Longitudinal Data Analysis via Linear Mixed Models

Edps/Psych/Soc 587

Carolyn J. Anderson

Department of Educational Psychology
ILLINOIS

Outline

- Introduction
- Approaches to Longitudinal Data Analysis
- Longitudinal HLM by Example
    - The Riesby Data
    - Exploratory Analysis
    - Model Selection
- Models for Serial Correlation

Reading and References

Reading: Snijders & Bosker, chapter 12

Additional References:

- Diggle, P.J., Liang, K.L., & Zeger, S.L. (2002). *Analysis of Longitudinal Data*, 2nd Edition. London: Oxford Science.
- Notes by Donald Hedeker. Available from his web-site http://tigger.uic.edu/˜hedeker.
- Hedeker, D., & Gibbons, R.D. (2006). *Longitudinal Data Analysis*. Wiley.
- Singer, J.D. & Willett, J.B. (2003). *Applied Longitudinal Data Analysis*. Oxford.
- Verbeke, G, & Molenberghs, G. (2000). *Linear Mixed Models for Longitudinal Data*. Springer.

Introduction

- Purpose: Study change and the factors that effect change.
- Data: Longitudinal data consist of repeated measurements on the same unit over time.
- Models: Hierarchical Linear Models (linear mixed models) with extensions for possible serial correlation and non-linear pattern of change.

Purpose: Study Nature of Change

Goal: Study change and the factors that effect intra- and inter-individual change.

- Differences found in cross-sectional data often explained as reflecting change in individuals.
- A model for cross-sectional data

$$Y_{i1} = \beta_0 + \beta_{cs} x_{i1} + \epsilon_{i1}$$

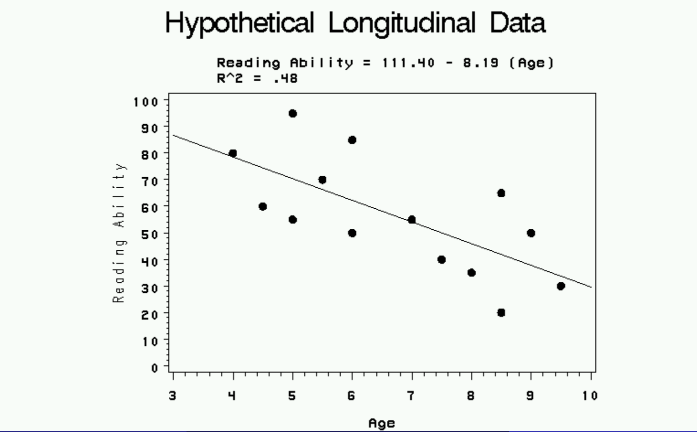
where $i = 1, \ldots, N$ (individuals) and x_{i1} is some time measure (e.g., age).

- Interpretation: $\beta_{cs} =$ difference in Y between 2 individuals that differ by 1 unit of time (x).

Cross-Sectional Data

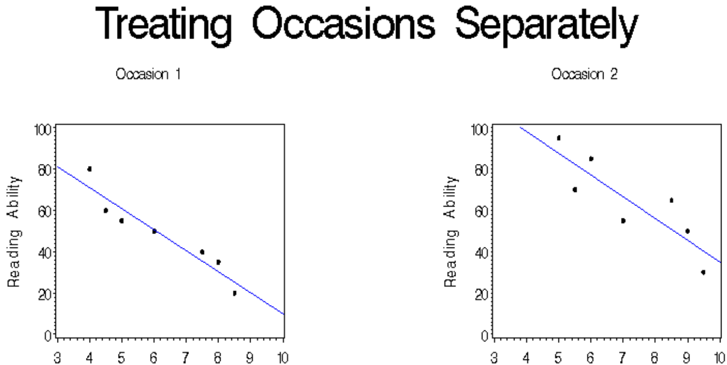
Ignoring longitudinal structure:

$$\widehat{(\text{reading})}_i = 111.40 - 8.19(\text{age})_i$$



Cross-Sectional Data (continued)

$$\text{Occasion 1:} \quad \widehat{(\text{reading})}_{i1} = 111.86 - 10.18(\text{age})_{i1}$$
$$\text{Occasion 2:} \quad \widehat{(\text{reading})}_{i2} = 140.01 - 10.50(\text{age})_{i2}$$



A Model for Longitudinal Data

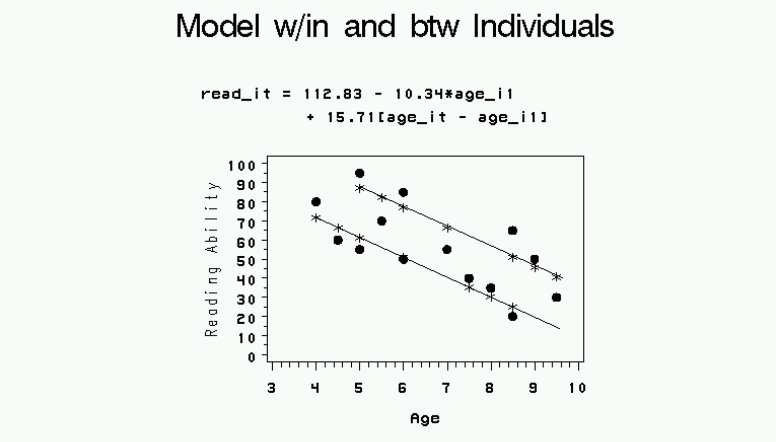
or repeated observations.

$$Y_{it} = \beta_0 + \beta_{cs} x_{i1} + \beta_l (x_{it} - x_{i1}) + \epsilon_{it}$$

- When $t = 1$, the model is the same as the cross-sectional model.
- $\beta_l =$ the expected change in Y over time per unit change in the time measure x (within individual differences).
- β_{cs} still reflects differences between individuals.
- β_{cs} and β_l reflect different processes.

A Model for Longitudinal Data

$$\widehat{(\text{reading})}_{it} = 112.83 - 10.34(\text{age})_{i1} + 15.71[(\text{age})_{it} - (\text{age})_{i1}]$$



Advantages: Longitudinal Data

More Powerful.

- Inference regarding β_{cs} is a comparison of individuals with the same value of x.
- Inference regarding β_l is a comparison of an individual's response at two times

  $\Longrightarrow$ Assuming y changes systematically with time and retains it's meaning.
- Each individual is their own control group.
- Often there is much more of variability between individuals than within individuals and the between variability is consistent over time.

Advantages: Longitudinal Data (continued)

Distinguish Among Sources of Variation.

Variation in Y may be due

- Between individuals differences.
- Within individuals:
  - Measurement error & unobserved covariates.
  - Serial correlation.
- A step toward showing causality.
  - Causal relativity (i.e., effect of cause relative to another).
  - Causal manipulation.
  - "Cause" precedes effect (i.e. temporal ordering).
  - Rule out all other possibilities.

See Schneider, Carnoy, Kilpatrick & Shavelson (2010). *Estimating Causal Effects Using Experimental and Observational Designs:A think Thank White Paper*. The Governing Board of the AERA Grant Program.

Studying Change

Longitudinal data is required to study the pattern of change and the factors that effect it, both within and between individuals.

- Level 1: How does the outcome change over time? (descriptive)
- Level 2: Can we predict differences between individuals in terms of how they change? (relational).

Time

- Time is a level 1 (micro level) predictor.

  The number of time points/occasions needed.

- Measure of time should be
  - Reliable
  - Valid
  - Makes sense for outcome and research questions.
  - The Meaning does not change over time.

Metric for Time

Example from Singer & Willett:

If you want to study the "longevity" of automobiles.

- Change in appearance of cars $\longrightarrow$ Age.
- Tire wear $\longrightarrow$ Miles.
- Wear of ignition system $\longrightarrow$ Trips (# of starts).
- Engine wear $\longrightarrow$ Oil changes.

Metric & Clock for Time

Example from Nicole Allen et al.

Study the change in arrest rates following passage of law in 1994 requiring coordinated responses to cases of domestic violence.

- Daily data from all municipalities in Illinois (excluding those in Cook) from 1996 to 2004.
- Zero point?
  - 1996?
  - When council (coordinated response) began?
  - Others
- Metric? (Daily, Weekly, Monthly, Quarterly, Yearly?)
- Level? (Municipality? County? Circuit?)

Three Major Approaches

to analyze longitudinal data.

Classic reference: Diggle, Liang & Zeger

- Marginal Analysis: Only interested in average response.
- Transition Models: Focus on how Y_{it} depends on past values of Y and other variables (i.e., conditional models, Markov models).
- Random Effects Models: Focus on how regression coefficients vary over individuals.

Marginal Analysis

Focus on average of the response variable:

$$\bar{Y}_{+t} = \frac{1}{N}\sum_{i=1}^{N} Y_{it}$$

and how the mean changes over time.

- In HLM terms, only interested in the fixed effects,

$$E(Y_{it}) = \boldsymbol{X}_i\boldsymbol{\Gamma}.$$

- Observations are correlated, so need to make adjustments to variance estimates, i.e., $\text{var}(\boldsymbol{Y}_i) = \boldsymbol{V}_i(\theta)$ where θ are parameters.
- "Sandwich estimator" or Robust estimation (of standard errors of parameters).

Transition Models

Focus on how Y_{it} depends on pervious values of Y (i.e., $Y_{i,(t-1)}$, $Y_{i,(t-2)}, \ldots$) and other variables.

- Model the Conditional Distribution of Y_{it},

$$E(Y_{it}|Y_{i,(t-1)}, \ldots, Y_{i,1}, \boldsymbol{x}) = \sum_{k=1}^{p} \beta_k x_{itk} + \sum_{k=1}^{(t-1)} \alpha_k Y_{i,k}$$

- Such models include assumptions about
  - Dependence of Y_{it} on x_{it}'s.
  - Correlation between repeated measures.

Transition Models (continued)

We have focused on continuous/numerical Y's, but when Y is categorical,

- "Stage sequential models" (e.g., must master addition and subtraction before can master multiplication).
- The "gateway hypothesis" of drug use.

Digression: When an event occurs is another type discrete outcome variable, but we're not considering such discrete variables in this class.

Random Effects Models

- Observations are correlated because repeated measurements are made on the same individual.
- Regression coefficients vary over individuals, i.e.,

$$\mathsf{E}(Y_{it}|\beta_{i1},\ldots,\beta_{ip}) = \sum_{k=1}^{p} \beta_{ik} x_{ikt}$$

- One individual's data does not contain enough information to estimate β_{ik}'s ; therefore, we assume a distribution for β_{ik}'s,

$$\boldsymbol{\beta}_i = \boldsymbol{X}_i \boldsymbol{\Gamma} + \boldsymbol{Z}_i \boldsymbol{U}_i$$

where $\boldsymbol{U}_i \sim \mathcal{N}(\mathbf{0}, \boldsymbol{T})$ *i.i.d.*.

Advantages of HLM

for Longitudinal Data

- Explicitly model individual change over time.
- Simultaneously and explicitly model between- and within-individual variation.
- Explanatory variables can be time-invariant or time-varying.
- Flexible modeling of covariance structure of the repeated measures.
- Many non-linear patterns can be represented by linear models (e.g., polynomial, spline).

Advantages of HLM

- Flexible treatment of time
  - Time can be treated as a continuous variable or as a set of fixed points.
  - Can have a different numbers of repeated observations. (implication: can handle missing data).
- Can extend HLM models to higher level structures (e.g., repeated measurements on students within classes, etc).
- Generalizations exist for non-linear data.

HLM for Longitudinal Data

Uses everything we've learned about HLM's, but requires a slight change in terminology and notation:

- Level 1 units are occasions of measurement and indexed by t (t for "time" where $t = 1, \ldots, r$).
- Level 2 units are individuals.
- $Y_{it} =$ measurement of response/dependent variable for individual i at time t.
- The level 1 model: within individual model.
- The level 2 model: between individuals model.

HLM for Longitudinal Data

One major change: May need a more complex model for the level 1 (within individual) residuals; that is,

$$\boldsymbol{R}_i \sim \mathcal{N}(0, \boldsymbol{\Sigma}_i)$$

where $\boldsymbol{\Sigma}_i = \sigma^2 \boldsymbol{I}$ (constant and uncorrelated) may be too simple.

One's that we'll explicitly cover are lag 1:

- Auto-correlated errors, AR(1).
- Moving average, MA(1).
- Auto-correlated, moving average ARMA(1,1).
- TOEP(#).

Longitudinal HLM by Example

The Riesby Data, from Hedeker's web-site (and used in Hedeker & Gibbons book, 2006).

- Drug Plasma Levels and Clinical Response.
- "Risby and associates (Riesby, *et al*, 1977) examined the relationship between Imipramine (IMI) and Desipramine (DMI) plasma levels and clinical response in 66 depressed inpatients (37 endogenous and 29 non-endogenous)."

The Riesby Data

Outcome variable: Hamilton Depression Score (HD).

Independent variables:

- Gender.
- D where $= 1$ for endogenous and $= 0$ of non-endogenous.
- IMI (imipramine) drug-plasma levels ($\mu g/1$). — Antidepressant given 225 mg/day, weeks 3-6.
- DMI (desipramine) drug-plasma levels ($\mu g/1$). — Metabolite of imipramine.

The Design

		Drug-Washout					
		day 0 wk 0	day 7 wk 1	day 14 wk 2	day 21 wk 3	day 28 wk 4	day 35 wk 5
Y_{it}	Hamilton Depression	HD_1	HD_2	HD_3	HD_4	HD_5	HD_6
Level 2	Gender	G					
	Diagnosis	D					
Level 1	IMI	—	—	IMI_3	IMI_4	IMI_5	IMI_6
	DMI	—	—	DMI_3	DMI_4	DMI_5	DMI_6
	N	61	63	65	65	63	58

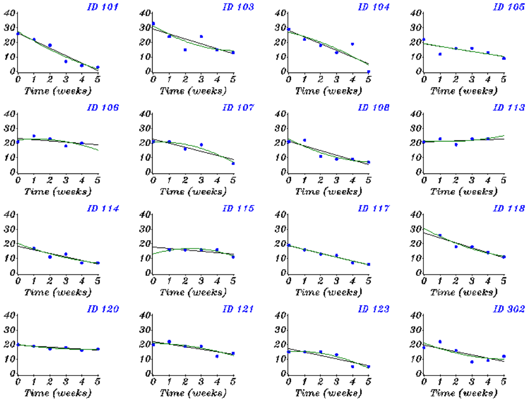
Note: $n = 6$ and $N = 66$.

More Information About the Topic

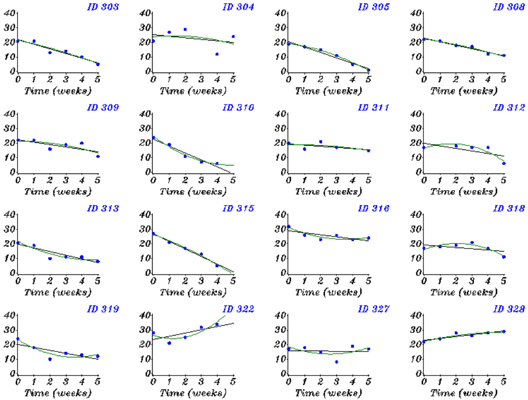
From a Psychiatrist friend:

- "Everyone uses Hamilton Depression Score"
- Good that both IMI and DMI are used. In the psychiatric literature, the sum is usually reported.
- Imipramine is an older drug, which has many undesirable side effects, but it works.
- Distinction between diagnosis with respect to drug not done (relevant to practice).

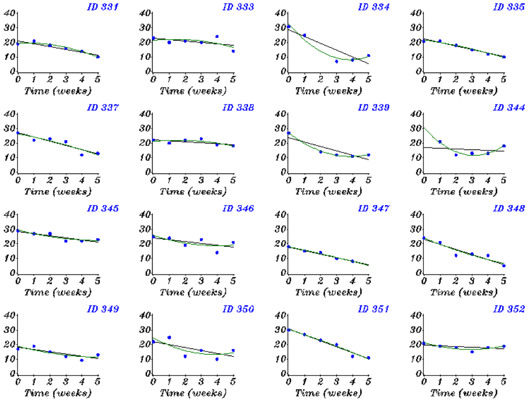
Exploring Individual Structures



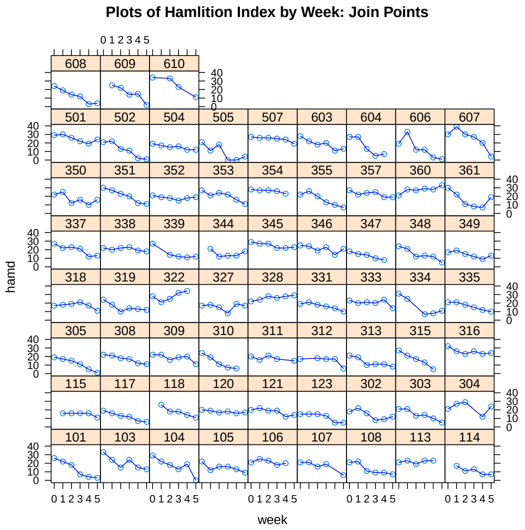
Exploring Individual Structures



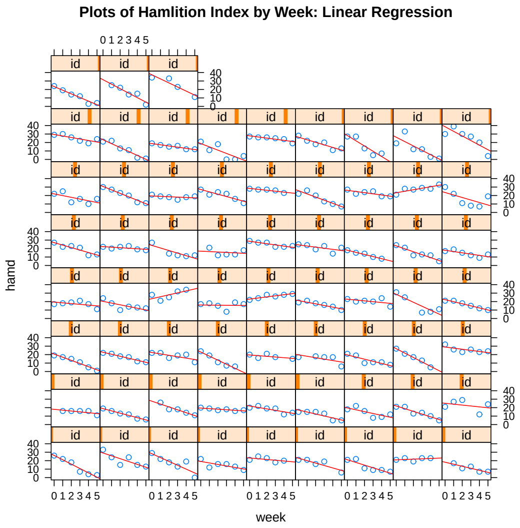
Exploring Individual Structures



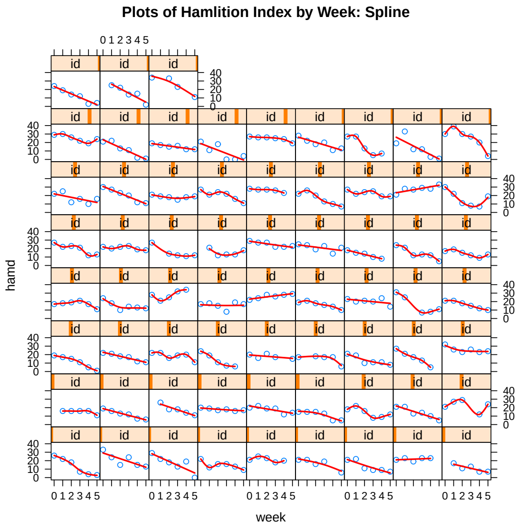
Exploring Individual Structures: R graphs



Exploring Individual Structures: R graphs

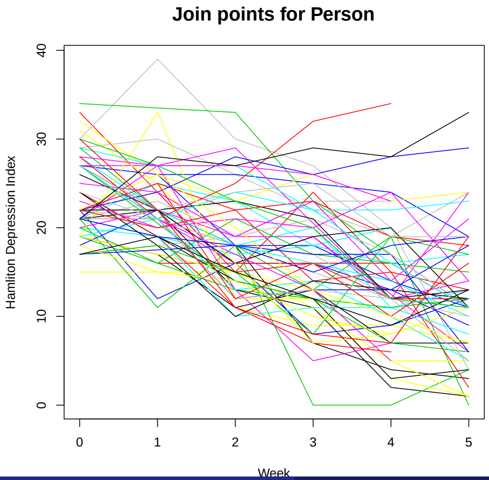


Exploring Individual Structures: R graphs

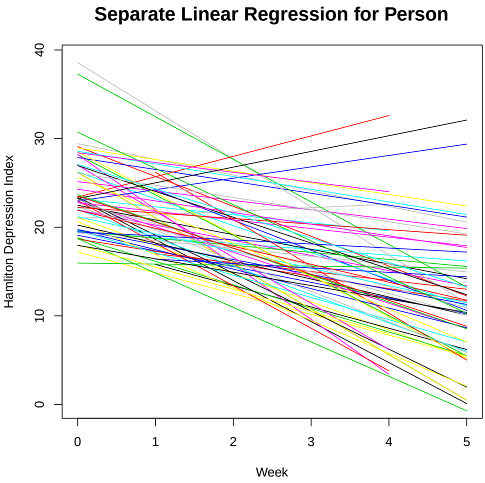


Join Points

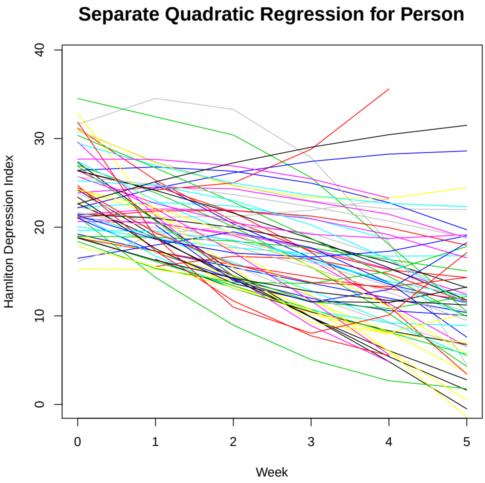
General linear decline and increasing variance.



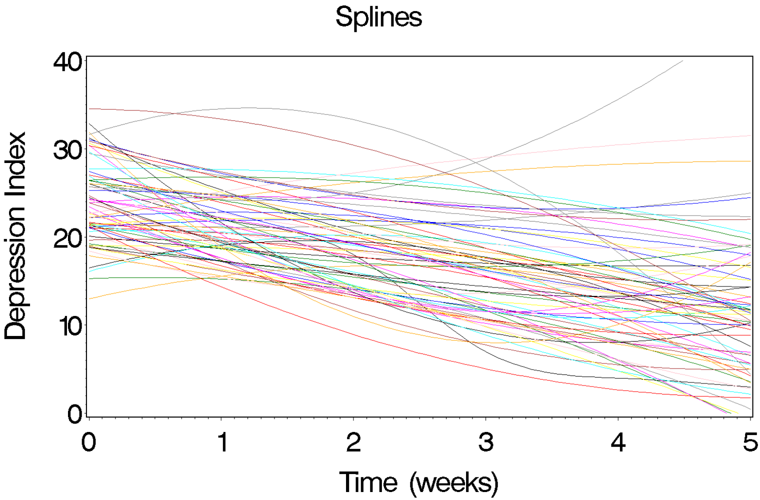
Overlay Individual Regressions



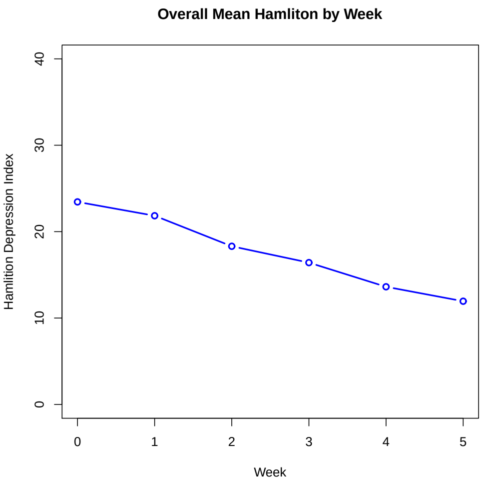
Overlay Quadratic Regressions



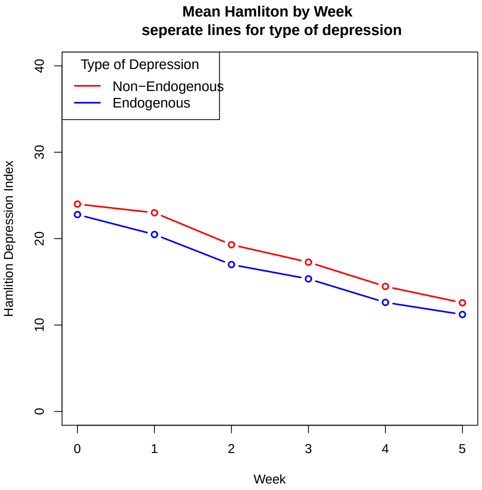
Overlay Individual Regressions



Exploring Mean Structure



Exploring Mean Structure (continued)



Exploring Individual Specific Models

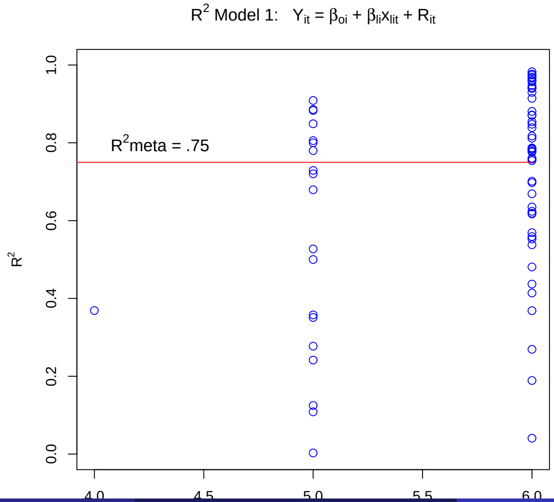
Based on the figures, a plausible for level 1

$$Y_{it} = \beta_{0i} + \beta_{i1}(\text{week})_{it} + \epsilon_{it}$$

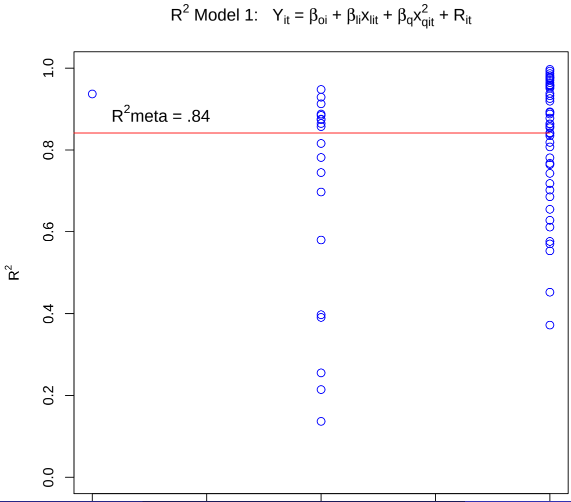
Using OLS, fit this model to each person's data and compute:

- $R^2_i = (\text{ssmodel})_i/(\text{sstotal})_i$.
- $R^2_{meta} = \sum_i(\text{ssmodel})_i/\sum_i(\text{sstotal})_i$.
- Try to see who improves.

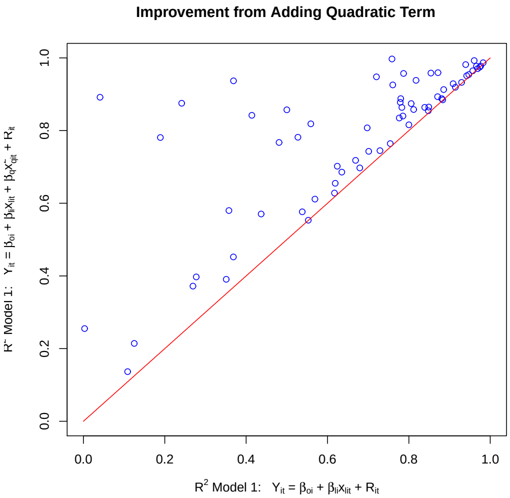
Linear Model: R_i^2 and R_{meta}^2



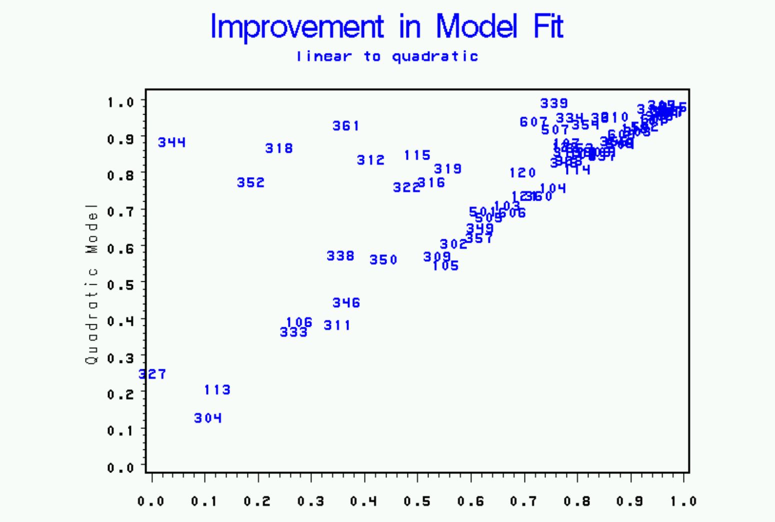
Quadratic: R_i^2 and R^2_{meta}



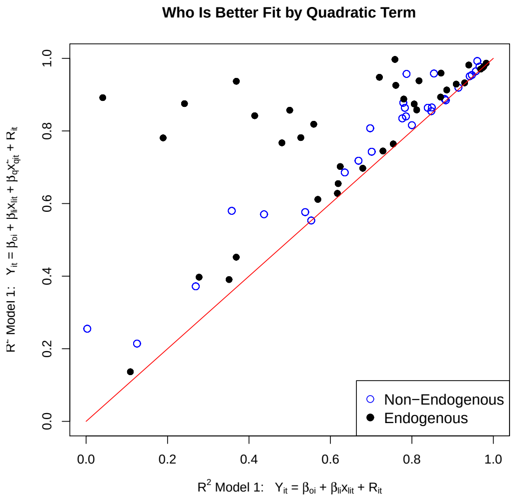
Comparison


Improvement from Adding Quadratic Term
R^2 Model 1: $Y_{it} = \beta_{oi} + \beta_{li}x_{lit} + \beta_q x^2_{qit} + R_{it}$
R^2 Model 1: $Y_{it} = \beta_{oi} + \beta_{li}x_{lit} + R_{it}$
0.0
0.2
0.4
0.6
0.8
1.0

Who Improved?



Who Improved?



F "Test" for Quadratic Term

- Reduced Model: $Y_{it} = \beta_o + \beta_1(\text{week})_{it} + \epsilon_{it}$
  $p = 2.$
- Full Model: $Y_{it} = \beta_o + \beta_1(\text{week})_{it} + \beta_2(\text{week})^2_{it} + \epsilon_{it}$
  $p^* = 1.$
- F-statistic:

$$F = \frac{(\sum_i (\text{sserror})(R)_i - (\text{sserror})(F)_i)/\sum_i p_i}{\sum_i (\text{sserror})(F)_i/\sum_i (n_i - p - p^*)} = \frac{1075.28/66}{1858.02/177} = 1.55$$

Comparing $F = 1.55$ to the $\mathcal{F}$–distribution with $df_{num} = 66$ and $df_{den} = 177$, the "p-value"$= .01$.
To be used with a Large "grain of sand" (assumptions violated)

Preliminary HLM

- Level 1: $Y_{it} = \beta_{oi} + \beta_{1i}(\text{week})_{it} + \beta_{2i}(\text{week})^2_{it} + R_{it}$
- Level 2:

$$\begin{aligned}\beta_{oi} &= \gamma_{00} + \gamma_{01}(\text{endog})_i \\ \beta_{1i} &= \gamma_{10} \\ \beta_{2i} &= \gamma_{20}\end{aligned}$$

- Preliminary Mixed Linear Model:

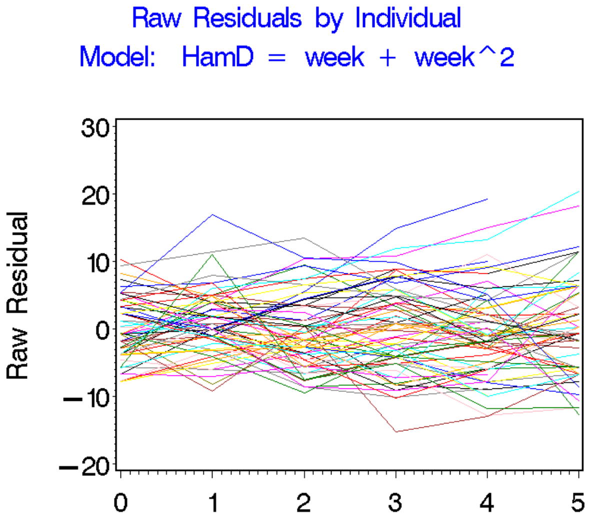
$$Y_{it} = \gamma_{00} + \gamma_{01}(\text{endog})_i + \gamma_{10}(\text{week})_{it} + \gamma_{20}(\text{week})^2_{it} + R_{it}$$

Exploring Random Effects

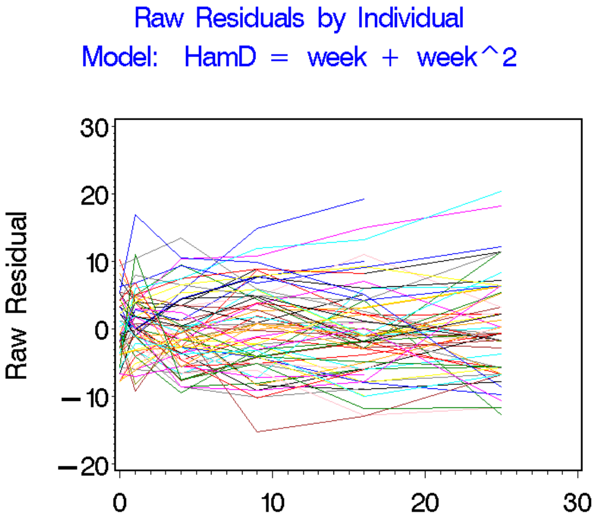
Fitting this model to each individual's data using ordinary least squares regression we look at

- Raw residuals,
$$\hat{R}_{it} = (Y_{it} - \hat{Y}_{it}) = \boldsymbol{Z}_i\boldsymbol{U}_i + \boldsymbol{R}_i.$$
- Squared residuals, $\hat{R}^2_{it}$.
- Correlations between residuals to look for serial correlation (i.e., need model for $\boldsymbol{\Sigma}_i$?)
$$\text{corr}(R_{it}, R_{it'}).$$

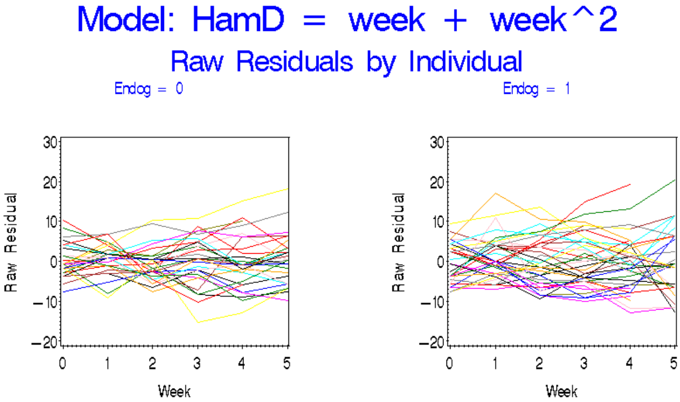
Raw Residuals by Week



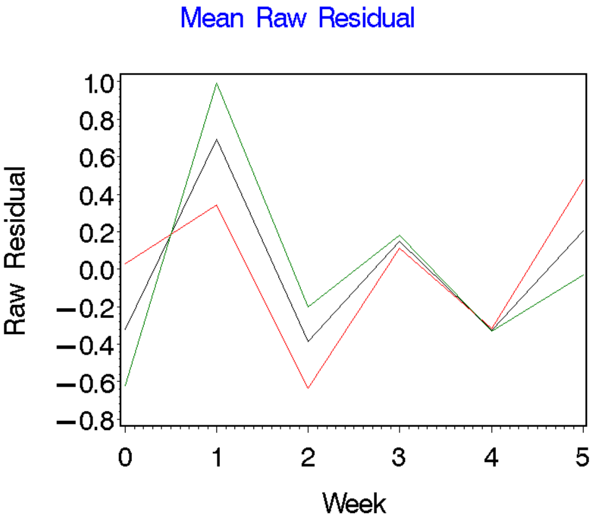
Raw Residuals by Week^2



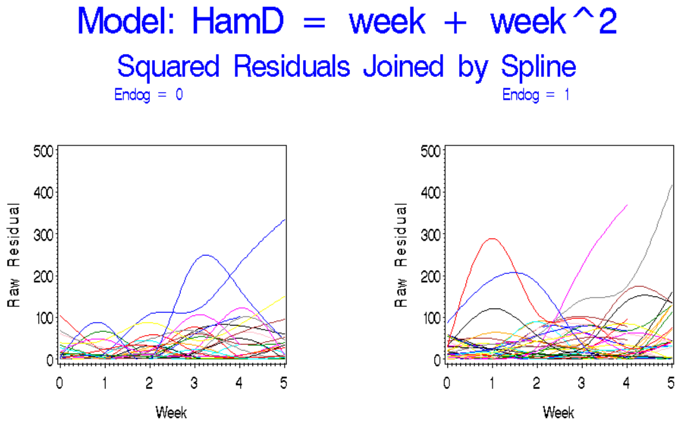
Raw Residuals with Endog



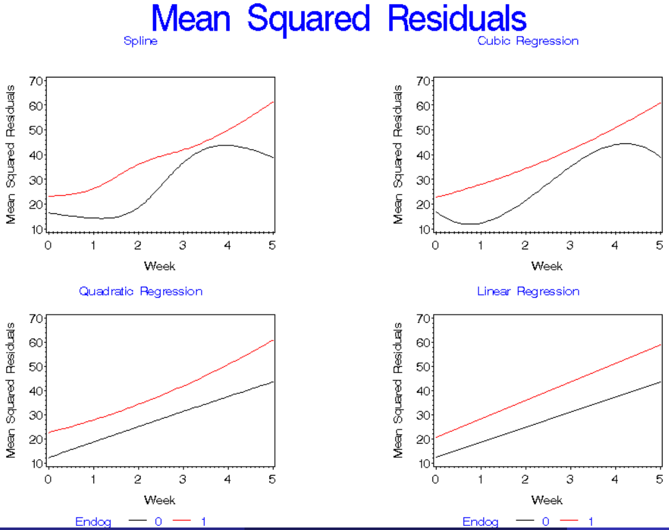
Mean Raw Residuals



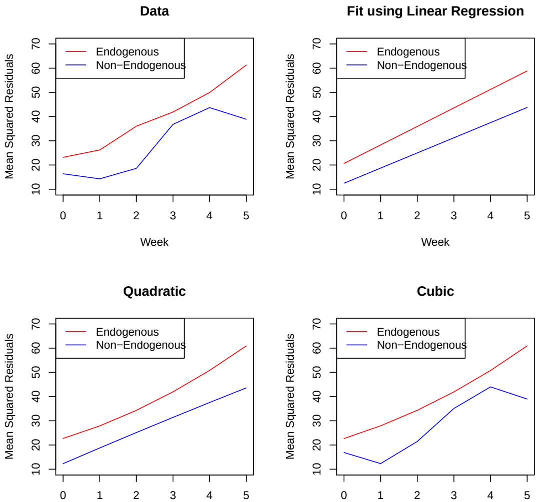
Squared Raw Residuals



Mean Squared Residuals



Mean Squared Residuals (R graph)



Variance Function

- Given Preliminary Mixed Linear Model:

$$Y_{it} = \gamma_{00} + \gamma_{01}(\text{endog})_i + \gamma_{10}(\text{week})_{it} + \gamma_{20}(\text{week})^2_{it} + R_{it}$$

- Assuming $R_{it} \sim \mathcal{N}(0, \sigma^2 \boldsymbol{I})$ and random intercept and slopes, i.e.,

$$(U_{0i}, U_{1i}, U_{2i})' \sim \mathcal{N}(\mathbf{0}, \boldsymbol{T})$$

- The variance of Y_{it} is 4th order polynomial

$$\begin{aligned} \text{var}(Y_{it}) &= \tau_0^2 + \tau_1^2\text{week}^2_{it} + \tau_2^2\text{week}^4_{it} + 2\tau_{01}\text{week}_{it} \\ &\quad + 2\tau_{02}\text{week}^2_{it} + 2\tau_{12}\text{week}^3_{it} + \sigma^2 \end{aligned}$$

Variance Function: 2 Random Effects

- Assuming $R_{it} \sim \mathcal{N}(0, \sigma^2 \boldsymbol{I})$ and random intercept and slope for week, i.e.,

$$(U_{0i}, U_{1i})' \sim \mathcal{N}(\mathbf{0}, \boldsymbol{T})$$

- The variance of Y_{it}

$$\text{var}(Y_{it}) = \tau_0^2 + \tau_1^2 \text{week}_{it}^2 + 2\tau_{01}\text{week}_{it} + \sigma^2$$

- How many random effects?

Need to also consider possible serial correlation.

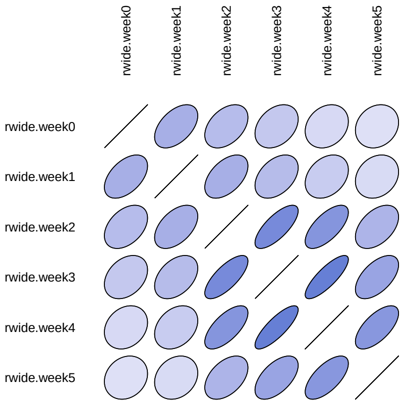
Correlation Between Time Points

Entries in the table are correlation$(\hat{R}_{it}, \hat{R}_{it'})$.

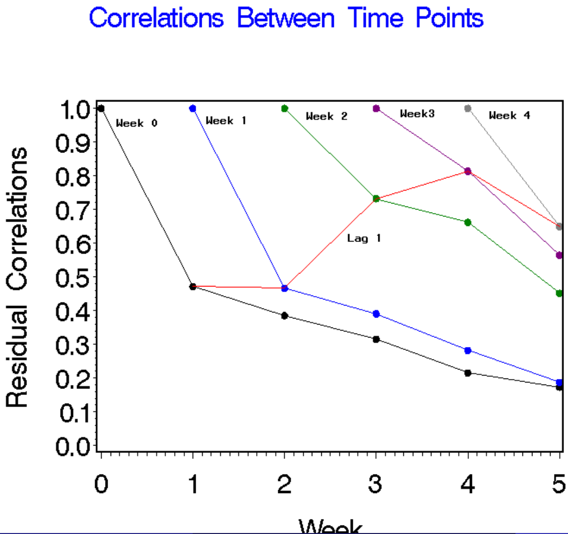
	week0	week1	week2	week3	week4	week5
week0	1.00					
week1	.47	1.00				
week2	.39	.47	1.00			
week3	.32	.39	.73	1.00		
week4	.22	.28	.66	.81	1.00	
week5	.17	.19	.45	.56	.65	1.00

Note: $46 \leq n \leq 66$ due to individuals with missing observations.

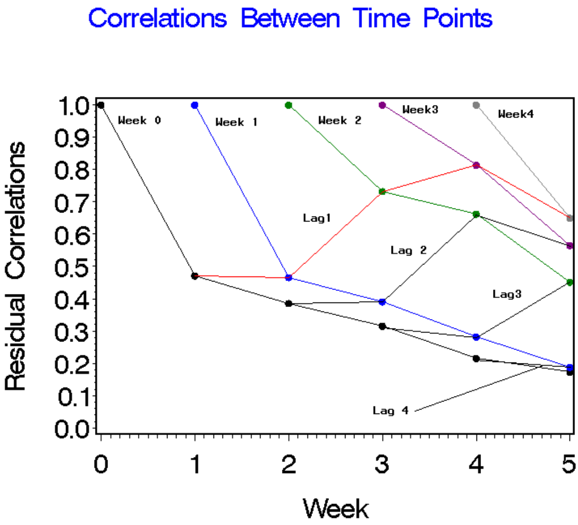
(R) Plot of Correlations



Plot of Correlations



Plot of Correlations: Lags



Mini-Outline (Next Steps)

- Before covering possible models for the level one, fit some HLM models to Riesby data (nothing new here).
- Consider some models for level 1 residuals.
- Simulation of data with different error structures.
- Analyze Riesby data using alternative error structures for level 1.

Model for Riesby Data

- Within Individual (level 1)

$$(\text{HamD})_{it} = \beta_{0i} + \beta_{1i}(\text{week})_{it} + \beta_{2i}(\text{week})^2_{it} + R_{it}$$

where $R_{it} \sim \mathcal{N}(0, \sigma^2)$.

- Between Individuals (level 2)

$$\begin{aligned} \beta_{0i} &= \gamma_{00} + \gamma_{01}(\text{endog})_i + U_{0i} \\ \beta_{1i} &= \gamma_{10} + \gamma_{11}(\text{endog})_i + U_{1i} \\ \beta_{2i} &= \gamma_{20} + U_{2i} \end{aligned}$$

where $\boldsymbol{U}_i \sim \mathcal{N}(\mathbf{0}, \boldsymbol{T})$.

Linear Mixed Model

- Scalar form

$$\begin{aligned}(\text{HamD})_{it} &= \gamma_{00} + \gamma_{10}(\text{week})_{it} + \gamma_{20}(\text{week})^2_{it} + \gamma_{01}(\text{endog})_i \\ &\quad + \gamma_{11}(\text{endog})_i(\text{week})_{it} + U_{0i} + U_{1i}(\text{week})_{it} \\ &\quad + U_{2i}(\text{week})^2_{it} + R_{it}\end{aligned}$$

- In matrix form,

$$\boldsymbol{Y}_i = \boldsymbol{X}_i\boldsymbol{\Gamma} + \boldsymbol{Z}_i\boldsymbol{U}_i + \boldsymbol{R}_i$$

Linear Mixed Model

$$\mathbf{Y}_i = \mathbf{X}_i\mathbf{\Gamma} + \mathbf{Z}_i\mathbf{U}_i + \mathbf{R}$$

$$\begin{pmatrix} (\text{HamD})_{i1} \\ (\text{HamD})_{i2} \\ \vdots \\ (\text{HamD})_{ir} \end{pmatrix} = \begin{pmatrix} 1 & (\text{week})_{i1} & (\text{week})^2_{i1} & D_i & D_i(\text{week})_{i1} \\ 1 & (\text{week})_{i2} & (\text{week})^2_{i2} & D_i & D_i(\text{week})_{i2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & (\text{week})_{ir} & (\text{week})^2_{ir} & D_i & D_i(\text{week})_{ir} \end{pmatrix} \begin{pmatrix} \gamma_{00} \\ \gamma_{10} \\ \gamma_{20} \\ \gamma_{01} \\ \gamma_{11} \end{pmatrix}$$

$$+ \begin{pmatrix} 1 & (\text{week})_{i1} & (\text{week})^2_{i1} \\ 1 & (\text{week})_{i2} & (\text{week})^2_{i2} \\ \vdots & \vdots & \vdots \\ 1 & (\text{week})_{ir} & (\text{week})^2_{ir} \end{pmatrix} \begin{pmatrix} U_{0i} \\ U_{1i} \\ U_{2i} \end{pmatrix} + \begin{pmatrix} R_{i1} \\ R_{i2} \\ \vdots \\ R_{ir} \end{pmatrix}$$

where $$D_i = \begin{cases} 0 & \text{Non-endogenous} \\ 1 & \text{Endogenous} \end{cases}$$

Marginal Model

$$(\text{HamD})_{it} \sim \mathcal{N}(\boldsymbol{X}_i\boldsymbol{\Gamma}, (\boldsymbol{Z}_i\boldsymbol{T}\boldsymbol{Z}_i' + \sigma^2\boldsymbol{I}))$$

The covariance matrix $(\boldsymbol{Z}_i\boldsymbol{T}\boldsymbol{Z}_i' + \sigma^2\boldsymbol{I})$

- The (k,t) element of $\boldsymbol{Z}_i = \{z_{itk}\}$ for $k = 0, (q-1)$ and $t = 0, \dots r$.
- The covariance matrix for $\boldsymbol{U}_i$:

$$\boldsymbol{T} = \begin{pmatrix} \tau_{00} & \tau_{10} & \tau_{20} \\ \tau_{10} & \tau_{11} & \tau_{12} \\ \tau_{20} & \tau_{12} & \tau_{22} \end{pmatrix}$$

- (t,t') element of $\boldsymbol{Z}_i\boldsymbol{T}\boldsymbol{Z}_i' = \sum_{k=0}^{q}\sum_{\ell=k}^{(q)} \tau_{k\ell} z_{itk} z_{it'\ell}$

Marginal Model: Our Example

- (t,t) element of $\boldsymbol{Z}_i\boldsymbol{T}\boldsymbol{Z}_i' + \sigma^2\boldsymbol{I}$

$$\begin{aligned}\text{var}(Y_{it}) &= \tau_0^2 + \tau_1^2(\text{week})_{it}^2 + \tau_2^2(\text{week})_{it}^4 + 2\tau_{01}(\text{week})_{it} \\ &\quad +2\tau_{02}(\text{week})_{it}^2 + 2\tau_{12}(\text{week})_{it}^3 + \sigma^2\end{aligned}$$

- (t,t') element of $\boldsymbol{Z}_i\boldsymbol{T}\boldsymbol{Z}_i' + \sigma^2\boldsymbol{I}$

$$\begin{aligned}\text{cov}(Y_{it}, Y_{it'}) &= \tau_{00} + \tau_{10}(\text{week}_{it} + \text{week}_{it'}) + \tau_{20}(\text{week}_{it}^2 + \text{week}_{it'}^2) \\ &\quad +\tau_{11}(\text{week}_{it})(\text{week}_{it'}) + \tau_{22}(\text{week}_{it}^2)(\text{week}_{it'}^2) \\ &\quad +\tau_{12}(\text{week}_{it}^2)(\text{week}_{it'}) + \tau_{12}(\text{week}_{it})(\text{week}_{it'}^2) \\ &\quad +\sigma^2\end{aligned}$$

- $\text{cov}(Y_{it}, Y_{i't}) = \text{cov}(Y_{it}, Y_{i't'}) = 0.$

Covariance Parameter Estimates

	Null	Some Fixed	Preliminary
Var: id (Intercept)	13.62	15.28	21.11
Var: Residual	37.96	19.03	10.50
Var: id week			11.23
Var: id weeksq			0.20
Cov: id (Intercept) week			-11.07
Cov: id (Intercept) weeksq			1.10
Cov: id week weeksq			-1.33

$\hat{\rho} = 13.62/(13.62 + 37.95) = .26$

Solution for fixed Effects

	Empty/Null		Preliminary HLM	
Effect	Estimate	Std Error	Estimate	Std Error
Intercept	17.66	0.56	24.58	0.72
Week			-2.66	0.51
Week*Week			0.05	0.09
Endog $= 0$			-1.81	1.04
Endog $= 1$			0	.
Week*Endog $= 0$			.02	0.43
Week*Endog $= 1$			0	.

Global Fit Statistics

Model	-2LnLike	AIC	BIC.new
Empty/Null	2501.1	2507.1	2515.4
Preliminary HLM	2204.0	2228.0	2263.0
Preliminary HLM w/ cubic week	2201.7	2227.7	2266.6

- To get the Preliminary with week$_{it}^3$ (as fixed effect) to converge, I had to do re-scale (i.e., week$_{it}^3/10$).
- So far, go with preliminary

Model Reduction: Random Effects

(i.e., Covariance Structure)

- Test whether need random term for $(\text{week})^2_{it}$,

$$H_0 : \tau_2^2 = \tau_{02} = \tau_{12} = 0 \quad \text{versus} \quad H_a : \text{ not } H_0$$

- The Reduced Model,

$$\begin{aligned}(\text{HamD})_{it} &= \gamma_{00} + \gamma_{10}(\text{week})_{it} + \gamma_{20}(\text{week})^2_{it} + \gamma_{01}(\text{endog})_i \\ &\quad + \gamma_{11}(\text{endog})_i(\text{week})_{it} + U_{0i} + U_{1i}(\text{week})_{it} + R_{it},\end{aligned}$$

has $-2\text{lnLike} = \text{Deviance} = 2214.5.$

Model Reduction: Random Effects

- Test statistic: difference between -2lnLike of full and reduced models,

$$2214.5 - 2204.5 = 10.5$$

- Sampling distribution is a mixture of χ^2_3 and χ^2_2,

$$p\text{-value} = .5(.015) + .5(.005) = .01$$

- Conclusion: Reject H_0.
- AIC favors the model with U_{2i} (i.e. $AIC = 2232.5$ vs 2228.0) while BIC.new favors the model without U_{2i} (i.e., $BIC.new = 2260.88$ vs 2263.0).

Model Reduction: Fitted Effects

- Do not remove $(\text{week})_{it}$ or $(\text{week})^2_{it}$, because of non-zero τ_1^2 and τ_2^2.
- Possible Reductions: "endog" and "endog×week".
- t-tests indicate don't need these ; however,
- Likelihood ratio test statistic for "week*endog" (i.e., $H_o : \gamma_{11} = 0$ versus $H_a : \gamma_{11} \neq 0$),

$$= 2204.015 - 2204.013 = .001$$

  $df = 1$, $p = .97$...retain H_o (i.e. drop the interaction).
- Test for "endog", in a few pages.

Reduced Model Covariance Parameters

	Preliminary	No Interaction
Var: id (Intercept)	21.11	21.08
Var: id week	11.23	11.22
Var: id weeksq	0.20	0.20
Cov: id (Intercept) week	-11.07	-11.06
Cov: id (Intercept) weeksq	1.10	1.10
Cov: id week weeksq	-1.33	-1.33
Var: Residual	10.50	10.50

Very similar without endog $\times$ week

Reduced Model fixed Effects

	Preliminary		No Interaction	
	est.	se	est	se
(Intercept)	25.46	$(1.17)^{***}$	25.48	$(1.07)^{***}$
week	-2.75	$(0.68)^{***}$	-2.76	$(0.64)^{***}$
weeksq	0.05	(0.09)	0.05	(0.09)
endog	1.83	(1.28)	1.79	(0.92)
week:endog	-0.02	(0.42)		

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Estimates are pretty similar with and without endog $\times$ week.

Do we need "endog"?

$$H_o : \gamma_{01} = 0 \quad \text{versus} \quad H_a : \gamma_{01} \neq 0$$

- t-test using the estimates from the model without the cross-level interaction,

$$t = \frac{-1.79}{.92} = -1.94, \quad df = 65.7, \quad p - \text{value} = .056$$

- Likelihood ratio test statistic,

$$2207.648 - 2204.015 = 3.633$$

  Comparing this to χ^2_1, p-value=.056.

- Conclusion: maybe/undecided about "endog", keep it for now.

Global Fit Statistics

Model	-2LnLike	AIC	BIC.new
Empty/Null	2501.1	2507.1	2515.4
Preliminary HLM	2204.0	2228.0	2263.0
No endog$\times$week	2214.5	2232.5	2260.9
No endog$\times$week & No endog	2207.6	2227.6	2254.8

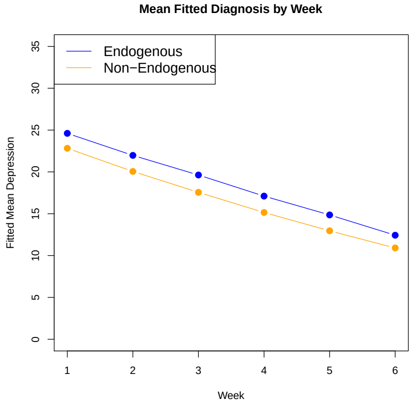
We'll go with this model (for now):

$$\begin{aligned} Y_{it} &= \beta_{0i} + \beta_{1i}(\text{week})_{it} + \beta_{2i}(\text{week})^2_{it} + R_{it} \\ \beta_{0i} &= \gamma_{00} + \text{endog}_i + U_{0i} \\ \beta_{1i} &= \gamma_{10} + U_{1i} \\ \beta_{2i} &= \gamma_{20} + U_{2i} \end{aligned}$$

What other analyses should we do to adequacy of this model?

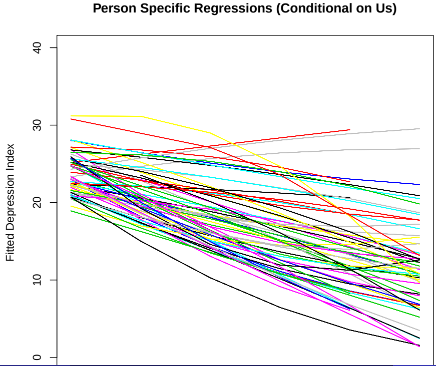
Interpretation

$$\widehat{(\text{HamD})}_{it} = 24.57 - 2.65(\text{week})_{it} + .05(\text{week})^2_{it} - 1.79D_i$$



Estimated Model Individuals

$$\begin{aligned}\widehat{(\text{HamD})}_{it} &= 25.48 - 2.76(\text{week})_{it} + .05(\text{week})^2_{it} \\ &\quad -1.79D_i + \hat{U}_{0i} + \hat{U}_{1i}(\text{week})_{it} + \hat{U}_{2i}(\text{week})^2_{it}\end{aligned}$$



Summary

- Just an HLM
- To study the nature of change & requires systematic change.
- Level 1:
  - Time is a level 1 variable (should have at least 3 time points).
  - Need to choose metric of time.
  - "time variant" variables are level 1 variables.
- Level 2:
  - Study differences between individuals.
  - "time invariant" variables are level 2 variables.
- What about covariance matrix for R_{it}? ...