"Random" or "Mixed" Effect Subject-Specific Linear Models

- The **linear mixed model** for longitudinal data is

$$Y_{ij} = \boldsymbol{x}'_{ij}\boldsymbol{\beta} + \underbrace{\boldsymbol{d}'_{ij}\boldsymbol{U}_i + Z_{ij}}_{\epsilon_{ij}}$$

where

$$\boldsymbol{U}_i \sim N(0, G) \quad \text{and} \quad Z_{ij} \sim N(0, \tau^2)$$

  - Z_{ij}'s are independent of $\boldsymbol{U}_i$, of each other and of X_i (and D_i)
  - $\boldsymbol{U}_i$'s are independent of X_i and D_i (include all $\boldsymbol{d}_{ij}$ within subject i)

- Note: The distributional assumption on $\boldsymbol{U}_i$ differentiates this model from the "fixed effects" models in previous note

- The $\boldsymbol{\beta}$'s are often called **fixed effects** (fixed across subjects) by biostatisticians;
  – Different from the use of "fixed effects" by econometricians in "fixed effect models"
- Like FE, $\boldsymbol{d}_{ij}$ is usually a set of observed covariates for the ith subject at the jth time, with **subject-specific** coefficient U_i
  – But here U_i's vary according to a distributional model

Conditional Models and Marginal Models

- Conditional on $\boldsymbol{U}_i$ (and (D_i, X_i))

$$\mathrm{E}(Y_{ij}|\boldsymbol{U}_i, D_i, X_i) = \boldsymbol{x}'_{ij}\boldsymbol{\beta} + \boldsymbol{d}'_{ij}\boldsymbol{U}_i$$

  is a **subject specific** model (**conditional model**) for the (mean response on the) ith subject

- However, we also can write

$$Y_{ij} = \boldsymbol{x}'_{ij}\boldsymbol{\beta} + \epsilon_{ij} \tag{1}$$

$$\epsilon_{ij} = \boldsymbol{d}'_{ij}\boldsymbol{U}_i + Z_{ij} \ , \tag{2}$$

- Models (1) and (2) together specify a **marginal model** $\mathrm{E}(Y_{ij}|X_i) = X_i\boldsymbol{\beta}$ for the mean and for the v-c-c model for the residuals ϵ_{ij}

- Unconditional Expectation (average over U):

$$\mathrm{E}(Y_{ij}|X_i) = \boldsymbol{x}'_{ij}\boldsymbol{\beta} + \mathrm{E}(\boldsymbol{d}'_{ij}\boldsymbol{U}_i + Z_{ij}) = \boldsymbol{x}'_{ij}\boldsymbol{\beta}$$

  – $\boldsymbol{x}_{ij}$: covariates for the ith subject at the jth measurement time that affect the **marginal mean** of Y_{ij}.

  – this is also a **population average** (PA) or **marginal** model

- This marginal model interpretation would **not** hold for fixed effects models (Why?)

- Note: For generalized linear model with non-linear link (eg, Poisson model for count data), $\boldsymbol{\beta}$ under marginal model is different to conditional model
  – Conditional model:

$$\log(\mathrm{E}(Y_{ij}|\boldsymbol{U}_i, D_i, X_i)) = \boldsymbol{x}'_{ij}\boldsymbol{\beta} + \boldsymbol{d}'_{ij}\boldsymbol{U}_i$$

  does not lead to
  – Marginal model (average over U):

$$\log(\mathrm{E}(Y_{ij}|X_i)) = \boldsymbol{x}'_{ij}\boldsymbol{\beta}$$

  (more details later)

- For the marginal model, the subject-level residuals are written

$$\boldsymbol{\epsilon}_i = D_i \boldsymbol{U}_i + \boldsymbol{Z}_i$$

where

$$\boldsymbol{\epsilon}_i = (\epsilon_{i1}, \ldots, \epsilon_{in_i})' \text{ is } n_i \times 1,$$

$$D_i = (\boldsymbol{d}_{i1}, \ldots, \boldsymbol{d}_{in_i})' \text{ is } n_i \times q$$

(q is number of variables in random effects, including intercept),

$$\boldsymbol{Z}_i = (Z_{i1}, \ldots, Z_{in_i})' \text{ is } n_i \times 1,$$

$$\boldsymbol{U}_i = (\boldsymbol{U}_{i1}, \ldots, \boldsymbol{U}_{iq})' \text{ is } q \times n_i.$$

and

$$\text{var}(D_i \boldsymbol{U}_i) = D_i \text{var}(U_i) D_i' = D_i G D_i' \quad \text{and} \quad \text{var}(\boldsymbol{Z}_i) = \tau^2 I_{n_i}$$

so that

$$V_i = \mathrm{var}(\boldsymbol{\epsilon}_i) = \mathrm{var}(D_i \boldsymbol{U}_i + \boldsymbol{Z}_i) = D_i G D_i' + \tau^2 I_{n_i}$$

- Important idea for subject-specific model-fitting:
  Equivalent to **marginal model**
  – $\mathrm{E}(Y_{ij}|X_i) = X_i\boldsymbol{\beta}$ for the mean
  – v-c-c model for the residuals ϵ_{ij}
- Implication: Can use all of the sound statistical theory and techniques (and software) for marginal models
  – joint **maximum likelihood** (ML) estimation for the parameters $(\boldsymbol{\beta}, G, \tau^2)$
  – **restricted maximum likelihood estimation** (ReML) for variance-covariance parameters (G, τ^2), and subsequent WLS for $\boldsymbol{\beta}$

- Note: **A more general model**:

$$\epsilon_{ij} = \boldsymbol{d}'_{ij}\boldsymbol{U}_i + W_{ij} + Z_{ij}$$

  - W_{ij}'s are correlated within subject
  - Z_{ij}s are i.i.d. measurement error
  - we will have

$$\boldsymbol{\epsilon}_i \sim N(0, V_i)$$

    where

$$V_i = D_i G D'_i + R_i$$

$$R_i = \mathrm{var}(\boldsymbol{W}_i + \boldsymbol{Z}_i)$$

  - In SAS:
    - `random`: specify $\boldsymbol{d}'_{ij}\boldsymbol{U}_i$
    - `repeated`: specify W_{ij}
    - `repeated` with `local` option: specify Z_{ij}

– Special case: When $\boldsymbol{d}_{ij} = 1$ (random intercept only), W_{ij} has an exponential correlation structure

$$\epsilon_{ij} = U_i + W_{ij} + Z_{ij}$$

giving the "exchangeable + exponential + measurement error" v-c-c structure

Linear Mixed Effects Models
Example: Depression Patients

- Study goal: relationship between Imipramine (IMI) and Desipramine (DMI) plasma levels and outcomes (Riesby and others, 1977)
- $m = 66$ depressed in-patients
- Subjects were assessed on days 0, 7, . . . , 35
- Main outcome: Hamilton depression score, measured at each time
- Baseline covariate: diagnosis group (endogenous versus non-endogenous depression)
- Time-varying covariates: imipramine (IMI) drug-plasma level and desipramine (DMI) drug-plasma level
- No drug was administered during weeks 0 and 1

- Some data description:

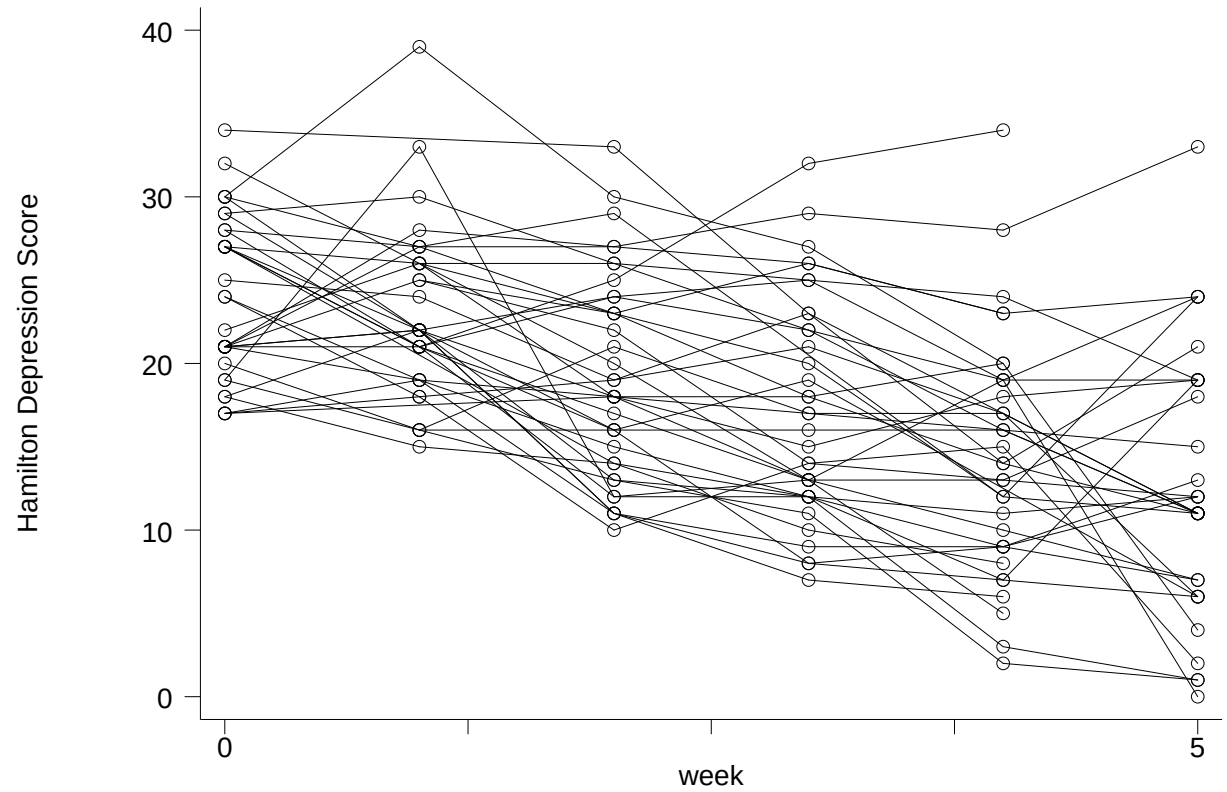
```
data riesby;
set riesby;
endwk=endog*week;
wkcen = week-2.5;
endwkcen = endog*wkcen;
wksqr = wkcen*wkcen;
run;
proc freq data=riesby;
table endog;
run;
proc freq data=riesby(where=(hamd ne .));
table week;
run;
```

endog	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	174	43.94	174	43.94
1	222	56.06	396	100.00

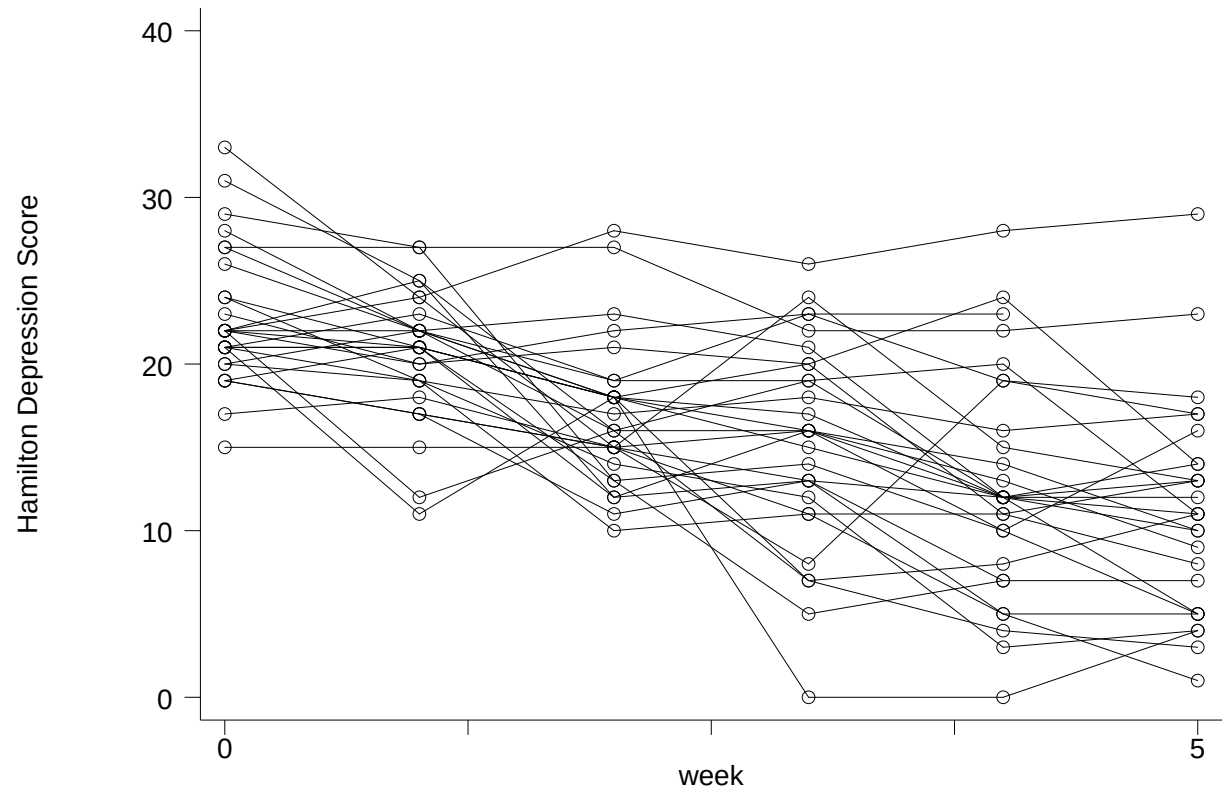
week	Freq.	Percent	Cum.
0	61	16.27	16.27
1	63	16.80	33.07
2	65	17.33	50.40
3	65	17.33	67.73
4	63	16.80	84.53
5	58	15.47	100.00
Total	375	100.00	

- Examine response by week and by diagnosis group

  Endogenous group:



Non-endogenous group:



- Some things to note:
  - Depression scores go down over time
  - The non-endogenous group has generally lower scores
  - The endogenous group may go down more rapidly over time
  - The standard deviation increases over time
  - Subjects "track" fairly well
  - Modeling such **trajectories** is what random (mixed) effects models are good for
- How can we account for these features of the data in a **mixed effects model**?
- What about letting Y_{ij} be the depression score for subject i at time j, and considering the model

$$Y_{ij} = \beta_0 + \beta_1 \texttt{endog}_i + \beta_2 \texttt{week}_{ij} + \beta_3 \texttt{endog}_i * \texttt{week}_{ij} + U_{i1} + U_{i2} \texttt{week}_{ij} + Z_{ij}$$

which has a random intercept U_{i1} and a random slope U_{i2}

- The model can be rewritten (in HLM form) as

$$Y_{ij} = \underbrace{(\beta_0 + \beta_1 \texttt{endog}_i + U_{i1})}_{b_{i0}} + \underbrace{(\beta_2 + \beta_3 \texttt{endog}_i + U_{i2})}_{b_{i1}} \texttt{week}_{ij} + Z_{ij}$$

where we can see that

  - $(\beta_0 + \beta_1 \texttt{endog}_i + U_{i1}) = b_{i0}$ is a subject-specific intercept
  - $(\beta_2 + \beta_3 \texttt{endog}_i + U_{i2}) = b_{i1}$ is a subject-specific slope w.r.t. week

- Now lets fit the model:
  - first use ReML for v-c-c parameters (G, τ^2)
  - then WLS for $\boldsymbol{\beta}$

- SAS code:

```
proc mixed data=riesby covtest;
class id;
model hamd=endog week endwk/ s outpm=pred;
random intercept week / subject=id type=un g ;
contrast 'F test for overall effect over time of endogenous'
endog 1 ,
endwk 1 ;
run;
```

  – `random intercept week`: Include random intercept and slope w.r.t. week
  – `type=un`: do not place any structural constraints on $\mathrm{cov}(\boldsymbol{U}_i) = G$

Output:

Estimated G Matrix

Row	Effect	id	Col1	Col2
1	Intercept	101	12.2506	-1.5150
2	week	101	-1.5150	2.1719

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z
UN(1,1)	id	12.2506	3.5158	3.48	0.0002
UN(2,1)	id	-1.5150	1.0640	-1.42	0.1545
UN(2,2)	id	2.1719	0.5406	4.02	<.0001
Residual		12.2109	1.1180	10.92	<.0001

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	22.4760	0.8074	64	27.84	<.0001
endog	1.9882	1.0864	243	1.83	0.0685
week	-2.3657	0.3170	64	-7.46	<.0001
endwk	-0.02680	0.4264	243	-0.06	0.9499

- Parameter interpretation:
  - **Population average** (marginal model or OLS) interpretation of $\widehat{\boldsymbol{\beta}}$ (thinking of $\boldsymbol{U}_i$ as part of the error $\boldsymbol{\epsilon}_i$):
    * $\widehat{\beta}_0 =$ estimated average depression score for non-engogenous subjects at week 0 (baseline)
    * $\widehat{\beta}_1 = 1.99$ is the estimated difference in **average depression scores** at baseline between endogenous and non-endogenous subjects
    * $\beta_2 = -2.37$ is the estimated slope (wrt week) of the **average depression scores** for non-endogenous subjects
  - **Subject-specific** interpretation of $\widehat{\boldsymbol{\beta}}$ (building on HLM form of the model):
    * $\widehat{\beta}_0 = 22.47$ is the estimated average (over $\boldsymbol{U}_i$) **subject-specific intercept** of depression scores for non-endogenous subjects
    * $\widehat{(\beta_0 + \beta_1)} = 22.48 + 1.99 = 24.47 =$ the estimated average (over $\boldsymbol{U}_i$) subject-specific intercept of depression scores for endogenous

subjects

* $\widehat{\beta}_2 = -2.37$ is the estimated average **subject-specific** slope of depression scores for non-endogenous subjects
* $(\beta_2 + \beta_3) = -2.37 - 0.03 = -2.40$ is the estimated average **subject-specific** slope of depression scores for endogenous subjects

- **Interpretation of variance components** $\widehat{G}$:
  - $\widehat{\text{var}}(U_{i1}) = 12.25$ is the estimated variance of **subject-specific** intercepts (at time 0) for depression score among subjects in either diagnosis group (after removing mean differences due to endogenous vs. non-endogenous groups)
  - $\widehat{\text{var}}(U_{i2}) = 2.17$ is the estimated variance of **subject-specific** slopes in depression score among subjects in either diagnosis group (again after removing differences in mean slope due to endogenous vs. non-endogenous groups)
  - $\widehat{\text{cov}}(U_{i1}, U_{i2}) = -1.52$ is the estimated covariance between subject-specific intercept and slope in depression score among subjects in either diagnosis group
  - Fitted random slope and intercept are negatively correlated.

  - **Example**: What is the (estimated) variance of subject-specific mean depression scores at week 5?

Recall,

$$Y_{ij} = \beta_0 + \beta_1 \texttt{endog}_i + \beta_2 \texttt{week}_{ij} + \beta_3 \texttt{endog}_i * \texttt{week}_{ij} + U_{i1} + U_{i2}\texttt{week}_{ij} + Z_{ij}$$

Then subject-specific mean depression scores at week t is

$$\beta_0 + \beta_1 \texttt{endog}_i + \beta_2 t + \beta_3 \texttt{endog}_i * t + U_{i1} + U_{i2} t$$

When $t = 5$, its variance (conditional on X) is

$$\begin{aligned} & \widehat{\text{var}}(U_{i1} + 5U_{i2}) \\ = \; & \widehat{\text{var}}(U_{i1}) + 5^2 \widehat{\text{var}}(U_{i2}) + 2 \times 5 \times \widehat{\text{cov}}(U_{i1}, U_{i2}) \\ = \; & 51.3 \end{aligned}$$

* larger than variance at baseline ($\widehat{\text{var}}(U_{i1})$), as expected from our original exploration of the data
* RE models deal with heteroscedasticity as a by-product

* Note: In Stata you can fit the same model with
  `. xtmixed hamd endog week endwk || id: week , cov(uns) , var`

* Details of Stata syntax:
  - `||` – divides specification of $\boldsymbol{x}'_{ij}\boldsymbol{\beta}$ from $\boldsymbol{d}'_{ij}\boldsymbol{U}_i$
  - `id: week` – include a random intercept and slope (wrt week) at level of `id`
  - `cov(uns)` – do not place any structural constraints on $\mathrm{cov}(\boldsymbol{U}_i) = G$

Random (Mixed) Effects Models
Important Ideas

Recall, with HLM form:

$$Y_{ij} = \underbrace{(\beta_0 + \beta_1 \texttt{endog}_i + U_{i1})}_{b_{i0}} + \underbrace{(\beta_2 + \beta_3 \texttt{endog}_i + U_{i2})}_{b_{i1}} \texttt{week}_{ij} + Z_{ij}$$

- Explicit model for individual change over time:
  - **within subject** variation:
    - Subject-specific slope with respect to week for subject i is:

$$b_{i1} = (\beta_2 + \beta_3 \texttt{endog}_i + U_{i2})$$

Subject-specific intercept for i is:

$$b_{i0} = (\beta_0 + \beta_1 \texttt{endog}_i + U_{i1})$$

- within-subject covariate is week

    * within-subject residuals Z_{ij} with within-subject variance τ^2
  - **between-subject** variation:
    * subjects differ with respect to their intercepts b_{i0} and slopes b_{i1}
    * between subject intercept covariate: `endog`
    * between subject intercept residual: U_{i1} with variance $G[1,1]$
    * between subject slope covariate: `endog`
    * between subject slope residual: U_{i2} with variance $G[2,2]$
  - **hierarchical linear models** (HLM) formulation:
    – 1st level (within): subject-specific intercept b_{i0} and slope b_{i1} (w.r.t. week)
    – 2nd level (between): models (b_{i0}, b_{i1}) as a function of between-subject covariates

- Variability among subjects reflects natural heterogeneity due to unmeasured factors:
  – children born at different birth weights

– have different growth rates

- In real data: Random effects often explain the observed correlation among repeated measures on a given subject:
  - given random effect, observations independent (often, not always)
  - all observations on same subject share the same unobserved random effects $\boldsymbol{U}_i$, hence correlation in observed responses Y_{ij}
  - correlation arises because we cannot see underlying $\boldsymbol{U}_i$'s (random effects) specific to each subject

- For this example, test yourself:
  – compute $\mathrm{var}(Y_{ij})$ and $\mathrm{var}(Y_{ik})$ at two different time points, t_j and t_k
  – compute $\mathrm{cov}(Y_{ij}, Y_{ik})$
  – compute $\mathrm{corr}(Y_{ij}, Y_{ik})$

- Two key assumptions:

  - Random effects $\boldsymbol{U}_i$ are assumed to arise from a **distribution**:
    - often multivariate normal, but not required
    - $\boldsymbol{U}_i$ independent across subjects

  - Random effects $\boldsymbol{U}_i$ are assumed to be **independent** of covariates X_i, unlike the fixed effects models
    - The "independence of random effects" assumption is much safer when the time-varying covariates are fixed by design or are the same for all subjects
    - e.g., week or week $\times$ group in the depression (Riesby) data

Linear Mixed Models for Longitudinal Data
Details of Model Building and Inference

- Model fitting and testing with random effects works pretty much as with models with other VCC structures (Note 4)
- Statistical inference for β is same with marginal model (Note 5)
- Recall the linear mixed model:

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

where

$$\epsilon_{ij} = \boldsymbol{d}'_{ij}\boldsymbol{U}_i + Z_{ij}$$

so that

$$\text{var}(\boldsymbol{\epsilon}_i) = D_i G D'_i + \tau^2 I_{n_i} = V_i(\boldsymbol{\gamma})$$

So, the variance parameters are $\boldsymbol{\gamma} = (G, \tau^2)$

- Note: with a more general model

$$\epsilon_{ij} = \boldsymbol{d}'_{ij}\boldsymbol{U}_i + W_{ij} + Z_{ij}$$

where

$$\mathrm{var}(\boldsymbol{\epsilon}_i) = D_i G D'_i + \mathrm{var}(W_{ij} + Z_{ij}) = D_i G D'_i + R_i = V_i(\boldsymbol{\gamma}),$$

$\boldsymbol{\gamma}$ will also include parameters for correlation structure of W_{ij}

- Can estimate variance parameters $\boldsymbol{\gamma}$ and fixed effects $\boldsymbol{\beta}$ by ML
- or by ReML:
  - Variance parameters $\boldsymbol{\gamma}$ via ReML
  - **fixed effects** $\boldsymbol{\beta}$ via WLS with fitted $V(\widehat{\boldsymbol{\gamma}})$

Testing fixed effects (β-coefficients)

- **Testing single β-coefficient:**
  Example: (Riesby depression data)
  We wish to test for the effect of endogenous versus non-endogenous diagnosis on depression just at week 0 (baseline).
  - t-statistics in the model output yield tests of individual parameters.

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
endog	1.9882	1.0864	243	1.83	0.0685

  - we obtain a t-statistic of 1.83 with $P\text{-value} = 0.069$
  - Conclusion: with borderline statistical significance that those with endogenous type have a higher level of depression at baseline than those without
  - What is H_0?, H_A?

- **Testing multiple β-coefficients jointly:**

  Example: We wished to test for the **overall effect over time** of endogenous versus non-endogenous diagnosis on depression

  – In the model

$$Y_{ij} = \beta_0 + \beta_1 \texttt{endog}_i + \beta_2 \texttt{week}_{ij} + \beta_3 \texttt{endog}_i * \texttt{week}_{ij} + U_{i1} + U_{i2}\texttt{week}_{ij} + Z_{ij}$$

  we would test

$$H_0 : \beta_1 = \beta_3 = 0 \quad \text{vs.} \quad H_A : \beta_1 \neq 0 \text{ and/or } \beta_3 \neq 0$$

Label	Num DF	Den DF	F Value	Pr > F
F test for endog and endwk	2	243	2.05	0.1304

– Conclusion (**subject-specific interpretation**):
Recall HLM form:

$$Y_{ij} = \underbrace{(\beta_0 + \beta_1 \texttt{endog}_i + U_{i1})}_{b_{i0}} + \underbrace{(\beta_2 + \beta_3 \texttt{endog}_i + U_{i2})}_{b_{i1}} \texttt{week}_{ij} + Z_{ij}$$

* No evidence in the data that the two diagnostic groups differ in their **mean subject-specific intercept** or **mean subject-specific slope** w.r.t. week
* Interpretation in terms of **population** of subject-specific trends: The **average** subject-specific trend does not differ significantly across the two diagnostic groups

– **Important Note:** Do not do LRT for $\boldsymbol{\beta}$ parameters if you are using ReML

– Note for DF computation in Stata:

  Unlike PROC MIXED in SAS, `xtmixed` in Stata does not try to compute a denominator df
  – hence produces Z- and χ^2-test (by `testparm` command) as opposed to t- and F-tests
  – a bit less precise in smaller samples

Testing random effects variance parameters

Example: Test whether subject-specific slope over time varies significantly from subject to subject

- We wish to test

$$H_0 : \text{var}(U_{i2}) = G[2,2] = 0 \quad \text{and} \quad \text{cov}(U_{i1}, U_{i2}) = G[1,2] = 0$$

versus

$$H_A : \text{var}(U_{i2}) = G[2,2] > 0$$

with a possibility that

$$\text{cov}(U_{i1}, U_{i2}) = G[1,2] \neq 0$$

- Slightly awkward situation statistically:

$$\mathrm{var}(U_{i2}) = 0$$

is a **boundary value** (because variances cannot be negative), which suggests a certain one-sided-ness to the problem

However,

$$\mathrm{cov}(U_{i1}, U_{i2}) = 0$$

is not a boundary value, but is a **model constraint** imposed by the fact that $\mathrm{var}(U_{i2}) = 0$

- **Remark**: Fitzmaurice, Laid and Ware (Section 8.5) have some discussion:

"In general, when testing null hypothesis that is on the boundary of the parameter space (e.g., the variance of a random effect equals zero), the usual null distribution for the likelihood ratio test is no

longer a chi-squared distribution with degrees of freedom equal to the difference between the number of parameters in the full and reduced models; instead, the null distribution is a mixture of chi-squared distributions."

- LRT for No random effect vs one random effect
  - The asymptotic null distribution of LRT statistics is a mixture of χ^2_0 and χ^2_1 with equal weights 0.5.
  - The χ^2_0 is the distribution which gives probability one to the value 0.
  - If χ^2_1 is used, p-values would be over-estimated, resulting in accepting the null hypothesis too often.
  - Recall (Note 4): We divide p-value by 2 to solve this issue

- LRT for q random effect vs $q+1$ random effect
  - The asymptotic null distribution of LRT statistics is a mixture of χ^2_q and χ^2_{q+1} with equal weights 0.5.
  - Conservatively, can use χ^2_{q+1}.
  - A more accurate way: use mixture of χ^2_q and χ^2_{q+1} to calculate p-value.
- LRT for q random effect vs $q+k$ random effect
  - The asymptotic null distribution of LRT statistics is a mixture of χ^2 random variables and other types of random variables.
  - Simulation methods are generally used to estimate its distribution.
  - Rule of thumb: Fitzmaurice, Laid and Ware (Section 8.5) recommended to use 0.10 critical level, for a level 0.05 test.

- True test statistic for our example is a mixture of χ^2 with 2 DF and χ^2 with < 2 DF
- To be conservative, use 2 DF
- ReML-based LRT
- From the original model fit, we had

```
-2 Res Log Likelihood          2214.0
```

- Refitting the model in SAS without the random slope and comparing the two likelihoods gives:

```
proc mixed data=riesby;
class id;
model hamd=endog week endwk/ s;
random intercept/ subject=id g ;
run;
```

```
-2 Res Log Likelihood          2282.5
```

R code for LRT:

```
> 1-pchisq(2282.5-2214.0,2)
[1] 1.332268e-15
```

- Conclusion: With $\chi^2 = 2282.5 - 2214.0 = 68.5$ on 2 df, we reject H_0 and conclude that slopes w.r.t. week vary significantly from subject to subject in this population

- A more accurate way to obtain p-value in R (use mixture of χ^2_1 and χ^2_2 distribution):

```
> 1-0.5*pchisq(2282.5-2214.0,2)-0.5*pchisq(2282.5-2214.0,1)
[1] 7.21645e-16
```

- Wald-based 95% CI's for random effects variances work as usual

$$\widehat{\text{var}}(U_{i2}) \pm 1.96 \times \widehat{\text{se}}(\widehat{\text{var}}(U_{i2}))$$

- **Note:** Because of covariance constraint, **do not** use Z-statistics associated with random effects variance parameters as test statistics

Model Evaluation
Does it Fit the Data?

- Questions:
  – Is the mean model correct?
  – Is the variance-covariance model correct
  (recalling that the variance-covariance model is **induced** by the specified random effects model)?

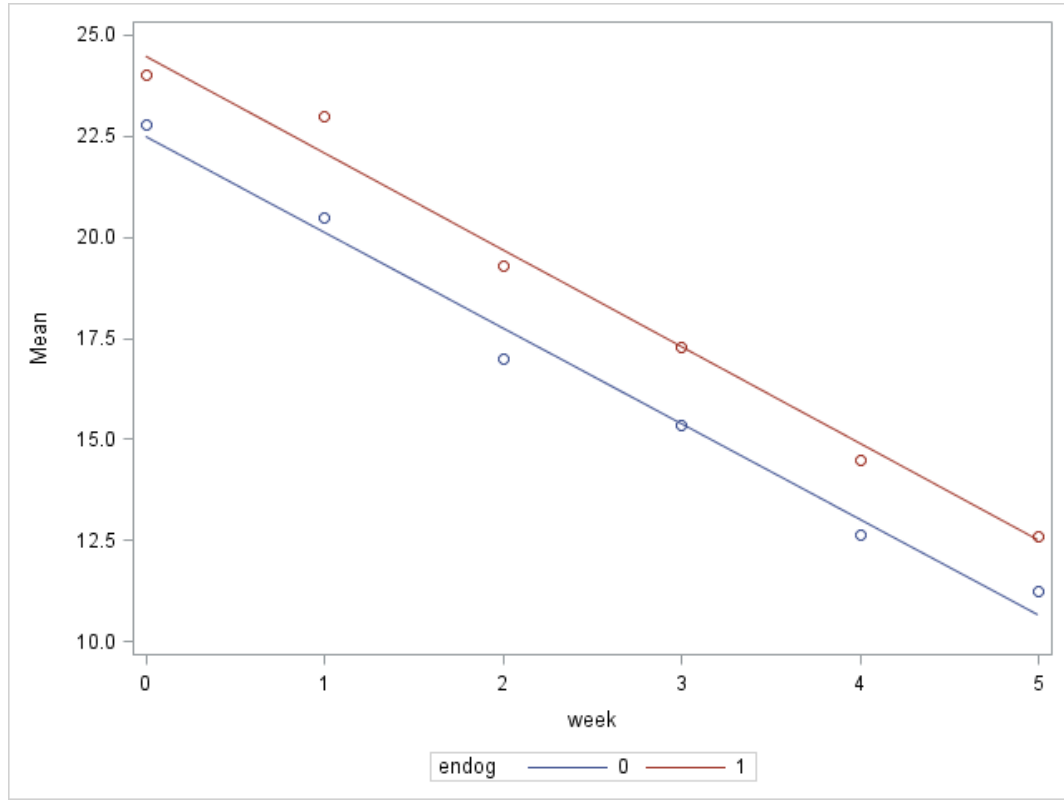
- **Evaluate mean mode**: Fitted **mean functions** (β's) can be compared to observed means to assess model fit.
  – ie, compare mean of observed y_ij vs estimated unconditional expectation $\mathrm{E}(Y_{ij}|X_i) = \boldsymbol{x}'_{ij}\boldsymbol{\beta}$

- SAS code:

```
proc means data=pred mean;
    class endog week;
    var Pred hamd;
    ods output Summary=mean;
run;

proc sgplot data=mean;
    series y=Pred_mean x=week/group=endog;
    scatter y=hamd_mean x=week/group=endog;
run;
```

Fitted response by week and by diagnosis group:



- Conclusion: The model including linear terms for week, a main effect for diagnosis, and their interaction captures the mean trends very well
  – seems interaction can be removed

- How to evaluate **fitted covariance model**?

$$\widehat{\text{var}}(\boldsymbol{\epsilon}_i) = D_i \widehat{G} D_i' + \widehat{\tau}^2 I_{n_i} = V_i(\boldsymbol{\gamma})$$

- **Question**: Is my random effects specification doing a good job of capturing the variance-covariance structure $V_i(\boldsymbol{\gamma})$ of the data?
  - **Cannot** use ACF to assess model fit
    * Variance of subject-specific mean depression at week 0:

$$\widehat{\text{var}}(U_{i1}) = 12.25$$

    * Variance of subject-specific mean depression at week 5:

$$\widehat{\text{var}}(U_{i1} + 5U_{i2}) = 51.3$$

    * Not a **stationary** model (which would require, e.g., constant variance over time)

– Just compare model fitted variances and covariances to observed ones:

  Recall, model fitted covariance matrix:

$$\hat{\text{var}}(Y_i|X_i) = V_i = D_i\hat{G}D_i' + \hat{\tau^2}I_{n_i}$$

  We can compare this with empirical covariance matrix obtained in exploration analysis (like Note 2B)

* Fitted variance parameter estimates

Estimated G Matrix

Row	Effect	id	Col1	Col2
1	Intercept	101	12.2506	-1.5150
2	week	101	-1.5150	2.1719

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z
UN(1,1)	id	12.2506	3.5158	3.48	0.0002
UN(2,1)	id	-1.5150	1.0640	-1.42	0.1545
UN(2,2)	id	2.1719	0.5406	4.02	<.0001
Residual		12.2109	1.1180	10.92	<.0001

* Now compute the fitted (**model-based**) covariance matrix using the estimated $\widehat{G}$ and the estimated $\widehat{\tau}^2$.

* R code:

```
> G=matrix(c(12.2506,-1.5150,-1.5150,2.1719),2,2)
> sigma2=12.2109
> R=diag(6)*sigma2
> D=matrix(c(1,1,1,1,1,1,0,1,2,3,4,5),6,2)
> D
     [,1] [,2]
[1,]    1    0
[2,]    1    1
[3,]    1    2
[4,]    1    3
[5,]    1    4
[6,]    1    5
> G
         [,1]    [,2]
[1,] 12.2506 -1.5150
[2,] -1.5150  2.1719
```

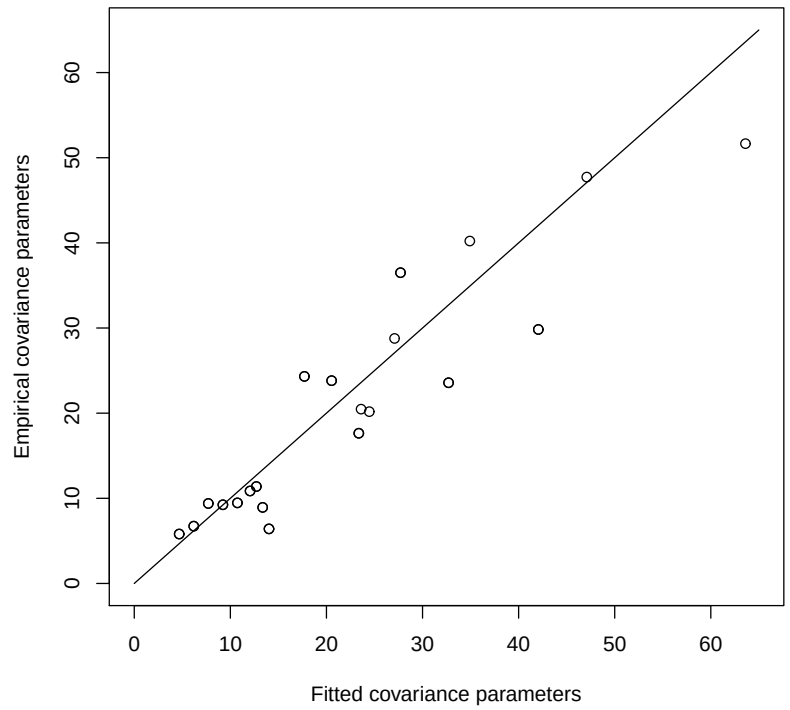
```
> R
         [,1]    [,2]    [,3]    [,4]    [,5]    [,6]
[1,] 12.2109  0.0000  0.0000  0.0000  0.0000  0.0000
[2,]  0.0000 12.2109  0.0000  0.0000  0.0000  0.0000
[3,]  0.0000  0.0000 12.2109  0.0000  0.0000  0.0000
[4,]  0.0000  0.0000  0.0000 12.2109  0.0000  0.0000
[5,]  0.0000  0.0000  0.0000  0.0000 12.2109  0.0000
[6,]  0.0000  0.0000  0.0000  0.0000  0.0000 12.2109
> V=D%*%G%*%t(D)+R
> V
         [,1]    [,2]    [,3]    [,4]    [,5]    [,6]
[1,] 24.4615 10.7356  9.2206  7.7056  6.1906  4.6756
[2,] 10.7356 23.6034 12.0494 12.7063 13.3632 14.0201
[3,]  9.2206 12.0494 27.0891 17.7070 20.5358 23.3646
[4,]  7.7056 12.7063 17.7070 34.9186 27.7084 32.7091
[5,]  6.1906 13.3632 20.5358 27.7084 47.0919 42.0536
[6,]  4.6756 14.0201 23.3646 32.7091 42.0536 63.6090
```

* Compute observed (**empirical**) covariance, removing effects of week and diagnosis:

```
> # Compute residuals, removing effects of week and diagnosis
> # treating week as categorical
> fit<-lm(hamd~endog*factor(week),data=data)
> data$hamdrs[!is.na(data$hamd)]=resid(fit)
>
> # into wide format
> wide<-reshape(data[,c(1,3,5)],v.names="hamdrs",idvar="id",
+ timevar="week",direction="wide")
>
> # covariance matrix
> Emp.V<-cov(wide[,-1],use="pairwise.complete.obs")
> Emp.V
          hamdrs.0  hamdrs.1  hamdrs.2 hamdrs.3  hamdrs.4  hamdrs.5
hamdrs.0 20.178571  9.463455  9.243478  9.39348  6.730263  5.805616
hamdrs.1  9.463455 20.471635 10.860362 11.38152  8.932270  6.410207
hamdrs.2  9.243478 10.860362 28.777027 24.31726 23.817108 17.641202
hamdrs.3  9.393480 11.381521 24.317259 40.21522 36.499201 23.573087
hamdrs.4  6.730263  8.932270 23.817108 36.49920 47.730616 29.828953
hamdrs.5  5.805616  6.410207 17.641202 23.57309 29.828953 51.652896
>
```

```
> ## plot
> plot(c(V),c(Emp.V),xlim=c(0,65),ylim=c(0,65),
+ xlab="Fitted covariance parameters",
+ ylab="Empirical covariance parameters")
> lines(c(0,65),c(0,65))
```

* Plot the empirical variances and covariance (21 of them) versus the model-based variance and covariances



* Conclusion:
  – While this is not a formal check, the general pattern of empirical variance-covariances is well-captured by the model.
  – may be room for improvement (eg, adding quadratic random effects)

Parameterization of Random Effects

- Recall that in the Riesby data, we estimated the variance of the random intercept to be

$$\widehat{\text{var}}(U_{i1}) = 12.25$$

- However, if we had put the intercept at week 5 (ie, subtract 5 from week before fitting model), we would have had

$$\widehat{\text{var}}(U^*_{i1}) = 51.40$$

- So, think carefully about how to center variable to get the interpretation that you would like
- One option: Place the intercept in the middle of the time sequence by **centering the time variable**

- **Example:** Refit the model to Riesby depression data, centering week at 2.5 weeks:

$$\begin{aligned} Y_{ij} &= \beta_0 + \beta_1 \texttt{endog}_i + \beta_2(\texttt{week}_{ij} - 2.5) + \beta_3 \texttt{endog}_i * (\texttt{week}_{ij} - 2.5) \\ &\quad + U_{i1} + U_{i2}(\texttt{week}_{ij} - 2.5) + Z_{ij} \end{aligned}$$

- SAS code:

```
proc mixed data=riesby dfbw covtest ;
     class id;
     model hamd = endog wkcen endwkcen / s;
     random intercept wkcen / subject=id type=un g ;
     run;
```

- Results:

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z
UN(1,1)	id	18.2503	3.6294	5.03	<.0001
UN(2,1)	id	3.9149	1.1081	3.53	0.0004
UN(2,2)	id	2.1719	0.5406	4.02	<.0001
Residual		12.2109	1.1180	10.92	<.0001

Fit Statistics

-2 Res Log Likelihood	2214.0
AIC (smaller is better)	2222.0

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	16.5618	0.8379	64	19.77	<.0001
endog	1.9212	1.1210	64	1.71	0.0914
wkcen	-2.3657	0.3170	307	-7.46	<.0001
endwkcen	-0.02680	0.4264	307	-0.06	0.9499

- The mean (for non-endogenous subjects) and variance of the within subject slope with respect to week are the same:

$$\widehat{\beta}_2 = -2.37 \quad \text{and} \quad \widehat{\text{var}}(U_{i2}) = 2.17$$

- But the mean (for non-endogenous subjects) and variance of the within subject **intercept** now refer to the level of depression at 2.5 weeks:

$$\widehat{\beta}_0 = 16.56 \quad \text{and} \quad \widehat{\text{var}}(U_{i1}) = 18.25$$

- Intercept and slope are now positively correlated
- Generally, centered estimation is more **numerically stable** in software computation
- Another reason to center time: to add **Higher order terms**
  - We have seen that the population mean trend over time is fairly **linear** in both diagnostic groups
  - But, some subjects could experience more curvi-linear trajectories
  - We can test this by adding a quadratic term to random effects

Higher Order Random Effects

- Suppose we would like to include a random quadratic term for each person in the Riesby data set
- SAS code:

```
proc mixed data=riesby dfbw covtest ;
     class id;
     model hamd = endog wkcen endwkcen wksqr / s;
     random intercept wkcen wksqr/ subject=id type=un g ;
     run;
```

  Note: We generally add an effect to $\boldsymbol{x}_{ij}$ whenever we add it to $\boldsymbol{d}_{ij}$

- Results:

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z
UN(1,1)	id	23.6860	4.9648	4.77	<.0001
UN(2,1)	id	3.7533	1.2212	3.07	0.0021

UN(2,2)	id	2.2132	0.5339	4.15	<.0001
UN(3,1)	id	-1.2074	0.5565	-2.17	0.0300
UN(3,2)	id	0.03432	0.1550	0.22	0.8247
UN(3,3)	id	0.2026	0.09576	2.12	0.0172
Residual		10.5008	1.1039	9.51	<.0001

Fit Statistics

-2 Res Log Likelihood	2206.2
AIC (smaller is better)	2220.2

- To test for the presence of quadratic individual trajectories, we test

$$H_0 : \text{var}(U_{i3}) = \text{cov}(U_{i3}, U_{i1}) = \text{cov}(U_{i3}, U_{i2}) = 0$$

versus

$$H_A : \text{var}(U_{i3}) > 0$$

This is (conservatively) 3 DF.

- Use LRT:
  Need to fit model under null hypothesis (with same fixed effects) in SAS:

```
proc mixed data=riesby dfbw covtest ;
     class id;
     model hamd = endog wkcen endwkcen wksqr / s;
     random intercept wkcen/ subject=id type=un g ;
     run;
```

```
                  -2 Res Log Likelihood          2216.9
```

  Using LRT based on ReML, we obtain

$$X^2 = 2216.9 - 2206.2 = 7.8$$

- LRT in R:

```
> 1-pchisq(2216.9 - 2206.2,3)
[1] 0.01346379
```

  which gives $P\text{-value} = 0.01$

- A more accurate LRT in R (mixture of DF):

```
> 1-0.5*pchisq(2216.9 - 2206.2,3)-0.5*pchisq(2216.9 - 2206.2,2)
[1] 0.009105968
```

- By LRT, we conclude that there is significant deviation from linear trends among the subjects in this population

- Interpretation:

$$\widehat{\text{var}}(U_{i1}) = 23.68$$

is the **variance of the within-subject intercept** at the **centered** value of week (at 2.5 weeks),

$$\widehat{\text{var}}(U_{i2}) = 2.21$$

is the **variance of the within-subject slope** at the **centered** value of week (at 2.5 weeks).

- **Note:** This slope estimate is not that much different than the one estimated without the quadratic term

$$\widehat{\text{var}}(U_{i3}) = 0.2026$$

is the **variance of the within-subject quadratic coefficient**.

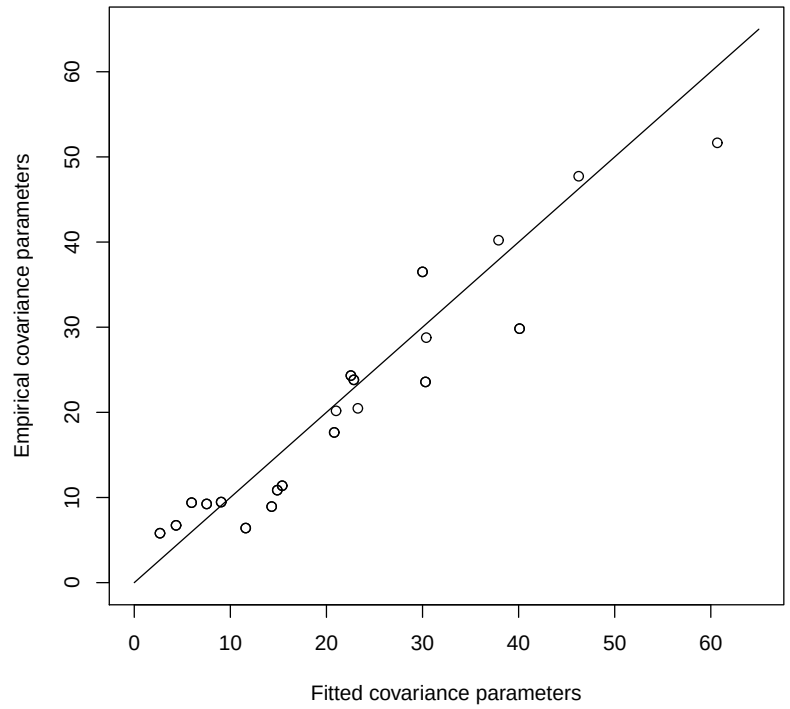
- **Remark:** To test quadratic random effect using LRT
  In most cases, you would add a quadratic term to the fixed effects:
  - If use ReML: Need to include the fixed effects quadratic term in **both model fits**
  - Or, use ML, and include the fixed effect of quadratic term in your test

- Finally, we might compute the new fitted covariance matrix and see if is better reflects the empirical covariance matrix
- R code:

```
> G=matrix(c(23.686 , 3.7533 , -1.2074, 3.7533, 2.2132 , 0.03432,
+ -1.2074 , 0.03432 , 0.2026),3,3)
> sigma2=10.5008
> R=diag(6)*sigma2
> D=matrix(c(rep(1,6),seq(-2.5,2.5,1),seq(-2.5,2.5,1)^2),6,3)
> D
     [,1] [,2] [,3]
[1,]    1 -2.5 6.25
[2,]    1 -1.5 2.25
[3,]    1 -0.5 0.25
[4,]    1  0.5 0.25
[5,]    1  1.5 2.25
[6,]    1  2.5 6.25
> G
        [,1]    [,2]     [,3]
[1,] 23.6860 3.75330 -1.20740
[2,]  3.7533 2.21320  0.03432
[3,] -1.2074 0.03432  0.20260
```

```
> R
        [,1]    [,2]    [,3]    [,4]    [,5]    [,6]
[1,] 10.5008  0.0000  0.0000  0.0000  0.0000  0.0000
[2,]  0.0000 10.5008  0.0000  0.0000  0.0000  0.0000
[3,]  0.0000  0.0000 10.5008  0.0000  0.0000  0.0000
[4,]  0.0000  0.0000  0.0000 10.5008  0.0000  0.0000
[5,]  0.0000  0.0000  0.0000  0.0000 10.5008  0.0000
[6,]  0.0000  0.0000  0.0000  0.0000  0.0000 10.5008
> V=D%*%G%*%t(D)+R
> V
          [,1]      [,2]      [,3]      [,4]      [,5]      [,6]
[1,] 21.001863  9.043662  7.532362  5.967162  4.348062  2.675062
[2,]  9.043662 23.267302 14.883282 15.394002 14.298663 11.597262
[3,]  7.532362 14.883282 30.387182 22.541662 22.849122 20.808762
[4,]  5.967163 15.394002 22.541663 37.910942 29.999442 30.309563
[5,]  4.348063 14.298663 22.849123 29.999443 46.250422 40.099662
[6,]  2.675062 11.597262 20.808762 30.309562 40.099662 60.679862

> ## plot
> plot(c(V),c(Emp.V),xlim=c(0,65),ylim=c(0,65),
+ xlab="Fitted covariance parameters",ylab="Empirical covariance parameters")
> lines(c(0,65),c(0,65))
```



Seems a little bit better than only including the linear random effect term

- Some good reasons to center covariates $\boldsymbol{x}_{ij}$ and $\boldsymbol{d}_{ij}$:
  - Makes the intercept at the central time point rather than at the beginning
  - Facilitates adding and interpreting interaction and/or squared terms
  - Controls high correlation among random effects: helps numerically and with interpretation
  - Avoids very large random or fixed effects covariates, which can threaten numeric stability
  - If you do not like the interpretation, you can always change the parameterization