Example 2: Empty Models and Level-2 Predictors in General Multilevel Models for Two-Level Nested Outcomes (complete syntax, data, and output available for R electronically)

The data for this example come from the High School and Beyond dataset (HSB4) used in McNeish (2023). Using 7,185 students from 160 schools (an average cluster size of ~45 students per school), we will examine the extent to which student math can be predicted from school-level variables of school size and public versus private. This example uses several custom functions to process output and add additional computations, as shown below (all available in the online files).

R Syntax for Packages; Importing and Preparing Data for Analysis

```
# Install and load required packages
if (!require(pacman)) install.packages("pacman")
pacman::p_load(readxl, TeachingDemos, nlme, lme4, lmerTest, performance, psychometric,
r2mlm)
# readxl: import .xls or .xlsx files; TeachingDemos: save output to text file
# nlme: multilevel models with or without random effects; lme4: multilevel models
# lmerTest: denominator DF and p-values for lme4 models, also univ and Multiv Wald tests
# performance: ICC1; psychometric: ICC2 using lme; r2mlm: R&S pseudo-R2

# Set working directory (to import and export files to)
setwd("C:/Dropbox/25_PSQF6272/PSQF6272_Example2")

# Load R functions for this class from R file in folder above working directory
source("../EDF9850_Functions.R") # Remove ../ for same folder

# Import Example 2 excel data file in sheet "HSB4"
Example2 = read_excel(path="Example2_Data.xlsx", sheet="HSB4")
# Convert to data frame to use for analysis
Example2 = as.data.frame(Example2)
# Filter to only cases complete on all variables to be used below
Example2 = Example2[complete.cases(Example2[ , c("math","size","private")]),]

# Use function from above to add school mean math to student-level-1 data
Example2 = addUnitMeans(data=Example2, unitVariable="schoolID",
                        meanVariables=c("math"), newNames=c("CM_math"))
# Create school-level-2 data frame with one row per school
Example2Means = unique(Example2[ ,
                       c("schoolID","NperschoolID","CM_math","size","private")])

print("R Descriptive Statistics from School-Level-2 Data")
univSummary(data=Example2Means, vars=c("NperschoolID","CM_math","size","private"))
```

```
Variable       N     Mean      SD   Variance     Min      Max
 NperschoolID 160   44.906  11.855    140.538  14.000   67.000
 CM_math      160   12.621   3.118      9.720   4.240   19.719
 size         160 1097.825 629.506 396278.347 100.000 2713.000
 private      160    0.438   0.498      0.248   0.000    1.000
```

```
print("R Descriptive Statistics from Student-Level-1 Data")
univSummary(data=Example2, vars=c("math","size","private"))
```

```
Variable    N     Mean      SD   Variance     Min      Max
 math     7185   12.748   6.878     47.310  -2.832   24.993
 size     7185 1056.862 604.172 365024.409 100.000 2713.000
 private  7185    0.493   0.500      0.250   0.000    1.000
```

Note that the means do not match for the level-2 predictors size and private across the level-2 and level-1 datasets (=unweighted vs weighted)

```
# Center and re-scale school size; private is already 0/1
Example2$size100 = (Example2$size-1000)/100
```

Model 0: Single-Level Empty Means, NO Random Intercept (2 parameters)

***(using the `gls` function from the `nlme` package to omit random intercept variance for now)*:**

$$Math_{pc} = \beta_0 + e_{pc}$$

```
print("R Model 0: Single-Level Empty Means, NO Random Intercept")
Model0 = gls(data=Example2, model=math~1)
glsSummary(Model0, CI=TRUE, digits=2)
```

nlme::gls estimates models without random effects. glsSummary formats output and adds −2LL, residual variance, and optional fixed effect confidence intervals (via CI=TRUE)

```
Model Fit for Predicting math:
       AIC      BIC      logLik  -2*logLik
 48106.917 48120.676 -24051.459  48102.917

Residual Variance (sigma^2):
47.310 σ²_e

Fixed Effects (with 95% CI):
Term          Est    SE   DF t.value p.value  Lower  Upper
Intercept  12.748 0.081 7184 157.099   <.001 12.589 12.907  beta0
```

Model 1: Two-Level Empty Means, WITH Random Intercept (3 parameters)

(using the `lmer` function from the `lme4` package instead):

Level 1: $Math_{pc} = \beta_{0c} + e_{pc}$

Level 2: $\beta_{0c} = \gamma_{00} + U_{0c}$

```
print("R Model 1: Two-Level Empty Means, WITH Random Intercept")
Model1 = lmer(data=Example2, REML=TRUE, formula=math~1+(1|schoolID))
print("Show results with all relevant empty model options turned on")
lmerSummary(Model1, CI=TRUE, ICC1=TRUE, LRT=TRUE,
            ICC2=TRUE, randCI=TRUE, Neff=TRUE)
# CI and randCI = fixed effects and random effects confidence intervals
# ICC1 and ICC2 = intraclass correlations 1 and 2
# LRT = likelihood ratio tests for each random effect variance from ranova
# Neff = design effect and effective sample size using mean cluster size
```

```
Model Fit for Predicting math:
       AIC        BIC     logLik  -2*logLik
 47122.793  47143.433 -23558.397  47116.793  → -2LL used in HW to verify model

Model for the Variance:
Groups     Name      Variance Std.Dev
 schoolID  Intercept    8.614   2.935  Level-2 Var(U0)
 Residual              39.148   6.257  Level-1 Var(e)
Number of obs: 7185 , groups: schoolID, 160

Fixed Effects (with 95% CI):
 Term          Est    SE      DF t.value p.value  Lower  Upper
 Intercept  12.637 0.244 156.647  51.707   <.001 12.154 13.120  gamma00
```

The fixed intercept is now the mean of the school means, which will differ from the overall sample mean (as given by the model without a random intercept variance) whenever level-2 units have different level-1 sizes (i.e., the clusters are "unbalanced").

```
ICC1 intraclass correlation:
    Adjusted ICC: 0.180 → after predictors
  Unadjusted ICC: 0.180 → empty means model
```

$$ICC_1 = \frac{\tau^2_{U_0}}{\tau^2_{U_0} + \sigma^2_e} = \frac{8.614}{8.614 + 39.148} = .\mathbf{180}$$

```
LRT for random effect(s):
               Npar     LL  -2LL   AIC   BIC     LRT DF p.value p.mix
<none>            3 -23558 47117 47123 47143
(1 | schoolID)    2 -24051 48103 48107 48121 986.124  1   <.001 <.001
```

Likelihood Ratio Test (LRT) Statistic $= -2(-23{,}558.40 + 24{,}051.46) = 986.124$

The LRT above is a likelihood ratio test comparing this model to a single-level regression (without a random intercept) using a chi-square (χ^2) distribution with a regular DF=1 distribution. For the more correct mixture version using DF=0,1, you can cut the *p*-value in half in this case, as given by p.mix above. This LRT tests the significance of the intraclass correlation (ICC_1), which in turn provides an effect size for the amount of constant dependency attributed to school mean differences in math.

```
Design Effect and Effective Sample Size:
  Design Effect (using mean n = 44.906 per cluster): 8.919
  Effective N: 805.622
```

This design effect means that our power to detect fixed slopes of level-1 person predictors will be approximately that of an independent sample of 806 students—only if the ICC_1 = 0 would we have power for level-1 effects based on the actual number of students. Power for fixed slopes of level-2 cluster predictors is based on the number of schools (160).

```
ICC2 reliability of the cluster means
(using mean n = 44.906 per cluster):
0.908
```

$$ICC_2 = \frac{\tau^2_{U_0}}{\tau^2_{U_0} + \sigma^2_e} = \frac{8.614}{8.614 + 39.15/44.906} = .\mathbf{908}$$

```
Random Effect 95% CIs (Fixed Effect +/- 1.96*SD(random effect variance)):
 Term      FixedEffect Variance    SD Lower  Upper
 Intercept      12.637    8.614 2.935 6.884 18.390
```

Random intercept 95% confidence interval:

$$CI = \gamma_{00} \pm z_{crit}\sqrt{\tau^2_{U_0}} = 12.637 \pm 1.96\sqrt{8.614} = 6.884\ to\ 18.390$$

This means that 95% of the schools are expected to have school mean math outcomes between 6.884 and 18.390 (around the average of 12.637).

Model 2: Add Main Effects of Level-2 School Size and Public vs Private (5 parameters)

Size is per 100 students, so the reference school has 1000 students (10*100)—this was incorrect before.

Level 1: $\boldsymbol{Math_{pc} = \beta_{0c} + e_{pc}}$

Level 2: $\boldsymbol{\beta_{0c} = \gamma_{00} + \gamma_{01}(size_c - 10) + \gamma_{02}(private_c) + U_{0c}}$

```
print("Model 2 using lme function from nlme package instead of lmer from lme4")
print("nlme has more options for residual correlation and heterogeneous variance")
print("but doesn't work as well with Satterthwaite DDF as lmerTest")
Model2lme = lme(data=Example2, method="REML", fixed=math~1+size100+private,
                random=~1|schoolID)
lmeSummary(Model2lme, CI=TRUE, effectSizes=TRUE)
```

```
print("R Model 2: Add Main Effects of Level-2 School Size and Public vs Private")
Model2 = lmer(data=Example2, REML=TRUE,
              formula=math~1+size100+private+(1|schoolID))
lmerSummary(Model2, CI=TRUE, effectSizes=TRUE)
# effectSizes adds Cohen's d and partial r per slope computed from t-value
Model Fit for Predicting math:
       AIC        BIC     logLik  -2*logLik
 47092.313  47126.711 -23541.156  47082.313

Model for the Variance:
 Groups    Name      Variance Std.Dev
 schoolID  Intercept    6.618   2.573  Level-2 Var(U0) leftover
 Residual              39.150   6.257  Level-1 Var(e)
Number of obs: 7185 , groups: schoolID, 160

Overall F-test for All Fixed Slopes:
 Model     NumDF   DenDF F.value p.value
 Model2        2 155.936  21.803   <.001

Fixed Effects (with 95% CI):
Term          Est    SE      DF t.value p.value  Lower  Upper
Intercept  11.182 0.321 156.273  34.827   <.001 10.548 11.816  gamma00
size100     0.061 0.039 155.936   1.574   0.117 -0.016  0.138  gamma01
private     3.151 0.490 152.002   6.436   <.001  2.184  4.119  gamma02

Effect Sizes for Fixed Slopes:
 Term     Estimate    SE p.value Cohen.d Partial.r
 size100     0.061 0.039   0.117   0.252     0.125
 private     3.151 0.490   <.001   1.044     0.463
```

The model F-test, $F(2, 156) = 21.803$, $p < .0001$, is analogous to the F-test of the model R^2 in single-level regression. Here, it provides a significance test for the R^2 values given below.

Intercept γ_{00} =

size100 γ_{01} =

private γ_{02} =

```
# Pseudo-R2 relative to empty model using custom function
PseudoR2(baseModel=Model1, altModel1=Model2)

Pseudo-R2
 Term       Model  Variance Pseudo.R2
 Intercept  Model1    8.614
 Intercept  Model2    6.618     0.232
 Residual   Model1   39.148
 Residual   Model2   39.150    -0.000
```

Pseudo-R^2 Results: The two fixed slopes of our two level-2 predictors accounted for 23% of the level-2 random intercept variance and ~0% of the student-level-1 residual variance (as expected for school-level-2 predictors).

```
print("Rights & Sterba R2 suite -- f2 is for random intercept")
r2mlm::r2mlm(model=Model2, bargraph=FALSE)
$R2s
       total within between
f1  0.000000      0      NA → Level-1 residual variance
f2  0.044375     NA 0.24307 → Level-2 random intercept variance
v   0.000000      0      NA
m   0.138188     NA 0.75693
f   0.044375     NA      NA
fv  0.044375      0      NA
fvm 0.182563     NA      NA
```

```
# Total R2 relative to empty model
TotalR2(altModel1=Model2)
 Model    Total.R2
 Model2      0.043
```

Total-R^2 Results: The 2 fixed slopes of our 2 level-2 predictors accounted for 4% of the total math variance (as approximated by .232*. 180 = 4.3% using ICC_1 from the empty means model).

Model 3: Add Interaction of Level-2 School Size and Public vs Private (6 parameters)

Level 1: $\boldsymbol{Math_{pc} = \beta_{0c} + e_{pc}}$

Level 2: $\boldsymbol{\beta_{0c} = \gamma_{00} + \gamma_{01}(size_c - 10) + \gamma_{02}(private_c) + \gamma_{03}(size_c - 10)(private_c) + U_{0c}}$

Slope of school size for private schools = $\boldsymbol{[\gamma_{01} + \gamma_{03}(private_c)](size_c - 10)}$

R Syntax and Output for Model 3:

```
print("R Model 3: Add Interaction of Level-2 School Size and Public vs Private")
Model3 = lmer(data=Example2, REML=TRUE,
              formula=math~1+size100*private+(1|schoolID))
lmerSummary(Model3, CI=TRUE)
```

The * in the fixed effects means "give me the two-way interaction and all lower-order main effects."

```
Model Fit for Predicting math:
       AIC        BIC     logLik  -2*logLik
 47095.398  47136.677 -23541.699  47083.398

Model for the Variance:
 Groups    Name       Variance Std.Dev
 schoolID  Intercept     6.569  2.563   Level-2 Var(U0) leftover
 Residual               39.151  6.257   Level-1 Var(e)
Number of obs: 7185 , groups: schoolID, 160

Overall F-test for All Fixed Slopes:
 Model  NumDF   DenDF F.value p.value
 Model3     3 151.894  15.261   <.001

Fixed Effects (with 95% CI):
Term               Est    SE      DF t.value p.value  Lower  Upper
Intercept       11.288 0.329 155.264  34.299   <.001 10.637 11.938  gamma00
size100          0.031 0.045 157.645   0.685   0.495 -0.058  0.119  gamma01
private          3.252 0.494 150.004   6.589   <.001  2.277  4.227  gamma02
size100:private  0.124 0.090 149.283   1.377   0.171 -0.054  0.302  gamma03
```

Intercept γ_{00} =

size100 γ_{01} =

private γ_{02} =

size100*private γ_{03} =

```
print("Size slope for private schools = size100 + size100:private")
Model3con1D = contest1D(Model3, L=c(0,1,0,1))
con1DSummary(Model3con1D, CI=TRUE, effectSizes=TRUE) # Custom output

Fixed Effect (with 95% CI):
   Est    SE      DF t.value p.value Lower Upper
 0.155 0.078 146.678   1.979   0.050 0.000 0.309
```

```
Effect Size:
   Est    SE p.value Cohen.d Partial.r
 0.155 0.078   0.050   0.327     0.161
```

```
# Pseudo-R2 relative to empty model using custom function
PseudoR2(baseModel=Model1, altModel1=Model2, altModel2=Model3)
```

```
Pseudo-R2 and Change in Pseudo-R2
 Term       Model  Variance Pseudo.R2 Change.R2
 Intercept  Model1    8.614
 Intercept  Model2    6.618     0.232
 Intercept  Model3    6.569     0.237     0.006
 Residual   Model1   39.148
 Residual   Model2   39.150    -0.000
 Residual   Model3   39.151    -0.000    -0.000
```

Pseudo-R^2 Results: The 3 fixed slopes of our 3 level-2 predictors accounted for 23.7% of the level-2 random intercept variance, an increase of 0.5%,. The alternative using **r2mlm** below provides a slightly higher estimate.

```
print("Rights & Sterba R2 and change -- f2 is for random intercept")
r2mlm::r2mlm_comp(modelA=Model2, modelB=Model3, bargraph=FALSE)
```

```
$`Model A R2s`
       total within between
f1  0.000000      0      NA
f2  0.044375     NA 0.24307
v   0.000000      0      NA
m   0.138188     NA 0.75693
f   0.044375     NA      NA
fv  0.044375      0      NA
fvm 0.182563     NA      NA

$`Model B R2s`
       total within between
f1  0.000000      0      NA
f2  0.046869     NA 0.25497
v   0.000000      0      NA
m   0.136954     NA 0.74503
f   0.046869     NA      NA
fv  0.046869      0      NA
fvm 0.183823     NA      NA

$`R2 differences, Model B - Model A`
         total within   between
f1   0.0000000      0        NA
f2   0.0024941     NA  0.011901
v    0.0000000      0        NA
m   -0.0012336     NA -0.011901
f    0.0024941     NA        NA
fv   0.0024941      0        NA
fvm  0.0012605     NA        NA
```

```
# Total R2 relative to empty model using custom function
TotalR2(altModel1=Model2, altModel2=Model3)
```

```
Total-R2 and Change in Total-R2
 Model  Total.R2 Change.R2
 Model2    0.043
 Model3    0.045     0.002
```

Total-R^2 Results: The 3 fixed slopes of our 3 level-2 predictors accounted for 4.53% of the total math variance, an increase of 0.2% relative to the main effects model.

Sample Results Section [indicates notes about what to customize or also include; note that SE and *p*-values are not needed if you provide a table for the model solution]

The extent to which the extent to which student math outcomes could be predicted from school-level variables of school size and public versus private status was examined in a series of multilevel models in which the 7,185 students were modeled as nested within their 160 schools. The number of students included per school ranged from 14 to 67 (M ≈ 45). Restricted Maximum likelihood (REML) within the lmer function in the lme4 R package (v. 2.0.6; Bates et al., 2015) was used in estimating and reporting all model parameters. The significance of fixed effects was evaluated with univariate and multivariate Wald tests using Satterthwaite denominator degrees of freedom, whereas random effects were evaluated via likelihood ratio tests (i.e., −2ΔLL with degrees of freedom equal to the number of new random effects variances and covariances). Alpha was chosen as .01. Model-implied fixed effects were requested via contest1D statements (v. 3.2.1; Kuznetsova et al., 2017). Fixed effect sizes were evaluated via pseduo-R^2 values for the proportion reduction in each variance component relative to a nested model without the predictors in question, as well as with total-R^2, the squared correlation between the actual math outcomes and those predicted by the model fixed effects.

As derived from an empty means, random intercept model, student math had an intraclass correlation of ICC = .180, indicating that 18.0% of the variance in student math was between schools, a significant amount, −2ΔLL(1) = 986.12, p < .0001. Given an average of 45 students per school in this sample, the ICC_1 = .180 translated into a design effect = 8.92, further indicating the need for a multilevel analysis. The school mean math outcomes had strong reliability, as evidenced by ICC_2 = .908. The fixed intercept was 12.637 (SE = 0.244), which represented the expected average school mean math outcome. A random intercept confidence interval (computed as the fixed intercept ± 1.96*SQRT[random intercept variance]) indicated that 95% of the schools were expected to have school mean math outcomes between 6.884 and 18.390 (around the average of 12.637).

We then added the school-level predictors of school size (per 100 students, centered so that 0 = 1000 students) and status (public = 0 versus private = 1). These two fixed effects accounted for significant variance overall, F(2, 157.31) = 21.80, p < .0001, including 23.2% of the level-2 random intercept variance and 4.3% of the total variance. The fixed intercept was 11.182 (SE = 0.321), which represented the expected school mean math outcome for public schools with 1000 students. The slope for school size was nonsignificantly positive, indicating that school mean math was expected to be nonsignificantly higher by 0.061 (SE = 0.039, p = .118) per 100 more students. The slope for public vs. private was significantly positive, indicating school mean math was predicted to be higher by 3.151 (SE = 0. 490, p < .0001) for private schools relative to public schools.

We then added an interaction between school size and type. The three fixed effects still accounted for significant variance overall, F(3, 154.92) = 15.26, p < .0001, including 23.7% of the level-2 random intercept variance (an increase of 0.6%) and 4.5% of the total variance (an increase of 0.2%). The interaction slope was nonsignificantly positive, indicating that the slope of school size for private schools (Est = 0.155, SE = 0.078, p = .050) was 0.124 nonsignificantly more positive than the size slope for public schools (Est = 0.031, SE = 0.045, p = .495). The slope for public vs. private remained significantly positive, indicating school mean math was predicted to be higher by 3.252 (SE = 0.494, p < .0001) for private schools relative to public schools (both of size 1000 students).

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software, 82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13