Example 3a: From Between-Person to Within-Person Models for Longitudinal Data
(complete syntax, data, and output available for SAS, STATA, and R electronically)

The models for this example come from Hoffman (2015) chapter 3 example 3a. We will be examining the extent to which a learning achievement outcome can be predicted from group (control as the reference vs. treatment) and time (pre-test as the reference vs. post-test) in a sample of 50 children. For an example results section, please see the end of chapter 3.

Here is the original wide-format data (one row per person), standard for RM ANOVA:

	PersonID	group	outcome1	outcome2
1	1	1	55.957284	57.806319
2	2	1	53.751908	58.943056
3	3	1	53.255502	59.149058
4	4	1	44.796029	44.561578
5	5	1	47.155838	50.087386

Here is what the dataset will become after transforming from wide-format to long-format (one row per occasion), standard for MLM:

	PersonID	occasion	group	outcome
1	1	1	1	55.957284
2	1	2	1	57.806319
3	2	1	1	53.751908
4	2	2	1	58.943056
5	3	1	1	53.255502
6	3	2	1	59.149058
7	4	1	1	44.796029
8	4	2	1	44.561578
9	5	1	1	47.155838
10	5	2	1	50.087386

STATA Syntax for Data Import and Manipulation:

```
// Define working directory for files
   cd "C:\Dropbox\24_PSQF6271\PSQF6271_Example3a"

// Import chapter 3a two-occasion wide-format data from excel
   clear // clear memory in case a dataset is already open
   import excel "Example3a.xlsx", firstrow case(preserve) clear

// Stack into long-format data
// List time-varying variables first, i(level2ID) j(newtimeID)
   reshape long outcome, i(PersonID) j(occasion)
   label variable occasion "occasion: 1=pre-test, 2=post-test"
   label variable outcome "outcome: Learning Outcome"

// Center predictors for analysis
   gen time  = time - 1
   gen treat = group - 1
   label variable time  "time: 0=pre-test, 1=post-test"
   label variable treat "treat: 0=control, 1=treatment"

// Sort by person and occasion
   sort PersonID occasion
```

R Syntax for Data Import and Manipulation (after loading packages *readxl*, *TeachingDemos*, *psych*, *nlme*, and *multcomp*):

```
# Define variables for working directory and data name
setwd("C:/Dropbox/24_PSQF6271/PSQF6271_Example3a/")

# Import chapter 3a two-occasion wide-format data from excel
Example3a_wide = read_excel(path="Example3a.xlsx")
# Convert to data frame as data frame without labels to use for analysis
Example3a_wide = as.data.frame(Example3a_wide)
```

```
# Stack into long-format (one row per occasion per person)
Example3a = reshape(Example3a_wide, direction="long",
                    v.names="outcome", idvar="PersonID", timevar="occasion",
                    varying=c("outcome1","outcome2"), times=c(1,2))

# Center predictors for analysis
Example3a$time  = Example3a$occasion-1
Example3a$treat = Example3a$group-1
# Labels as comments only
# occasion: 1=pre-test, 2=post-test
# outcome: Learning Outcome
# time: 0=pre-test, 1=post-test
# treat: 0=control, 1=treatment

# Sort by person and occasion (needed for correct R)
Example3a = Example3a[order(Example3a$PersonID, Example3a$occasion), ]
```

STATA Descriptive Statistics using `tabulate` (→ means in Lecture 3 slide 15):

```
display "Chapter 3a Example: Means by group and occasion for learning outcome"
tabulate group occasion, summarize(outcome)
```

```
Means, Standard Deviations and Frequencies of outcome: Learning Outcome

           |      occasion:
           |     1=pre-test,
           |     2=post-test
     group |         1          2 |     Total
-----------+----------------------+----------
         1 | 49.076798  54.899163 |  51.98798
           | 5.6852881  5.6282644 | 6.3241453
           |        25         25 |        50
-----------+----------------------+----------
         2 |  50.75874  58.623631 | 54.691186
           | 4.5354041   4.932377 | 6.1457869
           |        25         25 |        50
-----------+----------------------+----------
     Total | 49.917769  56.761397 | 53.339583
           | 5.1602464  5.5650746 | 6.3510064
           |        50         50 |       100
```

R Descriptive Statistics using `describeBy` from `psych` package:

```
print("Chapter 3a Example: Means by group and occasion for learning outcome")
describeBy(x=Example3a$outcome, digits=2,
           group=list(group=Example3a$group, occasion=Example3a$occasion))
```

```
Descriptive statistics by group
group: 1
occasion: 1
   vars  n  mean   sd median trimmed  mad   min   max range skew kurtosis   se
X1    1 25 49.08 5.69  48.22   48.95 5.42 37.53 59.55 22.02 0.25     -0.8 1.14
------------------------------------------------------------------------------------------------
group: 2
occasion: 1
   vars  n  mean   sd median trimmed  mad   min   max range skew kurtosis   se
X1    1 25 50.76 4.54  51.43   50.61 4.43 40.53 62.13  21.6 0.26     0.24 0.91
------------------------------------------------------------------------------------------------
group: 1
occasion: 2
   vars  n mean   sd median trimmed  mad   min   max range skew kurtosis   se
X1    1 25 54.9 5.63  54.12   54.79 5.46 44.56 67.11 22.54 0.18    -0.72 1.13
------------------------------------------------------------------------------------------------
group: 2
occasion: 2
   vars  n  mean   sd median trimmed  mad   min   max range  skew kurtosis   se
X1    1 25 58.62 4.93  58.54   58.75 4.24 47.43 68.62 21.19 -0.16    -0.25 0.99
```

Equation 3.1: Between-Person (e-only) Empty Model $y_{ti} = \beta_0 + e_{ti}$

STATA Syntax and Output using `mixed`:

STATA MIXED includes a random intercept by default, removed by **`noconstant`** after the | |. The pattern of variances and covariances over time (occasion = time ID variable) is controlled by **`residuals`**.

```
display "Eq 3a.1: Between-Person (e-only) Empty Model via STATA MIXED"
mixed outcome , || PersonID: , ///
      noconstant reml residuals(independent,t(occasion)) ///
      nolog dfmethod(residual) dftable(pvalue)  // Use dftable(default) to get CIs
```

```
Mixed-effects REML regression                   Number of obs     =        100
DF method: Residual                             DF:           min =      99.00
                                                              avg =      99.00
                                                              max =      99.00
                                                F(0,     99.00)   =          .
Log restricted-likelihood = -325.79022          Prob > F          =          .
------------------------------------------------------------------------------
     outcome |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _cons |   53.33958   .6351006    83.99   0.000     52.07941    54.59976  Fixed intercept
------------------------------------------------------------------------------
------------------------------------------------------------------------------
  Random-effects Parameters  |   Estimate   Std. Err.     [95% Conf. Interval]
-----------------------------+------------------------------------------------
             var(Residual) |   40.33528    5.70427      30.57082    53.21857  TOTAL variance
------------------------------------------------------------------------------
```

```
display "-2LL = " e(ll)*-2   // Print -2LL for model -- used in homework
-2LL = 651.58043
```

R Syntax and Output using `gls` from `nlme` Package:

```
print("Eq 3.1: Between-Person (e-only) Empty Model via R GLS")
empty_BP = gls(data=Example3a, method="REML", model=outcome~1,
               correlation=NULL) # Diagonal R matrix (no correlation)

print("Show results with -2LL, total original variance, and 95% CI")
-2*logLik(empty_BP); summary(empty_BP)
```

```
'log Lik.' 651.58043 (df=2) → -2LL for model -- used in homework

       AIC       BIC     logLik → LL for model instead (along with AIC and BIC)
  655.58043 660.77067 -325.79022

Coefficients:
                Value  Std.Error   t-value p-value
(Intercept) 53.339583 0.63510064 83.986032       0  → Fixed intercept

Residual standard error: 6.3510064 → Is total SD instead of total variance

summary(empty_BP)$sigma^2; confit(empty_BP, level=.95)

[1] 40.335282 → Total variance

                2.5 %    97.5 %
(Intercept) 52.094809 54.584357  → 95% CI for fixed intercept
```

R Matrix from SAS (would not print in STATA or R for an e-only model)

Estimated R Matrix for PersonID 1		
Row	**Col1**	**Col2**
1	40.3353	
2		40.3353

The **R matrix** shows the model-implied pattern of total variance and covariance over time. A BP (e-only) model implies constant variance over time with no covariance between occasions ("independent" in STATA or "diagonal" elsewhere). But this is a testable hypothesis!

Equation 3.2: Within-Person (U+e) Empty Model $y_{ti} = \beta_0 + U_{0i} + e_{ti}$

STATA Syntax and Output using `mixed`:

```
display "Eq 3a.2: Within-Person (U+e) Empty Model via STATA MIXED"
mixed outcome , || PersonID: , ///
      noconstant reml residuals(exchangeable,t(occasion)) ///
      nolog dfmethod(repeated) dftable(pvalue)
```

```
Mixed-effects REML regression                   Number of obs     =        100
Group variable: PersonID                        Number of groups  =         50
                                                Obs per group:
                                                              min =          2
                                                              avg =        2.0
                                                              max =          2
DF method: Repeated                             DF:           min =      49.00
                                                              avg =      49.00
                                                              max =      49.00
                                                F(0,       .  )     =      .
Log restricted-likelihood = -323.40474          Prob > F          =      .
------------------------------------------------------------------------------
     outcome |      Coef.   Std. Err.           DF       t    P>|t|
-------------+----------------------------------------------------------------
       _cons |   53.33958   .7260282          49.0    73.47   0.000  Fixed intercept
------------------------------------------------------------------------------
------------------------------------------------------------------------------
  Random-effects Parameters  |   Estimate   Std. Err.  [95% Conf. Interval]
-----------------------------+------------------------------------------------
PersonID:            (empty) |
-----------------------------+------------------------------------------------
Residual: Exchangeable       |
                     var(e) |   40.45905   6.025635   30.21654    54.17346  Var(U0)+Var(e) = Total Var
                     cov(e) |   12.25265   6.025634   .4426208    24.06267  Var(U0)
------------------------------------------------------------------------------
LR test vs. linear model: chi2(1) = 4.77                  Prob > chi2 = 0.0289
```

The **`chi(1)`** test above is a likelihood ratio test (LRT) comparing this model to the single-level model (without a random intercept, as **`linear model`**) using a chi-square (χ^2) distribution with DF=1 for Var(U0).

```
display "-2LL = " e(ll)*-2      // Print -2LL for model -- used in homework
-2LL = 646.80947
```

```
estat wcorrelation, covariance  // R matrix
```

```
Covariances for PersonID = 1:
         obs |      1       2
-------------+----------------
           1 | 40.459
           2 | 12.253  40.459
```

The **R matrix** for total variance and covariance over time now has a "compound symmetry" or "exchangeable" structure: equal variance on the diagonal, "CS" covariance on the off-diagonal as Var(U0). So Var(e) = Var(total) – Var(U0) = 40.459 – 12.253 = 28.206

```
estat wcorrelation              // RCORR matrix
```

```
Correlations:
         obs |      1       2
-------------+----------------
           1 |  1.000
           2 |  0.303   1.000
```

The **RCORR matrix** gives the ICC in the off-diagonal:

$$\text{ICC} = \frac{\tau^2_{U_0}}{\tau^2_{U_0} + \sigma^2_e} = \frac{12.253}{40.459} = .\mathbf{303}$$

The LRT above is a significance test of the intraclass correlation (ICC), which in turn provides an effect size for the amount of constant dependency attributed to person mean differences in the outcome.

R Syntax and Output using `gls` from `nlme` Package:

```
print("Eq 3.2: Within-Person (U+e) Empty Model via R GLS (CS correlation)")
empty_WP = gls(data=Example3a, method="REML", model=outcome~1,
               correlation=corCompSymm(form=~1|PersonID))
print("Show results with -2LL"); -2*logLik(empty_WP); summary(empty_WP)
```

```
'log Lik.' 646.80947 (df=3) → -2LL for model -- used in homework

       AIC       BIC     logLik
  652.80947 660.59483 -323.40474

Correlation Structure: Compound symmetry
Parameter estimate(s):
       Rho
0.30284193 → ICC for correlation across occasions

Coefficients:
               Value  Std.Error   t-value p-value
(Intercept) 53.339583 0.72602868 73.467598       0  → Fixed intercept
```

The **RCORR matrix** gives the ICC in the off-diagonal:

$$\text{ICC} = \frac{\tau^2_{U_0}}{\tau^2_{U_0} + \sigma^2_e} = \frac{12.253}{40.459} = .\mathbf{303}$$

```
print("Show R and RCORR matrices")
getVarCov(empty_WP); corMatrix(empty_WP$modelStruct$corStruct)[[5]]
```

```
Marginal variance covariance matrix
       [,1]   [,2]
[1,] 40.459 12.253
[2,] 12.253 40.459

           [,1]       [,2]
[1,] 1.00000000 0.30284193
[2,] 0.30284193 1.00000000
```

The **R matrix** for total variance and covariance over time now has a "compound symmetry" or "exchangeable" structure: equal variance on the diagonal, "CS" covariance on the off-diagonal as Var(U0).
So Var(e) = Var(total) – Var(U0) = 40.459 – 12.253 = 28.206

The **RCORR matrix** gives the ICC in the off-diagonal (as **rho** above)

```
print("Show likelihood ratio test comparing empty means variance model fit")
anova(empty_WP, empty_BP) # Order of "more", "fewer"
```

```
         Model df       AIC       BIC     logLik   Test   L.Ratio p-value
empty_WP     1  3 652.80947 660.59483 -323.40474
empty_BP     2  2 655.58043 660.77067 -325.79022 1 vs 2 4.7709583  0.0289
```

The LRT above is a significance test of the intraclass correlation (ICC), which in turn provides an effect size for the amount of constant dependency attributed to person mean differences in the outcome.

Equation 3.7 (top): Between-Person (e-only) Conditional Model

$$y_{ti} = \beta_0 + \beta_1(\text{Time}_{ti}) + \beta_2(\text{Treat}_i) + \beta_3(\text{Time}_{ti})(\text{Treat}_i) + e_{ti}$$

Top: ALL fixed effects contribute to predicted outcomes

Bottom: ONLY fixed effects containing the target slope contribute its prediction

Control Group Mean at Pre-Test: $\hat{y}_{ti} = \beta_0 + \beta_1(0) + \beta_2(0) + \beta_3(0)(0)$

Control Group Mean at Post-Test: $\hat{y}_{ti} = \beta_0 + \beta_1(1) + \beta_2(0) + \beta_3(0)(0)$

Treatment Group Mean at Pre-Test: $\hat{y}_{ti} = \beta_0 + \beta_1(0) + \beta_2(1) + \beta_3(0)(0)$

Treatment Group Mean at Post-Test: $\hat{y}_{ti} = \beta_0 + \beta_1(1) + \beta_2(1) + \beta_3(1)(1)$

Time Slope for Control Group: $\beta_1(\text{Time}_{ti}) + \beta_3(\text{Time}_{ti})(\text{Treat}_i) \rightarrow [\beta_1 + \beta_3(0)](\text{Time}_{ti})$

Time Slope for Treatment Group: $\beta_1(\text{Time}_{ti}) + \beta_3(\text{Time}_{ti})(\text{Treat}_i) \rightarrow [\beta_1 + \beta_3(1)](\text{Time}_{ti})$

Group Slope at Pre-Test: $\beta_2(\text{Treat}_i) + \beta_3(\text{Time}_{ti})(\text{Treat}_i) \rightarrow [\beta_2 + \beta_3(0)](\text{Treat}_i)$

Group Slope at Post-Test: $\beta_2(\text{Treat}_i) + \beta_3(\text{Time}_{ti})(\text{Treat}_i) \rightarrow [\beta_2 + \beta_3(1)](\text{Treat}_i)$

STATA Syntax using `mixed`:

```
display "Eq 3.7 (top): Between-Person (e-only) Conditional (Predictor) Model"
display "Not using i., manually dummy coding group and time"
mixed outcome c.time c.treat c.time#c.treat,  || PersonID: ,  ///
      noconstant reml  residuals(independent,t(occasion))     ///
      nolog dfmethod(residual) dftable(pvalue)
display "-2LL = " e(ll)*-2      // Print -2LL for model
// DF=3 Joint F-test for Model R2
   test (c.time=0)(c.treat=0)(c.time#c.treat=0), small

// Control Group Mean at Pre-Test
   lincom _cons*1 + c.time*0 + c.treat*0 + c.time#c.treat*0, small
// Control Group Mean at Post-Test
   lincom _cons*1 + c.time*1 + c.treat*0 + c.time#c.treat*0, small
// Treatment Group Mean at Pre-Test
   lincom _cons*1 + c.time*0 + c.treat*1 + c.time#c.treat*0, small
// Treatment Group Mean at Post-Test
   lincom _cons*1 + c.time*1 + c.treat*1 + c.time#c.treat*1, small
// Time Slope for Control Group
   lincom c.time*1 + c.time#c.treat*0, small
// Time Slope for Treatment Group
   lincom c.time*1 + c.time#c.treat*1, small
// Group Slope at Pre-Test
   lincom c.treat*1 + c.time#c.treat*0, small
// Group Slope at Post-Test
   lincom c.treat*1 + c.time#c.treat*1, small
// Time*Group Interaction Slope
   lincom c.time#c.treat*1, small
```

R Syntax and Output using `gls` from `nlme` Package:

```
print("Eq 3.7 (top): Between-Person (e-only) Conditional (Predictor) Model")
print("Not using factor variables, manually dummy coding group and time")
cond_BP = gls(data=Example3a, method="REML", model=outcome~1+time+treat+time:treat,
              correlation=NULL)
print("Show results with -2LL and total leftover variance")
-2*logLik(cond_BP); summary(cond_BP); summary(cond_BP)$sigma^2
```

'log Lik.' 602.50705 (df=5) → -2LL for model -- used in homework

```
       AIC       BIC     logLik
  612.50705 625.32879 -301.25352

Coefficients:
               Value Std.Error   t-value p-value  fixed effect
(Intercept) 49.076798 1.0435425 47.029036  0.0000  beta0
time         5.822365 1.4757920  3.945248  0.0002  beta1
treat        1.681942 1.4757920  1.139688  0.2573  beta2
time:treat   2.042527 2.0870850  0.978650  0.3302  beta3
```

[1] 27.224523 → total variance *leftover* after predictors (diagonal of R matrix, 0 off-diagonal)

```
print("Get and show joint test of model R2 using correct denominator DF")
modelF_BP=glht(model=cond_BP, df=48, linfct=c("time=0","treat=0","time:treat=0"))
BP=summary(modelF_BP, test=Ftest()); BP
```

```
Warning message:
In test(object):'df.residual' is not available for 'model' a Chisq test is performed
instead of the requested F test.

Global Test:
   Chisq DF        Pr(>Chisq)
1 50.676  3 0.000000000057339
```

```
print("Get and show hidden results for F, dfnum, dfden, and p-value")
BP$test$fstat; BP$test$df; BP$df

[1,] 16.892107 → F-value = Chisq / #slopes
[1] 3          → numerator DF = # slopes
[1] 48         → denominator DF based on DF=2 for intercept and treat

pf(BP$test$fstat,df1=BP$test$df,df2=BP$df,lower.tail=FALSE)
                [,1]
[1,] 0.0000001262522  → p-value for F-test
```

Equation 3.7 (top) Repeated: $y_{ti} = \beta_0 + \beta_1(\text{Time}_{ti}) + \beta_2(\text{Treat}_i) + \beta_3(\text{Time}_{ti})(\text{Treat}_i) + e_{ti}$

```
print("Get linear combination estimates, add correct denominator DF and 95% CIs")
estimates_BP = glht(model=cond_BP, df=48, linfct=rbind(
       "Control Group Mean at Pre-Test"     =c(1,0,0,0), # Multiply ordered fixed effects
       "Control Group Mean at Post-Test"    =c(1,1,0,0),
       "Treatment Group Mean at Pre-Test"   =c(1,0,1,0),
       "Treatment Group Mean at Post-Test"  =c(1,1,1,1),
       "Time Slope for Control Group"       =c(0,1,0,0),
       "Time Slope for Treatment Group"     =c(0,1,0,1),
       "Group Slope at Pre-Test"            =c(0,0,1,0),
       "Group Slope at Post-Test"           =c(0,0,1,1),
       "Time*Group Interaction Slope"       =c(0,0,0,1)))
summary(estimates_BP, test=adjusted("none"))

Linear Hypotheses:
                                       Estimate Std. Error t value    Pr(>|t|)  b=beta
Control Group Mean at Pre-Test == 0     49.0768     1.0435 47.0290   < 2.2e-16  b0
Control Group Mean at Post-Test == 0    54.8992     1.0435 52.6085   < 2.2e-16  b0+b1
Treatment Group Mean at Pre-Test == 0   50.7587     1.0435 48.6408   < 2.2e-16  b0+b2
Treatment Group Mean at Post-Test == 0  58.6236     1.0435 56.1775   < 2.2e-16  b0+b1+b2+b3

Time Slope for Control Group == 0        5.8224     1.4758  3.9452   0.0002591  b1
Time Slope for Treatment Group == 0      7.8649     1.4758  5.3293 0.000002604  b1+b3
Group Slope at Pre-Test == 0             1.6819     1.4758  1.1397   0.2600719  b2
Group Slope at Post-Test == 0            3.7245     1.4758  2.5237   0.0149764  b2+b3
Time*Group Interaction Slope == 0        2.0425     2.0871  0.9787   0.3326596  b3
(Adjusted p values reported -- none method)
```

These results assume independent occasions… what happens if that's not the case?

Equation 3.7 (bottom): Within-Person (U+e) Conditional Model

$$y_{ti} = \beta_0 + \beta_1(\text{Time}_{ti}) + \beta_2(\text{Treat}_i) + \beta_3(\text{Time}_{ti})(\text{Treat}_i) + U_{0i} + e_{ti}$$

STATA Syntax using `mixed`:

```
display "Eq 3.7 (bottom): Within-Person (U+e) Conditional (Predictor) Model"
display "Not using i., manually dummy coding group and time"
mixed outcome c.time c.treat c.time#c.treat,  || PersonID: ,  ///
      noconstant reml  residuals(exchangeable,t(occasion))    ///
      nolog dfmethod(residual) dftable(pvalue)
display "-2LL = " e(ll)*-2      // Print -2LL for model
estat wcorrelation, covariance  // R matrix
estat wcorrelation              // RCORR matrix
// DF=3 Joint F-test for Model R2
   test (c.time=0)(c.treat=0)(c.time#c.treat=0), small

// Control Group Mean at Pre-Test
   lincom _cons*1 + c.time*0 + c.treat*0 + c.time#c.treat*0, small
// Control Group Mean at Post-Test
   lincom _cons*1 + c.time*1 + c.treat*0 + c.time#c.treat*0, small
```

```
// Treatment Group Mean at Pre-Test
   lincom _cons*1 + c.time*0 + c.treat*1 + c.time#c.treat*0, small
// Treatment Group Mean at Post-Test
   lincom _cons*1 + c.time*1 + c.treat*1 + c.time#c.treat*1, small
// Time Slope for Control Group
   lincom c.time*1 + c.time#c.treat*0, small
// Time Slope for Treatment Group
   lincom c.time*1 + c.time#c.treat*1, small
// Group Slope at Pre-Test
   lincom c.treat*1 + c.time#c.treat*0, small
// Group Slope at Post-Test
   lincom c.treat*1 + c.time#c.treat*1, small
// Time*Group Interaction Slope
   lincom c.time#c.treat*1, small
```

R Syntax and Output using `gls` from `nlme` Package:

```
print("Eq 3.7 (bottom): Within-Person (U+e) Conditional (Predictor) Model")
print("Not using factor variables, manually dummy coding group and time")
cond_WP = gls(data=Example3a, method="REML", model=outcome~1+time+treat+time:treat,
              correlation=corCompSymm(form=~1|PersonID) )
print("Show results with -2LL"); -2*logLik(cond_WP); summary(cond_WP)
```

```
'log Lik.' 544.69899 (df=6) → -2LL for model -- used in homework
       AIC       BIC    logLik
  556.69899 572.08508 -272.3495

Correlation Structure: Compound symmetry
      Rho
0.83672484 → ICC for correlation across occasions

Coefficients:
               Value  Std.Error   t-value p-value  fixed effect
(Intercept) 49.076798 1.04354249 47.029036  0.0000  beta0
time         5.822365 0.59632799  9.763696  0.0000  beta1
treat        1.681942 1.47579195  1.139688  0.2573  beta2
time:treat   2.042527 0.84333513  2.421963  0.0173  beta3
```

The **RCORR matrix** gives the ICC in the off-diagonal:

$$\text{ICC} = \frac{\tau^2_{U_0}}{\tau^2_{U_0} + \sigma^2_e} = \frac{22.779}{27.779} = .\mathbf{837}$$

```
print("Show R and RCORR matrices")
getVarCov(cond_WP); corMatrix(cond_WP$modelStruct$corStruct)[[5]]
```

```
Marginal variance covariance matrix
       [,1]   [,2]
[1,] 27.225 22.779
[2,] 22.779 27.225

           [,1]       [,2]
[1,] 1.00000000 0.83672484
[2,] 0.83672484 1.00000000
```

The **R matrix** for total *leftover* variance and covariance over time now has a "compound symmetry" or "exchangeable" structure: equal variance on the diagonal, "CS" covariance on the off-diagonal as Var(U0).
So Var(e) = Var(total) – Var(U0) = 27.225 – 22.779 = 4.446

The **RCORR matrix** gives the ICC in the off-diagonal (as **rho** above)

```
print("Get and show joint test of model R2 using correct denominator DF")
modelF_WP=glht(model=cond_WP, df=48, linfct=c("time=0","treat=0","time:treat=0"))
WP=summary(modelF_WP, test=Ftest()); WP # Print results
```

```
Global Test:
   Chisq DF Pr(>Chisq)
1 272.93  3 7.1727e-59
```

```
print("Get and show hidden results for F, dfnum, dfden, and p-value")
WP$test$fstat; WP$test$df; WP$df
          [,1]
[1,] 90.976467 → F-value = Chisq / #slopes
[1] 3          → numerator DF = # slopes
[1] 48         → denominator DF based on DF=2 for intercept and treat

pf(WP$test$fstat,df1=WP$test$df,df2=WP$df,lower.tail=FALSE)
             [,1]
[1,] 8.1574706e-20  → p-value for F-test
```

Equation 3.7 (bottom) Repeated:

$$y_{ti} = \beta_0 + \beta_1(\text{Time}_{ti}) + \beta_2(\text{Treat}_i) + \beta_3(\text{Time}_{ti})(\text{Treat}_i) + U_{0i} + e_{ti}$$

```
print("Get linear combination estimates, add correct denominator DF")
estimates_WP = glht(model=cond_BP, df=48, linfct=rbind(
       "Control Group Mean at Pre-Test"     =c(1,0,0,0), # Multiply ordered fixed effects
       "Control Group Mean at Post-Test"    =c(1,1,0,0),
       "Treatment Group Mean at Pre-Test"   =c(1,0,1,0),
       "Treatment Group Mean at Post-Test"  =c(1,1,1,1),
       "Time Slope for Control Group"       =c(0,1,0,0),
       "Time Slope for Treatment Group"     =c(0,1,0,1),
       "Group Slope at Pre-Test"            =c(0,0,1,0),
       "Group Slope at Post-Test"           =c(0,0,1,1),
       "Time*Group Interaction Slope"       =c(0,0,0,1)))
summary(estimates_WP, test=adjusted("none"))
confint(estimates_WP, level=.95, calpha=univariate_calpha())
```

```
Linear Hypotheses:
                                     Estimate Std. Error t value           Pr(>|t|)  b=beta
Control Group Mean at Pre-Test == 0    49.07680    1.04354 47.0290          < 2.2e-16  b0
Control Group Mean at Post-Test == 0   54.89916    1.04354 52.6085          < 2.2e-16  b0+b1
Treatment Group Mean at Pre-Test == 0  50.75874    1.04354 48.6408          < 2.2e-16  b0+b2
Treatment Group Mean at Post-Test == 0 58.62363    1.04354 56.1775          < 2.2e-16  b0+b1+b2+b3

Time Slope for Control Group == 0       5.82237    0.59633  9.7637 0.0000000000005509  b1
Time Slope for Treatment Group == 0     7.86489    0.59633 13.1889          < 2.2e-16  b1+b3
Group Slope at Pre-Test == 0            1.68194    1.47579  1.1397            0.26007  b2
Group Slope at Post-Test == 0           3.72447    1.47579  2.5237            0.01498  b2+b3
Time*Group Interaction Slope == 0       2.04253    0.84334  2.4220            0.01926  b3
(Adjusted p values reported -- none method)
```

```
print("Show likelihood ratio test comparing conditional means variance model fit")
anova(cond_WP, cond_BP) # Order of "more", "fewer"
```

```
        Model df       AIC       BIC     logLik   Test   L.Ratio p-value
cond_WP     1  6 556.69899 572.08508 -272.34950
cond_BP     2  5 612.50705 625.32879 -301.25352 1 vs 2 57.808052  <.0001
```

The LRT above is a significance test of the *conditional* intraclass correlation (ICC), which in turn provides an effect size for the amount of constant dependency attributed to person mean differences *leftover* in the outcome.

Let's compare the two sets of fixed effects and their F-tests (left = BP vs right = WP)—how do they differ?

BP Coefficients:

	Value	Std.Error	t-value	p-value
(Intercept)	49.076798	1.0435425	47.029036	0.0000
time	5.822365	1.4757920	3.945248	0.0002
treat	1.681942	1.4757920	1.139688	0.2573
time:treat	2.042527	2.0870850	0.978650	0.3302

```
[1,] 16.892107 → F-value = Chisq / #slopes
[1] 3          → numerator DF = # slopes
[1] 48         → denominator DF based on DF=2
[1,] 0.0000001262522  → p-value for F-test
```

WP Coefficients:

	Value	Std.Error	t-value	p-value
(Intercept)	49.076798	1.04354249	47.029036	0.0000
time	5.822365	0.59632799	9.763696	0.0000
treat	1.681942	1.47579195	1.139688	0.2573
time:treat	2.042527	0.84333513	2.421963	0.0173

```
[1,] 90.976467 → F-value = Chisq / #slopes
[1] 3          → numerator DF = # slopes
[1] 48         → denominator DF based on DF=2
[1,] 8.1574706e-20  → p-value for F-test
```

Eq 3.7 (bottom): $y_{ti} = \beta_0 + \beta_1(\text{Time}_{ti}) + \beta_2(\text{Treat}_i) + \beta_3(\text{Time}_{ti})(\text{Treat}_i) + U_{0i} + e_{ti}$

Think about person as a factor in a three-way factorial design of person by time by treatment…
What other terms that could possibly be included are missing? Are they really missing???