Longitudinal Data Analysis via Linear Mixed Model
Edps/Psych/Stat 587

Carolyn J. Anderson

Department of Educational Psychology

© Board of Trustees, University of Illinois

Fall 2013
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:

- Notes by Donald Hedeker. Available from his web-site http://tigger.uic.edu/~hedeker.
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:

\[
\hat{\text{reading}}_i = 111.40 - 8.19(\text{age})_i;
\]

Hypothetical Longitudinal Data

Reading Ability = 111.40 - 8.19 (Age)

\[ R^2 = .48 \]
Cross-Sectional Data (continued)

Occasion 1: \[
\text{reading}_{i1} = 111.86 - 10.18(\text{age})_{i1}
\]

Occasion 2: \[
\text{reading}_{i2} = 140.01 - 10.50(\text{age})_{i2}
\]

Treating Occasions Separately
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).

- \( \beta_{cs} \) still reflects differences between individuals.

- \( \beta_{cs} \) and \( \beta_l \) reflect different processes.
A Model for Longitudinal Data

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

Model w/in and btw Individuals
Advantages: Longitudinal Data

More Powerful.

- Inference regarding $\beta_{cs}$ is a comparison of individuals with the same value of $x$.

- Inference regarding $\beta_I$ is a comparison of an individual's response at two times

  $\Rightarrow$ Assuming $y$ changes **systematically** with time and retains its **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.

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.
Metric for Time

Example from Singer & Willett:

If you want to study the “longevity” of automobiles.

- Change in appearance of cars $\rightarrow$ Age.
- Tire wear $\rightarrow$ Miles.
- Wear of ignition system $\rightarrow$ Trips (# of starts).
- Engine wear $\rightarrow$ 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 analyzing 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., a conditional model).

- **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,
  $$E(Y_{it}) = X_i \Gamma.$$

- Observations are correlated, so need to make adjustments to variance estimates, i.e., $\text{var}(Y_i) = V_i(\theta)$ where $\theta$ are parameters.

- “Sandwich estimator” or Robust estimation.
Transition Models

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

- Model the **Conditional Distribution** of \( Y_{it} \),

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

- 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.,
  \[
  E(Y_{it}|\beta_1, \ldots, \beta_p) = \sum_{k=1}^{p} \beta_{ik}x_{ikt}
  \]

- One individual’s data does not contain enough information to estimate \(\beta_{ik}\)'s; therefore, we assume a distribution for \(\beta_{ik}\)'s,
  \[
  \beta_i = Z_i\Gamma + U_i
  \]
  where \(U_i \sim \mathcal{N}(0, T)\) 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,

\[ R_i \sim \mathcal{N}(0, \Sigma_i) \]

where \( \Sigma_i = \sigma^2 I \) (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(\#).
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

<table>
<thead>
<tr>
<th>Drug-Washout</th>
<th>day 0</th>
<th>day 7</th>
<th>day 14</th>
<th>day 21</th>
<th>day 28</th>
<th>day 35</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>wk 0</td>
<td>wk 1</td>
<td>wk 2</td>
<td>wk 3</td>
<td>wk 4</td>
<td>wk 5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$Y_{it}$</th>
<th>Hamilton Depression</th>
<th>$HD_1$</th>
<th>$HD_2$</th>
<th>$HD_3$</th>
<th>$HD_4$</th>
<th>$HD_5$</th>
<th>$HD_6$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level 2</td>
<td>Gender</td>
<td>$G$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diagnosis</td>
<td></td>
<td>$D$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level 1</td>
<td>IMI</td>
<td>—</td>
<td>—</td>
<td>$IMI_3$</td>
<td>$IMI_4$</td>
<td>$IMI_5$</td>
<td>$IMI_6$</td>
</tr>
<tr>
<td>DMI</td>
<td>—</td>
<td>—</td>
<td>$DMI_3$</td>
<td>$DMI_4$</td>
<td>$DMI_5$</td>
<td>$DMI_6$</td>
<td></td>
</tr>
<tr>
<td>$N$</td>
<td>61</td>
<td>63</td>
<td>65</td>
<td>65</td>
<td>63</td>
<td>58</td>
<td></td>
</tr>
</tbody>
</table>

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.
- Impipramine 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
Overlay Individual Regressions

Linear Regressions

![Graph showing individual linear regressions over time.](image)
Overlay Individual Regressions

Quadratic Regressions

Depression Index

Time (weeks)
Overlay Individual Regressions

Splines

Depression Index

Time (weeks)
Exploring Mean Structure

General linear decline and increasing variance.

Riesby Data (n=66)
Exploring Mean Structure (continued)

Riesby Data (averages)

![Graph showing the trend of Hamilton Depression Index over weeks.](image)
Exploring Mean Structure (continued)

Mean Hamilton Depression Index
Endogenous and Non-endogenous

![Graph showing mean Hamilton Depression Index over weeks for endogenous and non-endogenous categories.](image-url)
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:

- \( R^2_i = \frac{\text{SSMOD}_i}{\text{SSTO}_i} \).
- \( R^2_{\text{meta}} = \frac{\sum_i \text{SSMOD}_i}{\sum_i \text{SSTO}_i} \).
- And if there are two possible (nested) models, Compute an \( F \)-statistic.
- Try to see who improves.
Linear Model: $R^2_i$ and $R^2_{\text{meta}}$

Fit statistics for the model

$$Y_{it} = \beta_0i + \beta_{i1}(\text{week})_{it} + \epsilon_{it}$$
Quadratic: $R^2_i$ and $R^2_{meta}$

$$Y_{it} = \beta_{0i} + \beta_{i1}(\text{week})_{it} + \beta_{i2}(\text{week})^2_{it} + \epsilon_{it}$$

**Within Individual Fits**

*Model: HamD = week week^2*
Comparison

Changes in $R^2_i$
Quadratic & Linear Models

![Graph showing changes in $R^2_i$ between quadratic and linear models.]
Who Improved?
Who Improved? (continued)

Improvement in Model Fit: Linear to quadratic
$F$ “Test” for Quadratic Term

- **Reduced Model:** $Y_{it} = \beta_0 + \beta_1(\text{week})_{it} + \epsilon_{it}$
  
  $p = 2$.

- **Full Model:** $Y_{it} = \beta_0 + \beta_1(\text{week})_{it} + \beta_2(\text{week})_{it}^2 + \epsilon_{it}$
  
  $p^* = 1$.

- **$F$-statistic:**

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

  Comparing $F = 1.55$ to the $F$–distribution with $df_{num} = 66$ and $df_{den} = 177$, the “$p$-value” = .01.
Preliminary HLM

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

- **Level 2:**
  
  $\beta_{oi} = \gamma_{00} + \gamma_{01}(\text{endog})_i$
  
  $\beta_{1i} = \gamma_{10}$
  
  $\beta_{2i} = \gamma_{20}$

- **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,
  \[ \hat{R}_{it} = (Y_{it} - \hat{Y}_{it}) = Z_i U_i + R_i. \]

- Squared residuals, \( \hat{R}_{it}^2 \).

- Correlations between residuals to look for serial correlation (i.e., need model for \( \Sigma_i \)?)
  \[ \text{corr}(R_{it}, R_{it'}). \]
Raw Residuals by Week

Raw Residuals by Individual

Model: HamD = week + week^2
Raw Residuals by Week

Raw Residuals by Individual

Model: $\text{HamD} = \text{week} + \text{week}^2$
Raw Residuals with Endog

**Model:** $\text{HamD} = \text{week} + \text{week}^2$

Raw Residuals by Individual

Endog = 0

Endog = 1
Mean Raw Residuals

Mean Raw Residual

Week

Raw Residual

0 1 2 3 4 5

-0.8 -0.6 -0.4 -0.2 0.0 0.2 0.4 0.6 0.8 1.0

C.J. Anderson (Illinois)

Longitudinal Data Analysis via Linear Mixed Model

Fall 2013
Squared Raw Residuals

**Model:** $\text{HamD} = \text{week} + \text{week}^2$

Squared Residuals Joined by Spline

Endog = 0  
Endog = 1
Mean Squared Raw Residuals

Mean Squared Residuals

Spline

Cubic Regression

Quadratic Regression

Linear Regression
Variance Function

Given Preliminary Mixed Linear Model:

\[ Y_{it} = \gamma_0 + \gamma_{10}(endog)_{it} + \gamma_{11}(week)_{it} + \gamma_{20}(week)^2_{it} + R_{it} \]

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

\[ (U_{0i}, U_{1i}, U_{2i})' \sim \mathcal{N}(0, T) \]

The variance of \( Y_{it} \)

\[ \text{var}(Y_{it}) = \tau_0^2 + \tau_1^2 \text{week}_{it}^2 + \tau_2^4 \text{week}_{it}^4 + 2\tau_{01} \text{week}_{it} + 2\tau_{02} \text{week}_{it}^2 + 2\tau_{12} \text{week}_{it}^3 + \sigma^2 \]
Variance Function: 2 Random Effects

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

\[ (U_{0i}, U_{1i})' \sim \mathcal{N}(0, T) \]

- 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($\hat{R}_{it}, \hat{R}_{it'}$).

<table>
<thead>
<tr>
<th></th>
<th>week0</th>
<th>week1</th>
<th>week2</th>
<th>week3</th>
<th>week4</th>
<th>week5</th>
</tr>
</thead>
<tbody>
<tr>
<td>week0</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>week1</td>
<td>.47</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>week2</td>
<td>.39</td>
<td>.47</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>week3</td>
<td>.32</td>
<td>.39</td>
<td>.73</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>week4</td>
<td>.22</td>
<td>.28</td>
<td>.66</td>
<td>.81</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>week5</td>
<td>.17</td>
<td>.19</td>
<td>.45</td>
<td>.56</td>
<td>.65</td>
<td>1.00</td>
</tr>
</tbody>
</table>

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

Correlations Between Time Points

Residual Correlations

Week 0  Week 1  Week 2  Week 3  Week 4

Lag 1

Week
Plot of Correlations: Lags

Correlations Between Time Points
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)

\[(\text{HamD})_{it} = \beta_0i + \beta_1i(\text{week})_{it} + \beta_2i(\text{week})_{it}^2 + R_{it}\]

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

- **Between Individuals** (level 2)

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

where \(U_i \sim \mathcal{N}(0, \mathbf{T})\).
Linear Mixed Model

- Scalar form

\[
(HamD)_{it} = \gamma_0 + \gamma_{10}(week)_{it} + \gamma_{20}(week)_{it}^2 + \gamma_{01}(endog)_i + \gamma_{11}(endog)_i(week)_{it} + U_{0i} + U_{1i}(week)_{it} + U_{2i}(week)_{it}^2 + R_{it}
\]

- In matrix form,

\[
Y_i = X_i\Gamma + Z_iU_i + R_i
\]
**Linear Mixed Model**

\[ Y_i = X_i \Gamma + Z_i U_i + R \]

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

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

\[ D_i = \begin{cases} 
0 & \text{Non-endogenous} \\
1 & \text{endogenous} 
\end{cases} \]
Marginal Model

\[(\text{HamD})_{it} \sim \mathcal{N}(X_i \Gamma, (Z_i T Z'_i + \sigma^2 I))\]

The covariance matrix \((Z_i T Z'_i + \sigma^2 I)\)

- The \((k, t)\) element of \(Z_i = \{z_{itk}\}\) for \(k = 0, (q - 1)\) and \(t = 0, \ldots r\).

- The covariance matrix for \(U_i\):

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

- \((t, t')\) element of \(Z_i T Z'_i = \sum_{k=0}^{q} \sum_{\ell=k}^{(q)} \tau_{k\ell} z_{itk} Z_{i\ell'}\)
Marginal Model: Our Example

- $(t, t)$ element of $Z_iTZ'_i + \sigma^2 I$

\[
\text{var}(Y_{it}) = \tau_0^2 + \tau_1^2(\text{week})_{it}^2 + \tau_2^2(\text{week})_{it}^4 + 2\tau_{01}(\text{week})_{it}
\]
\[
+ 2\tau_{02}(\text{week})_{it}^2 + 2\tau_{12}(\text{week})_{it}^3 + \sigma^2
\]

- $(t, t')$ element of $Z_iTZ'_i + \sigma^2 I$

\[
\text{cov}(Y_{it}, Y_{it'}) = \tau_{00} + \tau_{10}(\text{week}_{it} + \text{week}_{it'}) + \tau_{20}(\text{week}_{it}^2 + \text{week}_{it'}^2)
\]
\[
+ \tau_{11}(\text{week}_{it})(\text{week}_{it'}) + \tau_{22}(\text{week}_{it}^2)(\text{week}_{it'}^2)
\]
\[
+ \tau_{12}(\text{week}_{it}^2)(\text{week}_{it'}) + \tau_{12}(\text{week}_{it})(\text{week}_{it'}^2) + \sigma^2
\]

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

<table>
<thead>
<tr>
<th>Cov Parm</th>
<th>Empty/Null</th>
<th>Preliminary</th>
<th>HLM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
<td>Std Error</td>
<td>Estimate</td>
</tr>
<tr>
<td>UN(1,1)</td>
<td>13.62</td>
<td>3.58</td>
<td>9.91</td>
</tr>
<tr>
<td>UN(2,1)</td>
<td>-1.25</td>
<td>2.41</td>
<td></td>
</tr>
<tr>
<td>UN(2,2)</td>
<td>6.67</td>
<td>2.77</td>
<td></td>
</tr>
<tr>
<td>UN(3,1)</td>
<td>-0.04</td>
<td>0.42</td>
<td></td>
</tr>
<tr>
<td>UN(3,2)</td>
<td>-0.94</td>
<td>0.48</td>
<td></td>
</tr>
<tr>
<td>UN(3,3)</td>
<td>0.19</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td>Residual</td>
<td>37.95</td>
<td>3.05</td>
<td>10.50</td>
</tr>
</tbody>
</table>

\[
\hat{\rho} = \frac{13.62}{(13.62 + 37.95)} = .26
\]
### Solution for fixed Effects

<table>
<thead>
<tr>
<th>Effect</th>
<th>Empty/Null Estimate</th>
<th>Empty/Null Std Error</th>
<th>Preliminary HLM Estimate</th>
<th>Preliminary Std Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>17.66</td>
<td>0.56</td>
<td>24.58</td>
<td>0.72</td>
</tr>
<tr>
<td>Week</td>
<td>-2.66</td>
<td>0.51</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Week*Week</td>
<td>0.05</td>
<td>0.09</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Endog = 0</td>
<td>-1.81</td>
<td>1.04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Endog = 1</td>
<td>0</td>
<td>.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Week*Endog = 0</td>
<td>0.02</td>
<td>0.43</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Week*Endog = 1</td>
<td>0</td>
<td>.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

hline
### Global Fit Statistics

<table>
<thead>
<tr>
<th>Model</th>
<th>$-2\text{LnLike}$</th>
<th>AIC</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Empty/Null</td>
<td>2501.1</td>
<td>2507.1</td>
<td>2513.7</td>
</tr>
<tr>
<td>Preliminary HLM</td>
<td>2204.0</td>
<td>2228.0</td>
<td>2254.3</td>
</tr>
<tr>
<td>Preliminary HLM w/ cubic week</td>
<td>2192.7</td>
<td>2226.7</td>
<td>2264.0</td>
</tr>
</tbody>
</table>
Model Reduction: Random Effects

(i.e., Covariance Structure)

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

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

- The Reduced Model,

\[
(\text{HamD})_{it} = \gamma_{00} + \gamma_{10}(\text{week})_{it} + \gamma_{20}(\text{week})_{it}^2 + \gamma_{01}(\text{endog})_i \\
+ \gamma_{11}(\text{endog})_i(\text{week})_{it} + U_{0i} + U_{1i}(\text{week})_{it} + R_{it},
\]

has \(-2\ln\text{Like} = 2214.5\).
Model Reduction: Random Effects

- Test statistic: difference between \(-2\ln\text{Like}\) of full and reduced models,
  \[
  2214.5 - 2204.5 = 10.5
  \]

- Sampling distribution is a mixture of \(\chi^2_3\) and \(\chi^2_2\),
  \[p\text{-value} = .5(.015) + .5(.005) = .01\]

- Conclusion: Reject \(H_0\).

- AIC favors the model without \(U_{2i}\) while BIC favors the model with \(U_{2i}\).
Model Reduction: Fitted Effects

- **Do not remove** \((\text{week})_{it}\) or \((\text{week})^2_{it}\), because of non-zero \(\tau_1^2\) and \(\tau_2^2\).

- Possible Reductions: “endog” and “endog*week”.

- \(t\)-tests indicate don’t need these; however,

- Likelihood ratio test statistic for “week*endog” (i.e., \(H_0: \gamma_{11} = 0\) versus \(H_a: \gamma_{11} \neq 0\)),

\[
= 2204.015 - 2204.013 = .002
\]

\(df = 1\), \(p\)-value large...retain \(H_0\) (i.e. drop the interaction).

- After removing “week*endog”, we do a likelihood ratio test for “endog”
### Reduced Model Covariance Parameters

<table>
<thead>
<tr>
<th>Cov Parm</th>
<th>Preliminary Estimate</th>
<th>HLM Std Error</th>
<th>No interaction Estimate</th>
<th>Std Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>UN(1,1)</td>
<td>9.91</td>
<td>3.49</td>
<td>9.01</td>
<td>3.49</td>
</tr>
<tr>
<td>UN(2,1)</td>
<td>-1.25</td>
<td>2.41</td>
<td>-1.25</td>
<td>2.42</td>
</tr>
<tr>
<td>UN(2,2)</td>
<td>6.67</td>
<td>2.77</td>
<td>6.67</td>
<td>2.77</td>
</tr>
<tr>
<td>UN(3,1)</td>
<td>-0.04</td>
<td>.42</td>
<td>-.04</td>
<td>.42</td>
</tr>
<tr>
<td>UN(3,2)</td>
<td>-0.94</td>
<td>.48</td>
<td>-.94</td>
<td>.49</td>
</tr>
<tr>
<td>UN(3,3)</td>
<td>0.19</td>
<td>.09</td>
<td>.20</td>
<td>.10</td>
</tr>
<tr>
<td>Residual</td>
<td>10.50</td>
<td>1.04</td>
<td>10.51</td>
<td>1.10</td>
</tr>
</tbody>
</table>
Reduced Model fixed Effects

<table>
<thead>
<tr>
<th>Cov Parm</th>
<th>Preliminary</th>
<th>HLM</th>
<th>No interaction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
<td>Std Error</td>
<td>Estimate</td>
</tr>
<tr>
<td>Intercept</td>
<td>24.58</td>
<td>.72</td>
<td>24.57</td>
</tr>
<tr>
<td>Week</td>
<td>-2.66</td>
<td>.51</td>
<td>-2.65</td>
</tr>
<tr>
<td>Week*Week</td>
<td>.05</td>
<td>.09</td>
<td>.05</td>
</tr>
<tr>
<td>Endog = 0</td>
<td>-1.81</td>
<td>1.04</td>
<td>-1.79</td>
</tr>
<tr>
<td>Endog = 1</td>
<td>0</td>
<td>.</td>
<td>0</td>
</tr>
<tr>
<td>Week*Endog = 0</td>
<td>.02</td>
<td>.43</td>
<td></td>
</tr>
<tr>
<td>Week*Endog = 1</td>
<td>0</td>
<td>.</td>
<td></td>
</tr>
</tbody>
</table>

Estimates are pretty similar.
Do we need “endog”?

\[ H_0 : \gamma_{01} = 0 \quad \text{versus} \quad H_a : \gamma_{01} \neq 0 \]

- \( t \)-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 \( \chi^2_1 \), \( p \)-value = .056.

Conclusion: maybe/undecided about “endog”, keep it for now.
### Global Fit Statistics

<table>
<thead>
<tr>
<th>Model</th>
<th>-2LnLike</th>
<th>AIC</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Empty/Null</td>
<td>2501.1</td>
<td>2507.1</td>
<td>2513.7</td>
</tr>
<tr>
<td>Preliminary HLM</td>
<td>2204.0</td>
<td>2228.0</td>
<td>2254.3</td>
</tr>
<tr>
<td>No “endog*week”</td>
<td>2204.0</td>
<td>2226.0</td>
<td>2250.1</td>
</tr>
<tr>
<td>No “endog*week” and no “endog”</td>
<td>2207.6</td>
<td>2227.6</td>
<td>2249.5</td>
</tr>
</tbody>
</table>

We’ll go with this model:

\[
Y_{it} = \beta_{0i} + \beta_{1i}(week)_{it} + \beta_{2i}(week)_{it}^2 + R_{it}
\]

\[
\beta_{0i} = \gamma_{00} + \text{endog}_i + U_{0i}
\]

\[
\beta_{1i} = \gamma_{10} + U_{1i}
\]

\[
\beta_{2i} = \gamma_{20} + U_{2i}
\]

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

\[ (\text{HamD})_{it} = 24.57 - 2.65(\text{week})_{it} + 0.05(\text{week})^2_{it} - 1.79D_i \]
Estimated Model Individuals

\[
(\text{HamD})_{it} = 24.57 - 2.65(\text{week})_{it} + 0.05(\text{week})_{it}^2 \\
- 1.79D_i + \hat{U}_0 + \hat{U}_1(\text{week})_{it} + \hat{U}_2(\text{week})_{it}^2
\]