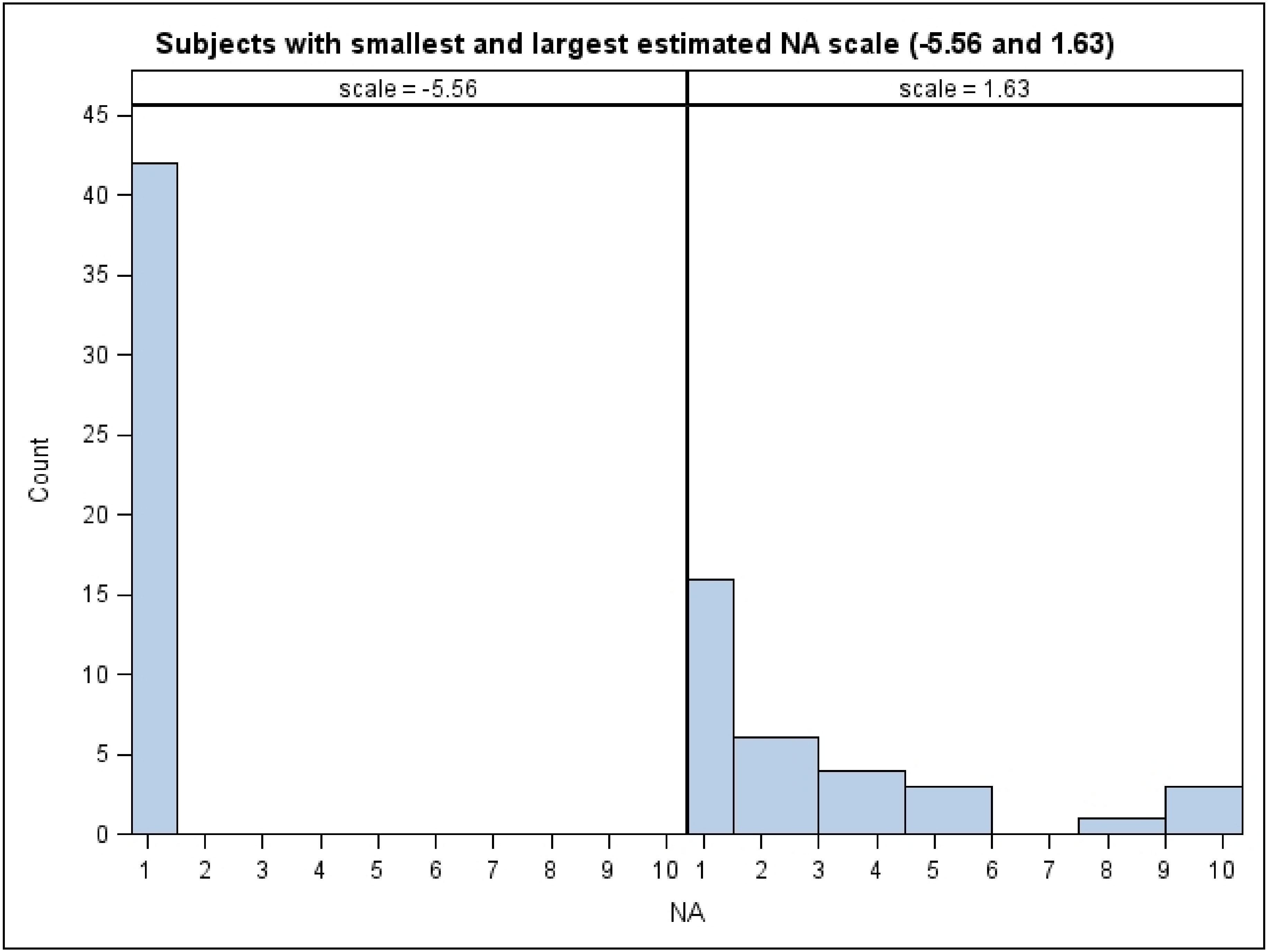
- Positive Affect mood scale (mean=6.797 and sd=1.935)
  - Before signal: I felt Happy
  - Before signal: I felt Relaxed
  - Before signal: I felt Cheerful
  - Before signal: I felt Confident
  - Before signal: I felt Accepted by Others

- Negative Affect mood scale (mean=3.455 and sd=2.253)
  - Before signal: I felt Sad
  - Before signal: I felt Stressed
  - Before signal: I felt Angry
  - Before signal: I felt Frustrated
  - Before signal: I felt Irritable

$\Rightarrow$ items rated on 1 (not al all) to 10 (very much) scale



Subjects with smallest and largest estimated PA scale (-4.03 and 2.11)
scale = -4.03
scale = 2.11
Count
PA
0
5
10
15
20
25
1 2 3 4 5 6 7 8 9 10
1 2 3 4 5 6 7 8 9 10


Subjects with smallest and largest estimated NA scale (-5.56 and 1.63)
scale = -5.56
scale = 1.63
Count
NA
0
5
10
15
20
25
30
35
40
45
1
2
3
4
5
6
7
8
9
10

Subject-level Independent Variables

	mean	std dev	min	max
`Smoker`	.508	.500	0	1
`Male`	.449	.498	0	1

- `Smoker`: gave at least one report of a smoking event in the week of EMA measurement (about half of the subjects)

- `Male`: a bit more females than males in this sample

parameter	Positive Affect estimate	se	$p <$	Negative Affect estimate	se	$p <$
<u>Mean</u>						
Intercept β_0	6.741	.094	.001	3.609	.118	.001
`Male` β_1	.296	.114	.01	-.603	.136	.001
`Smoker` β_2	-.188	.115	.10	.283	.136	.04
<u>WS variance</u>						
Intercept τ_0	.706	.060	.001	.824	.077	.001
`Male` τ_1	-.276	.072	.001	-.453	.093	.001
`Smoker` τ_2	.078	.071	.27	.238	.092	.01
<u>BS variance</u>						
Intercept α_0	.292	.102	.004	.908	.067	.001
`Male` α_1	-.103	.121	.40	-.319	.113	.005
`Smoker` α_2	.198	.120	.10	.111	.110	.31
<u>Scale</u>						
BS variance of scale σ^2_ω	.506	.039	.001	.908	.065	.001
covariance $\sigma_{\upsilon\,\omega}$	-.361	.046	.001	.661	.073	.001

What about smoking?

- `Smoker` does not consider smoking level (just whether or not a subject provided at least one smoking event)

- 234 with smoking events: average=5, median=3, range = 1 to 42

- Perhaps, smoking level needs to be considered

- `PropSmk` = proportion of occasions (both random prompts and smoking events) that were smoking events

  `PropSmk` = n_smk / (n_smk + n_random)

Model with `Smoker` and `Psmk`

`PropSmk` = n_smk / (n_smk + n_random)

N=234 with n_smk > 0 (and `Smoker` = 1)

min = .014, 25% quartile = .05, median = .08, 75% quartile = .18

`Psmk` = `PropSmk` - min(`PropSmk`)

Model: $\texttt{Mood}_{ij} = \beta_0 + \beta_1\texttt{Smoker} + \beta_2\texttt{Psmk} + \ \ldots \ + \upsilon_i + \epsilon_{ij}$

subject	`Smoker`	`Psmk`	mean (with other covariates = 0)
non-smoker	0	0	β_0
min smoker	1	0	$\beta_0 + \beta_1$
light smoker	1	.05	$\beta_0 + \beta_1 + .036\beta_2$
medium smoker	1	.08	$\beta_0 + \beta_1 + .066\beta_2$
high smoker	1	.18	$\beta_0 + \beta_1 + .166\beta_2$

$\Rightarrow$ piecewise linear model for means

Similar models for BS and WS variance

BS Variance Model: $\exp(\alpha_0 + \alpha_1 \texttt{Smoker} + \alpha_2 \texttt{Psmk} + \ldots)$

WS Variance Model: $\exp(\tau_0 + \tau_1 \texttt{Smoker} + \tau_2 \texttt{Psmk} + \ \ldots \ + \omega_i)$

subject	`Smoker`	`Psmk`	BS variance	WS variance
non-smoker	0	0	$\exp(\alpha_0)$	$\exp(\tau_0 + \omega_i)$
min smoker	1	0	$\exp(\alpha_0 + \alpha_1)$	$\exp(\tau_0 + \tau_1 + \omega_i)$
light smoker	1	.036	$\exp(\alpha_0 + \alpha_1 + .036\alpha_2)$	$\exp(\tau_0 + \tau_1 + .036\tau_2 + \omega_i)$
med smoker	1	.066	$\exp(\alpha_0 + \alpha_1 + .066\alpha_2)$	$\exp(\tau_0 + \tau_1 + .066\tau_2 + \omega_i)$
high smoker	1	.166	$\exp(\alpha_0 + \alpha_1 + .166\alpha_2)$	$\exp(\tau_0 + \tau_1 + .166\tau_2 + \omega_i)$

Note: other covariates set to zero

	Positive Affect			Negative Affect		
parameter	estimate	se	$p <$	estimate	se	$p <$
Mean						
Intercept β_0	6.740	.094	.001	3.607	.117	.001
`Male` β_1	.299	.114	.01	-.599	.135	.001
`Smoker` β_2	-.192	.141	.18	.462	.168	.007
`PSmk` β_3	.018	.742	.98	-1.530	.791	.054
WS variance						
Intercept τ_0	.704	.059	.001	.820	.077	.001
`Male` τ_1	-.272	.071	.001	-.444	.092	.001
`Smoker` τ_2	.157	.086	.07	.407	.112	.001
`Psmk` τ_3	-.693	.430	.11	-1.446	.554	.01
BS variance						
Intercept α_0	.293	.102	.004	.800	.100	.001
`Male` α_1	-.115	.123	.35	-.319	.115	.006
`Smoker` α_2	.157	.149	.30	.183	.135	.18
`Psmk` α_3	.370	.812	.65	-.657	.653	.31
Scale						
BS variance of scale σ_ω^2	.503	.038	.001	.893	.064	.001
covariance $\sigma_{\upsilon\,\omega}$	-.356	.047	.001	.647	.071	.001

- Previous analyses focused on one measurement wave and the effect of smoking level on mood variance from random prompts (between-subjects or cross-sectional effect)

- What about as subjects change their own level of smoking? (within-subjects or longitudinal effect)

- What about smoking-attributable change in mood? (mood responses from smoking events)

EMA Study of Adolescents (Mermelstein, NCI)

- 461 adolescents (9th and 10th graders; 55% female); former and current smoking experimenters, and regular smokers
- Carry PDA for a week, answer questions when randomly prompted, or event-record when smoking (mutually exclusive)
- baseline, 6-month, 15-month, 2-year, and 5-year follow-ups

Interest: characterizing determinants of change in positive and negative affect associated with smoking events, especially across time

$\Rightarrow$ analysis of 158 subjects with two or more waves, where at each wave subject had two or more smoking events

158 subjects with two or more waves
at each wave subject had two or more smoking events

- total of 4,727 smoking events

- 65, 30, 33, 30 subjects had data at two, three, four and five waves

- number of subjects across waves:
  126 (baseline), 93 (6 mo), 95 (15 mo), 101 (2 yr), and 87 (5 yr)

- average number of smoking events across waves:
  6.90 (range = 2 to 42)
  7.53 (2 to 32)
  9.74 (2 to 43)
  10.14 (2 to 49)
  13.90 (2 to 64)

Dependent Variables - mood reports for smoking events

- Positive Affect (`PA`) mood scale (5 items)
  - Before smoking I felt: Happy, Relaxed, Cheerful, Confident, Accepted by Others
- Negative Affect (`NA`) mood scale (5 items)
  - Before smoking I felt: Sad, Stressed, Angry, Frustrated, Irritable
- items rated on 1 (not al all) to 10 (very much) scale
- also rated for "Now after smoking: I feel"
- difference (now-before) is measure of reported mood change associated with smoking
- `PA` mood change averages = .75, .54, .34, .41, .41 across waves
- `NA` mood change averages = -.46, -.45, -.33, -.44, -.32 across waves

Mixed Model for the mood y of subject i ($i = 1, 2, \ldots, N$ subjects) at occasion j ($j = 1, 2, \ldots, n_i$ smoking events):

$$\begin{aligned} y_{ij} \; = \; & (\beta_0 + \upsilon_{0i}) + (\beta_1 + \upsilon_{1i})\texttt{Wave}_j + \beta_2\texttt{Male}_i \\ & +\beta_3\texttt{AvgRate}_i + \beta_4\texttt{DevRate}_{ij} + \epsilon_{ij} \end{aligned}$$

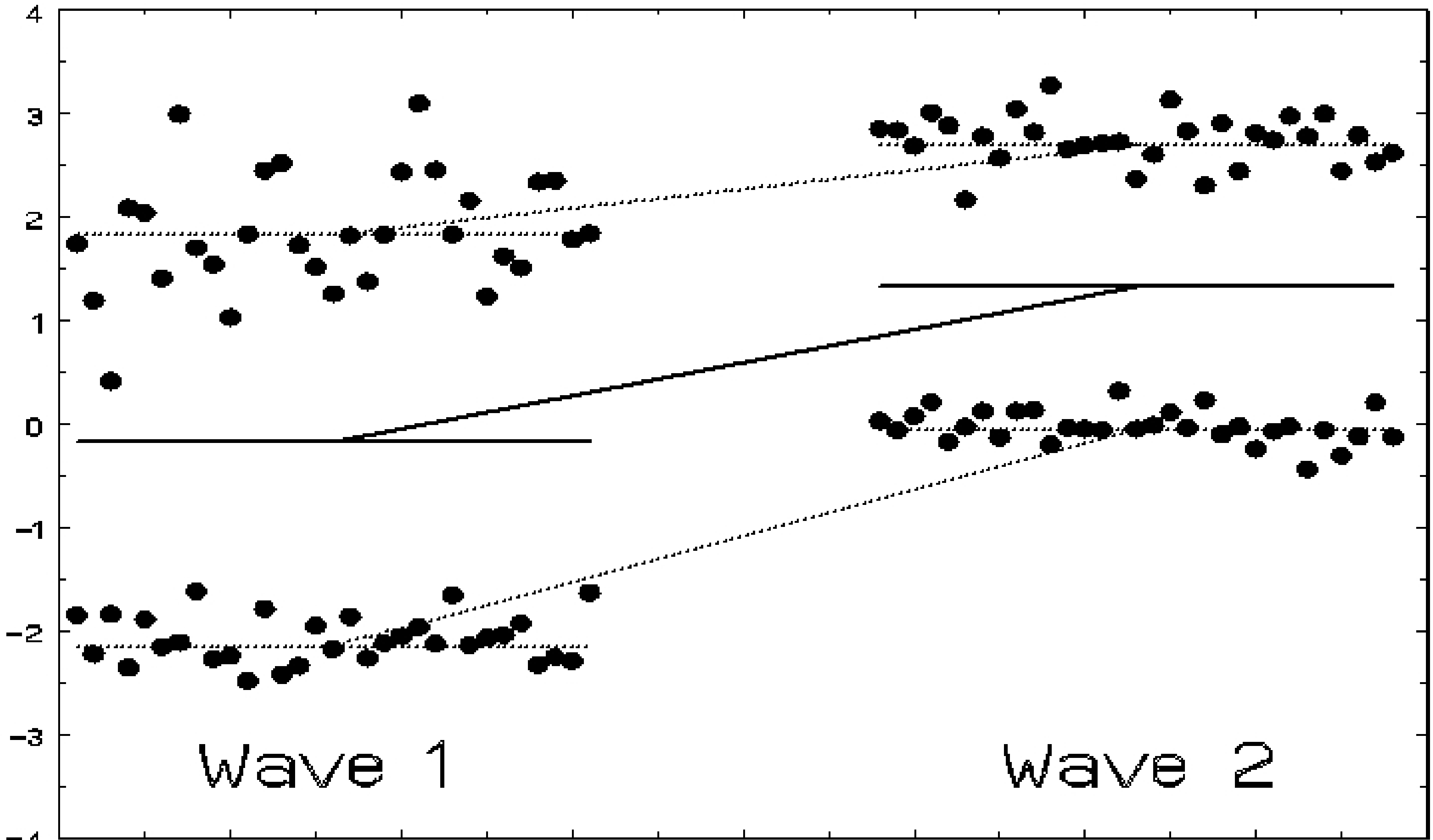
- $\texttt{Wave}_j$ (0=baseline, .5=6 mos, 1.25=15 mos, 2=2yrs, 5=5yrs)
- $\texttt{Male}_i$ (0=female, 1=male)
- Smoking level
  - $\texttt{SmkRate}_{ij}$ = per wave daily smoking rate (ln units)
  - BS version $\texttt{AvgRate}_i$ = subject average of $\texttt{SmkRate}_{ij}$
  - WS version $\texttt{DevRate}_{ij} = (\texttt{SmkRate}_{ij} - \texttt{AvgRate}_i)$ = per wave deviation in the daily smoking rate

Error variance model $\epsilon_{ij} \sim N(0, \sigma^2_\epsilon)$ WS variance

$$\log(\sigma^2_{\epsilon_{ij}}) \;=\; \tau_0 + \tau_1 \texttt{Wave}_j + \tau_2 \texttt{Male}_i + \tau_3 \texttt{AvgRate}_i + \tau_4 \texttt{DevRate}_{ij} + \omega_i$$

log-linear model of within-subject variance, with subject-specific perturbation $\omega_i \sim N(0, \sigma^2_\omega)$

- WS variance follow a log-normal distribution at the subject level
- skewed nonnegative nature of log-normal makes it a reasonable choice for representing variances
- random scale effect ω_i allowed to be correlated with random intercept υ_{0i} and trend υ_{1i}



- population intercept and trend (solid line)
- random intercept and trend for 2 subjects (dotted lines)
- error variance varies across time and subjects (random scale)

Smoking-related Positive and Negative Affect Change

estimates, standard errors (se), and p-values

	Positive Affect			Negative Affect		
Mean Model	est	se	$p <$	est	se	$p <$
Intercept β_0	.691	.110	.001	-.432	.093	.001
Wave β_1	-.013	.017	.44	.004	.013	.78
Male β_2	.129	.083	.13	-.057	.070	.41
AvgRate β_3	-.169	.060	.006	.071	.053	.19
DevRate β_4	-.161	.030	.001	.059	.027	.03
Error Var Model	est	se	$p <$	est	se	$p <$
Intercept τ_0	.921	.172	.001	1.043	.210	.001
Wave τ_1	-.162	.017	.001	-.121	.018	.001
Male τ_2	.210	.153	.172	.215	.193	.27
AvgRate τ_3	-.226	.106	.034	-.337	.133	.012
DevRate τ_4	-.322	.049	.001	-.319	.055	.001

Smoking-related Positive and Negative Affect Change

estimates, standard errors (se), and p-values

Random effect	Positive Affect			Negative Affect		
(co)variances	est	se	$p <$	est	se	$p <$
Intercept $\sigma^2_{\upsilon_0}$	.284	.062	.001	.125	.040	.002
Wave $\sigma^2_{\upsilon_1}$	.014	.004	.001	.003	.002	.12
Scale σ^2_{ω}	.752	.103	.001	1.26	.167	.001
Int, Wave $\sigma_{\upsilon_0 \upsilon_1}$	-.043	.014	.003	-.010	.007	.18
Int, Scale $\sigma_{\upsilon_0 \omega}$	.213	.057	.001	-.208	.052	.001
Wave, Scale $\sigma_{\upsilon_1 \omega}$	-.004	.015	.77	.011	.013	.39

Second or third thoughts?

- analysis treats observations (level-1) within subjects (level-2)

$$y_{ij} = (\beta_0 + \upsilon_{0i}) + (\beta_1 + \upsilon_{1i})\texttt{Wave}_j + \beta_2\texttt{Male}_i + \beta_3\texttt{AvgRate}_i + \beta_4\texttt{DevRate}_{ij} + \epsilon_{ij}$$

$$\sigma^2_{\epsilon_{ij}} = (\tau_0 + \tau_1\texttt{Wave}_j + \tau_2\texttt{Male}_i + \tau_3\texttt{AvgRate}_i + \tau_4\texttt{DevRate}_{ij} + \omega_i)$$

- however, observations (level-1) are nested within waves (level-2) within subjects (level-3)

3-level Model of Smoking-related Positive and Negative Affect Change; estimates, standard errors (se), and p-values

	Positive Affect			Negative Affect		
Mean Model	est	se	$p <$	est	se	$p <$
Intercept β_0	.708	.106	.001	-.447	.091	.001
Wave β_1	-.020	.016	.22	.002	.013	.90
Male β_2	.119	.082	.15	-.057	.069	.41
AvgRate β_3	-.174	.059	.004	.083	.050	.10
DevRate β_4	-.081	.052	.12	.071	.039	.08

Error Var Model	est	se	$p <$	est	se	$p <$
Intercept τ_0	.893	.174	.001	1.048	.211	.001
Wave τ_1	-.158	.017	.001	-.117	.018	.001
Male τ_2	.218	.156	.16	.235	.193	.22
AvgRate τ_3	-.229	.107	.034	-.361	.132	.007
DevRate τ_4	-.314	.049	.001	-.321	.055	.001

3-level Model of Smoking-related Positive and Negative Affect Change; estimates, standard errors (se), and p-values

Random effect (co)variances	Positive Affect est	se	$p <$	Negative Affect est	se	$p <$
Subject level						
Intercept $\sigma^2_{\upsilon_{(3)}}$	.130	.031	.001	.084	.023	.001
Scale σ^2_{ω}	.780	.106	.001	1.28	.166	.001
Int, Scale $\sigma_{\upsilon_{(3)}\omega}$	.186	.040	.001	-.189	.041	.001
	$(r = .59)$			$(r = -.58)$		
Wave level						
Intercept $\sigma^2_{\upsilon_{(2)}}$	.090	.021	.001	.028	.012	.022

Mixed-effects Proportional Odds Model: ordinal response Y_{ij} of subject i $(i = 1, 2, \ldots, N)$ on occasion j $(j = 1, 2, \ldots, n_i)$

$$\lambda_{ijc} = \log \left[\frac{P_{ijc}}{1 - P_{ijc}} \right] = \gamma_c - \left[\boldsymbol{x}'_{ij} \boldsymbol{\beta} + \upsilon_i \right]$$

$P_{ijc} = \Pr(Y_{ij} \le c)$ cumulative probabilities for C categories of Y

$\boldsymbol{x}_{ij} = p \times 1$ vector of regressors (no 1 for the intercept)

$\boldsymbol{\beta} = p \times 1$ vector of regression coefficients

$\gamma_1 < \gamma_2 < \ldots < \gamma_{C-1}$ strictly increasing thresholds

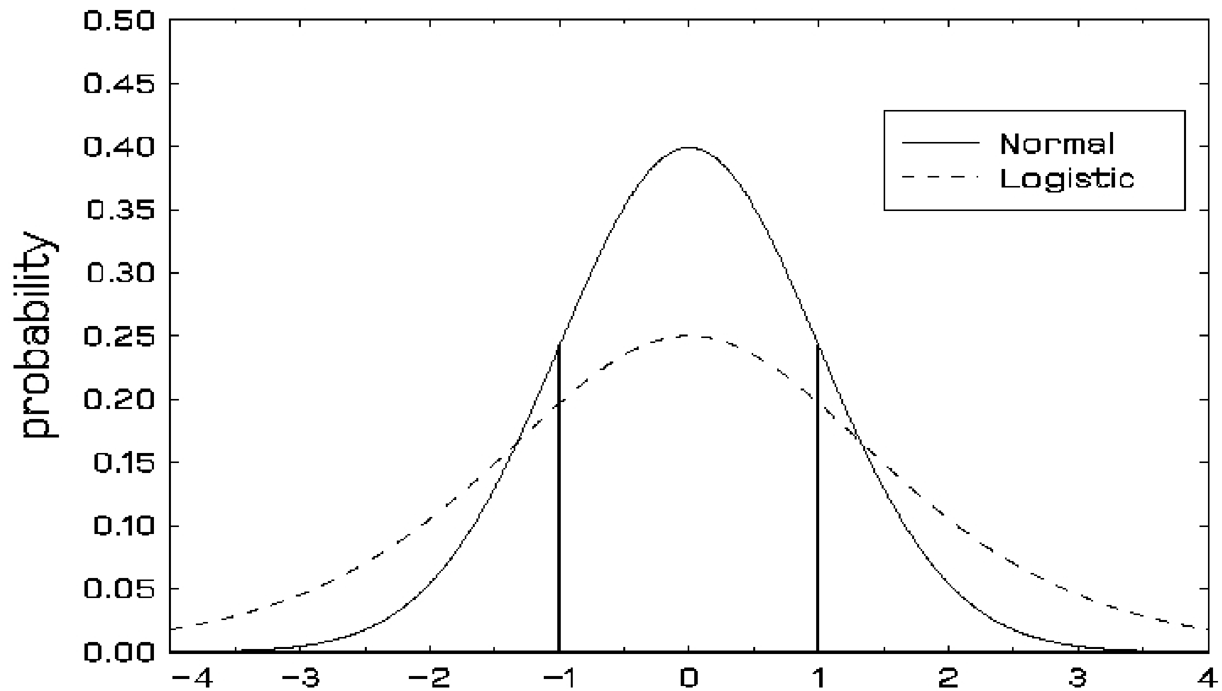
$\upsilon_i \sim N(0, \sigma_\upsilon^2)$ BS variance

Ordinal Response and Threshold Concept

Continuous y_{ij} - unobservable latent variable - related to ordinal response Y_{ij} via "threshold concept"

- threshold values $\gamma_1, \gamma_2, \ldots, \gamma_{C-1}$ ($\gamma_0 = -\infty$ and $\gamma_C = \infty$)
- C = number of ordered categories

Response occurs in category c, $\quad Y_i = c \quad$ if $\quad \gamma_{c-1} < y_{ij} < \gamma_c$



The Threshold Concept in Practice

"How was your day?"
(what is your level of satisfaction today?)

- Satisfaction may be continuous, but we sometimes emit an ordinal response:

Model for Latent Continuous Responses

Model with p covariates for the latent response strength y_{ij}:

$$y_{ij} = \boldsymbol{x}'_{ij}\boldsymbol{\beta} + \upsilon_i + \epsilon_{ij}$$

where $\upsilon_i \sim N(0, \sigma^2_\upsilon)$, BS variance, and WS errors

- $\epsilon_{ij} \sim$ standard normal (mean 0 and $\sigma^2_\epsilon = 1$)
  mixed-effects ordinal probit regression
- $\epsilon_{ij} \sim$ standard logistic (mean 0 and $\sigma^2_\epsilon = \pi^2/3$)
  mixed-effects ordinal logistic regression

Mixed-effects Ordinal Location Scale Model

$$\lambda_{ijc} = \frac{\gamma_c - (\boldsymbol{x}'_{ij}\boldsymbol{\beta} + \upsilon_i)}{\sigma_{\epsilon_{ij}}}$$

BS variance $\quad \sigma^2_{\upsilon_{ij}} = \exp(\boldsymbol{u}'_{ij}\boldsymbol{\alpha}) \quad$ or $\quad \log(\sigma^2_{\upsilon_{ij}}) = \boldsymbol{u}'_{ij}\boldsymbol{\alpha}$

WS variance $\quad \sigma^2_{\epsilon_{ij}} = \exp(\boldsymbol{w}'_{ij}\boldsymbol{\tau} + \omega_i) \quad$ or $\quad \log(\sigma^2_{\epsilon_{ij}}) = \boldsymbol{w}'_{ij}\boldsymbol{\tau} + \omega_i$

- $\boldsymbol{u}_{ij}$ and $\boldsymbol{w}_{ij}$ include covariates (and $\mathbf{1}$ only for $\boldsymbol{u}_i$)
- random location effects $\upsilon_i \sim N(0, \sigma^2_\upsilon)$
- random scale effects $\omega_i \sim N(0, \sigma^2_\omega)$

Ecological Momentary Assessment (EMA) Study of Adolescent Smokers (Mermelstein)

- 461 adolescents (9th and 10th graders); former and current smoking experimenters, and regular smokers

- Carry PDA for a week, answer questions when prompted

  average = 30 answered prompts (range = 7 to 71)

- $\Sigma_i^N n_i = 14,105$ total number of observations

Outcome: "I Felt Sad"

Interest: characterizing determinants of affect level, as well as BS and WS affect heterogeneity

I Felt Sad: marginal response frequencies and percentages

`Sad`	Frequency	Percent
1	6087	43.15
2	2269	16.09
3	1716	12.17
4	813	5.76
5	439	3.11
6	671	4.76
7	773	5.48
8	579	4.10
9	292	2.07
10	466	3.30

$\Rightarrow$ items rated on 1 (not al all) to 10 (very much) scale

	mean	std dev	min	max
Subject-level independent variables				
`Male`	.449	.498	0	1
`Smoker`	.508	.500	0	1
`Psmk` (234 smokers)	.131	.117	.014	.583
`AloneBS`	.517	.196	.024	.950
Prompt-level independent variables				
`AloneWS`	0	.461	-.950	.976

`Smoker`: gave at least one report of a smoking event in the week of EMA measurement (about half of the subjects)

`Psmk`: proportion of occasions (random prompts and smoking events) that were smoking events = n_smk / (n_smk + n_random)

For occasion-varying `Alone`, BS and WS decomposition:

$$X_{ij} = \bar{X}_i + (X_{ij} - \bar{X}_i)$$

Proportional odds mixed model

estimates, standard errors (se), and p-values

parameter	estimate	se	$p <$
`Male` β_1	-.716	.161	.0001
`Smoker` β_2	.477	.198	.017
`PSmk` β_3	-1.253	.942	.19
`AloneBS` β_4	1.082	.410	.009
`AloneWS` β_5	.527	.036	.0001
BS variance α_0	.965	.074	.0001

In terms of the BS variance, $\hat{\sigma}_\upsilon^2 = \exp(.965) = 2.625$

Intraclass correlation (ICC)
$\text{ICC} = 2.625/(2.625 + \pi^2/3) \ = \ .44$

Mixed location scale model

parameter	No random scale est	se	$p <$		With random scale est	se	$p <$
Location							
`Male` β_1	-.501	.136	.001		-.498	.157	.002
`Smoker` β_2	.358	.167	.04		.370	.183	.05
`PSmk` β_3	-.912	.775	.24		-.850	.833	.31
`AloneBS` β_4	.938	.338	.006		.876	.356	.02
`AloneWS` β_5	.460	.034	.001		.359	.039	.001
WS variance							
`Male` τ_1	-.318	.043	.001		-.401	.110	.001
`Smoker` τ_2	.325	.052	.001		.371	.135	.006
`Psmk` τ_3	-.909	.282	.002		-1.116	.659	.09
`AloneBS` τ_4	-.562	.108	.001		-.422	.281	.14
`AloneWS` τ_5	-.117	.044	.008		-.109	.044	.02
BS variance							
Intercept α_0	.772	.221	.001		.936	.251	.001
`Male` α_1	-.586	.155	.001		-.717	.144	.001
`Smoker` α_2	.079	.186	.67		.130	.173	.46
`Psmk` α_3	-.196	.867	.83		-.499	.759	.52
`AloneBS` α_4	.165	.379	.67		.288	.348	.41
Scale							
variance σ_ω^2					1.001	.084	.001
covariance $\sigma_{\upsilon\,\omega}$					-.506	.099	.001

Smoking effect on WS variance

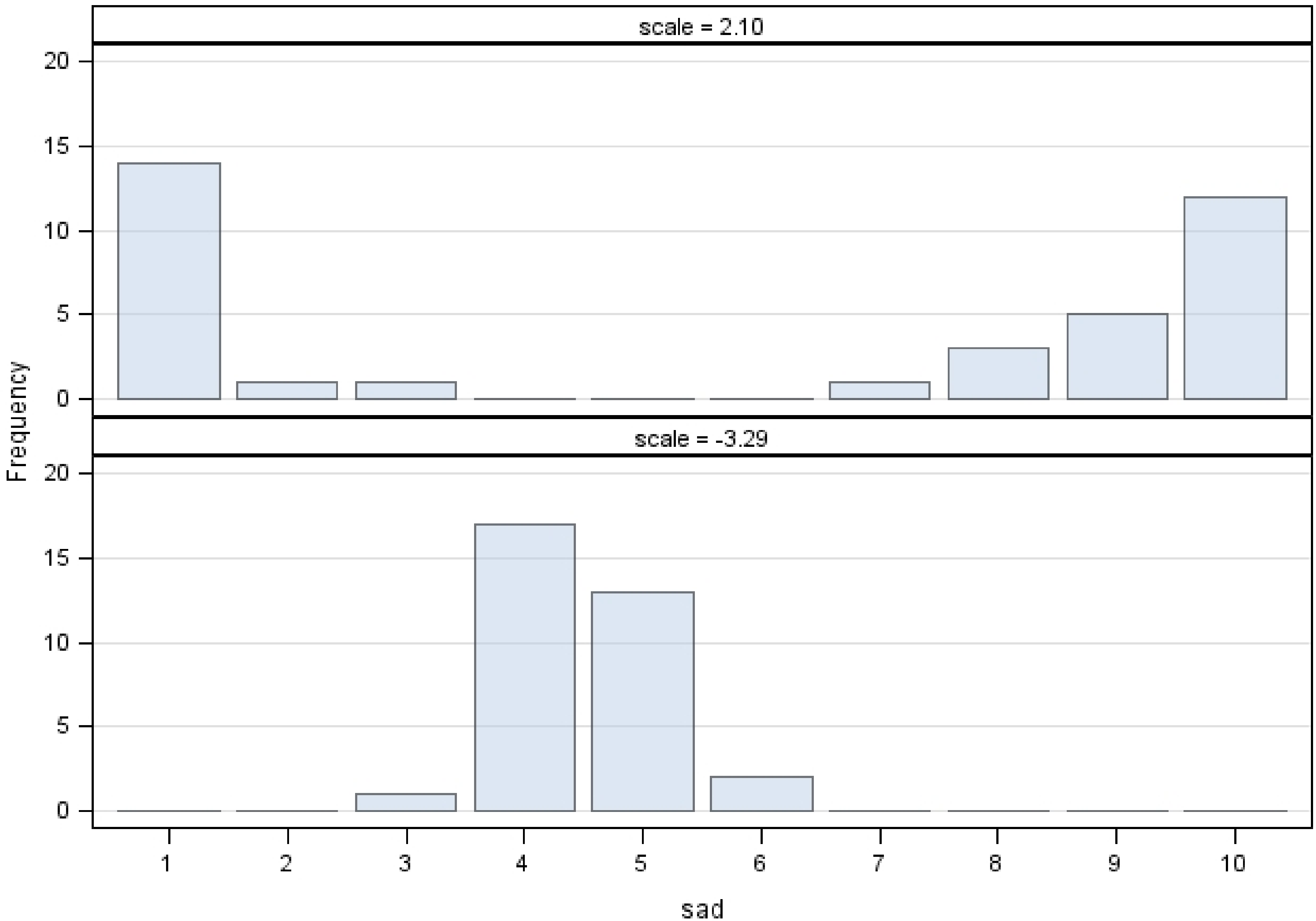
`Smoker` = .371 positive effect (increased mood variation)
`Psmk` = -1.116 negative effect (decreased mood variation)

`Psmk` value with zero effect on mood variation = .371/1.116 = .332

Of 234 smokers:
`Psmk` median = .081
`Psmk` 90% percentile = .3
`Psmk` 95% percentile = .367

$\Rightarrow$ most smokers elicited more varied response than non-smokers

Subjects with largest and smallest scale estimates



Summary

- More applications where interest is on modeling variance

Hedeker, D., Mermelstein, R.J., & Demirtas, H. (2008). An application of a mixed-effects location scale model for analysis of Ecological Momentary Assessment (EMA) data. *Biometrics, 64*, 627-634.

Hedeker, D., Mermelstein, R.J., & Demirtas, H. (2012). Modeling between- and within-subject variance in EMA data using mixed-effects location scale models. *Statistics in Medicine, 31*, 3328-3336.

Hedeker, D. & Mermelstein, R.J. (2012). Mood changes associated with smoking in adolescents: An application of a mixed-effects location scale model for longitudinal EMA data. In G. R. Hancock & J. Harring (Eds.), *Advances in Longitudinal Methods in the Social and Behavioral Sciences* (pp. 59-79). Information Age Publishing, Charlotte, NC.

Hedeker, D. & Nordgren, R. (2013). MIXREGLS: A program for mixed-effects location scale analysis. *Journal of Statistical Software, 52(12)*, 1-38.

Kapur, K., Li, X., Blood, E.A., & Hedeker, D. (2015). Bayesian mixed-effects location and scale models for multivariate longitudinal outcomes: An application to ecological momentary assessment data. *Statistics in Medicine, 34*, 630-651.

Li, X. & Hedeker, D. (2012). A three-level mixed-effects location scale model with an application to Ecological Momentary Assessment (EMA) data. *Statistics in Medicine, 31*, 3192-3210.

Pugach, O., Hedeker, D., Richmond, M.J., Sokolovsky, A., & Mermelstein, R.J. (2014). Modeling mood variation and covariation among adolescent smokers: Application of a bivariate location-scale mixed-effects model. *Nicotine and Tobacco Research, 16, Supplement 2*, S151-S158.

- Ordinal outcomes

Hedeker, Demirtas, & Mermelstein (2009). A mixed ordinal location scale model for analysis of Ecological Momentary Assessment (EMA) data. *Statistics and Its Interface, 2*, 391-402.

Hedeker, D., Mermelstein, R.J., Demirtas, H., & Berbaum, M.L. (under review). A mixed-effects location-scale model for ordinal questionnaire data.

More Examples of Variance Models in Health Studies

- Lin, Raz, & Harlow (1997) Linear mixed models with heterogeneous within-cluster variances, Biometrics. *Determinants of menstrual cycle length variability in women (which may be associated with fertility and long-term risk of chronic disease).*

- Carroll (2003) Variances are not always nuisance parameters, Biometrics. *Drug assay validation, measurement error in nutrient intake.*

- Elliott (2007) Identifying latent clusters of variability in longitudinal data, Biostatistics. *Clusters based on within-subject variation in affect of recovering MI patients.*

- Elliott, Sammel, & Faul (2010) Associations between variability of risk factors and health outcomes in longitudinal studies, Statistics in Medicine. *Residual variability in longitudinal recall data associated with dementia risk in elderly.*

- Rast & Zimprich (2011) Modeling within-person variance in reaction time data of older adults, Journal of Gerontopsychology and Geriatric Psychiatry.

- Coffman, Allen, & Woolson (2012) Mixed-effects regression modeling of real-time momentary pain assessments in osteoarthritis (OA) patients, Health Services and Outcomes Research Methodology. *Pain variability in patients with osteoarthritis.*

- Breslin (2014) Five indices of emotion regulation in participants with a history of nonsuicidal self-injury: A daily diary study, Behavior Therapy.

- Need a fair amount of BS and WS data, but modern data collection procedures are good for this. Also, from analysis of Riesby depression data ($N = 66, n_i = 4$ to 6):

The data of the two highest and lowest scale estimates from analysis of the Riesby data

id	$\tilde{\theta}_{2i}$	hd0	hd1	hd2	hd3	hd4	hd5
606	1.585	19	33	12	12	3	1
505	1.532	21	11	18	0	0	4
335	−1.317	21	21	18	15	12	10
308	−1.365	22	21	18	17	12	11

- Simulations with small datasets (*e.g.*, 20 subjects with 5 obs) often leads to non-convergence; this improves dramatically as numbers increase (*e.g.*, 100 subjects with 10 obs)

- Important to include random scale for correct inference of WS variance covariates (Leckie et al., 2014, *Jrn Educ Beh Stat)*