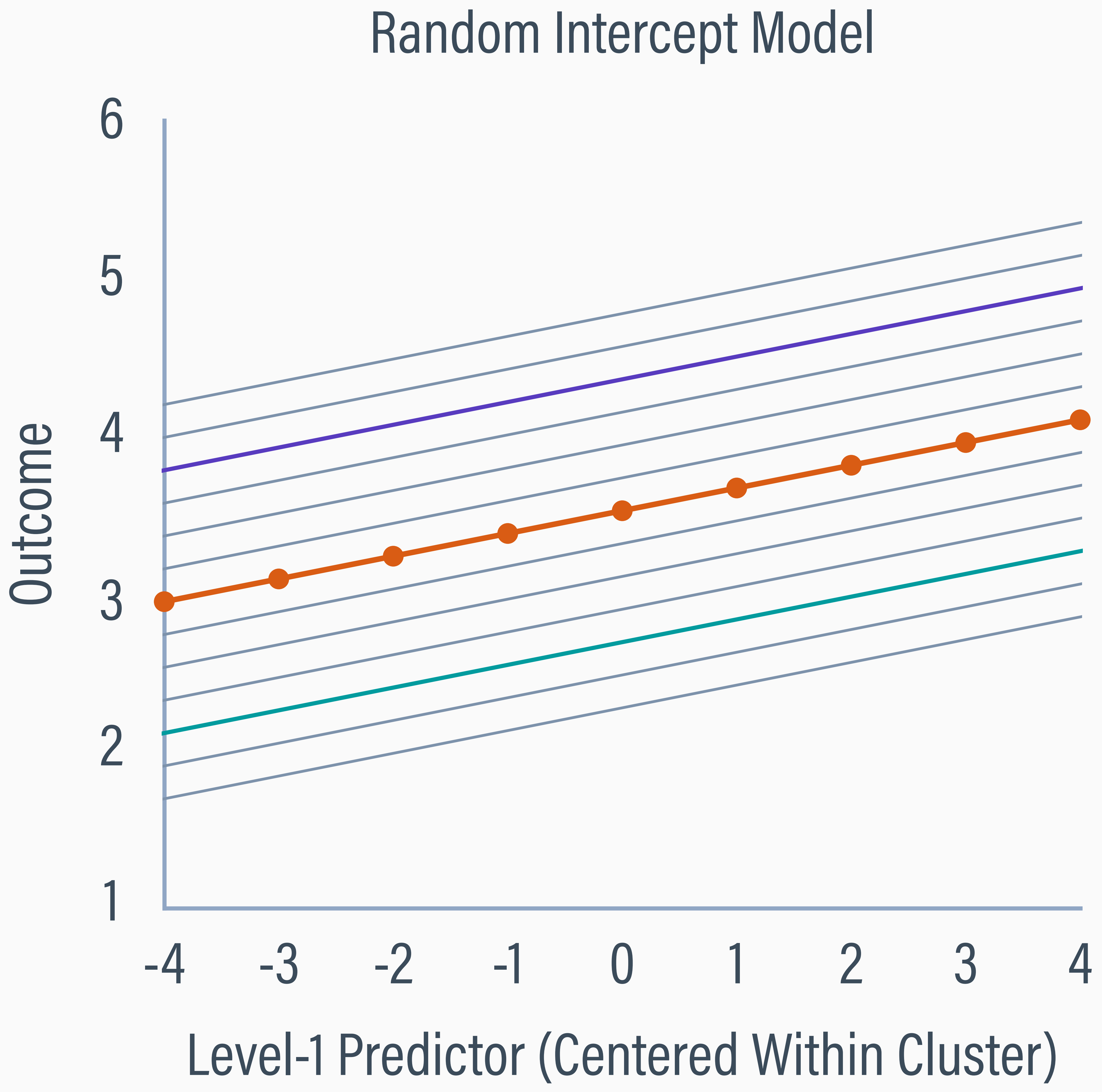
MODULE 5

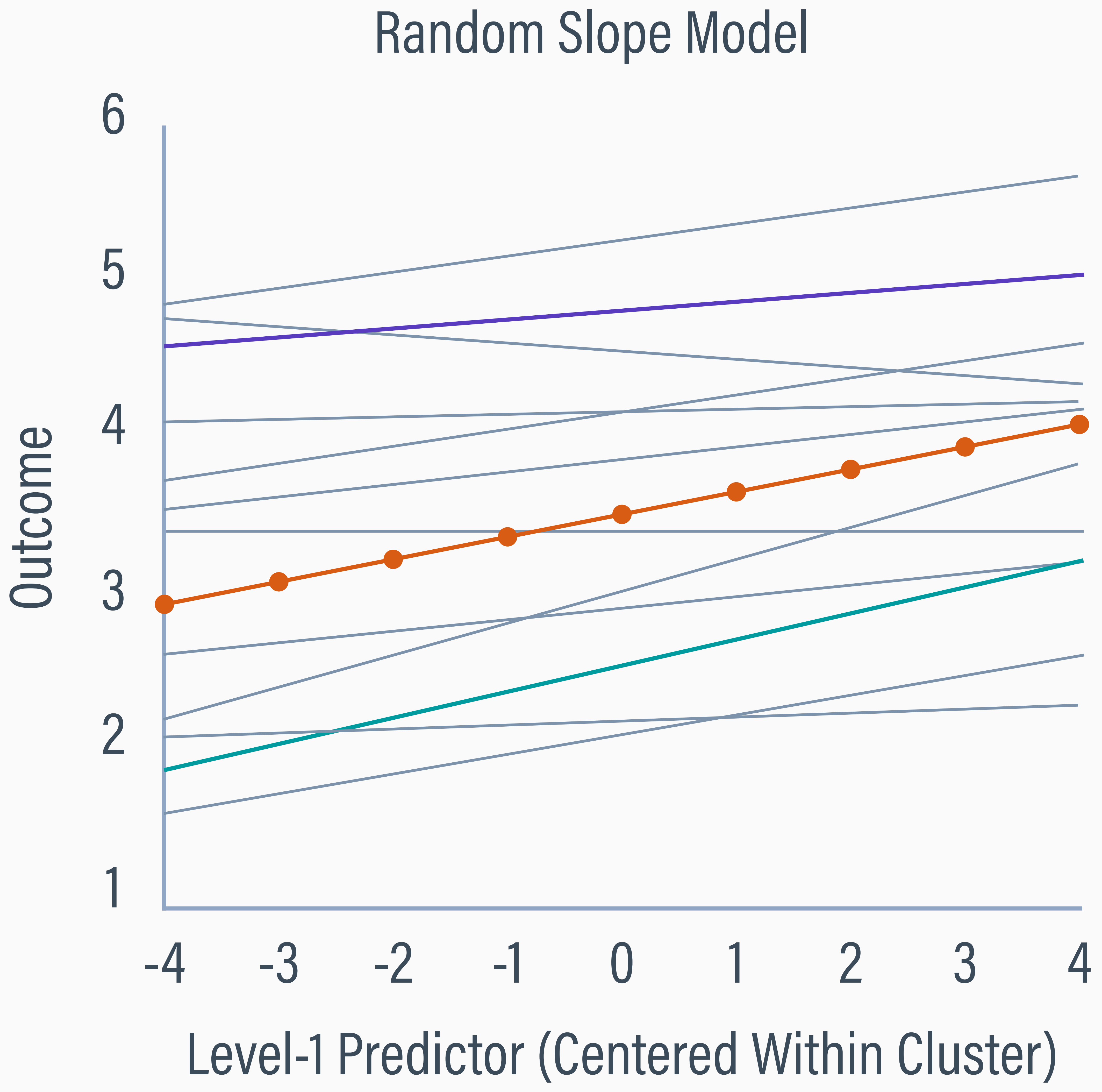
RANDOM COEFFICIENT MODELS

RANDOM SLOPE COEFFICIENTS

- Each cluster contains multiple observations, effectively forming its own dataset and allowing for cluster-specific slopes
- Random slope (random coefficient) models allow the effect of a level-1 predictor to vary across level-2 units
- The influence of a level-1 predictor on the outcome may be stronger in some clusters than in others
- Level-2 predictors cannot have cluster-varying slopes because they have only one observed value per cluster

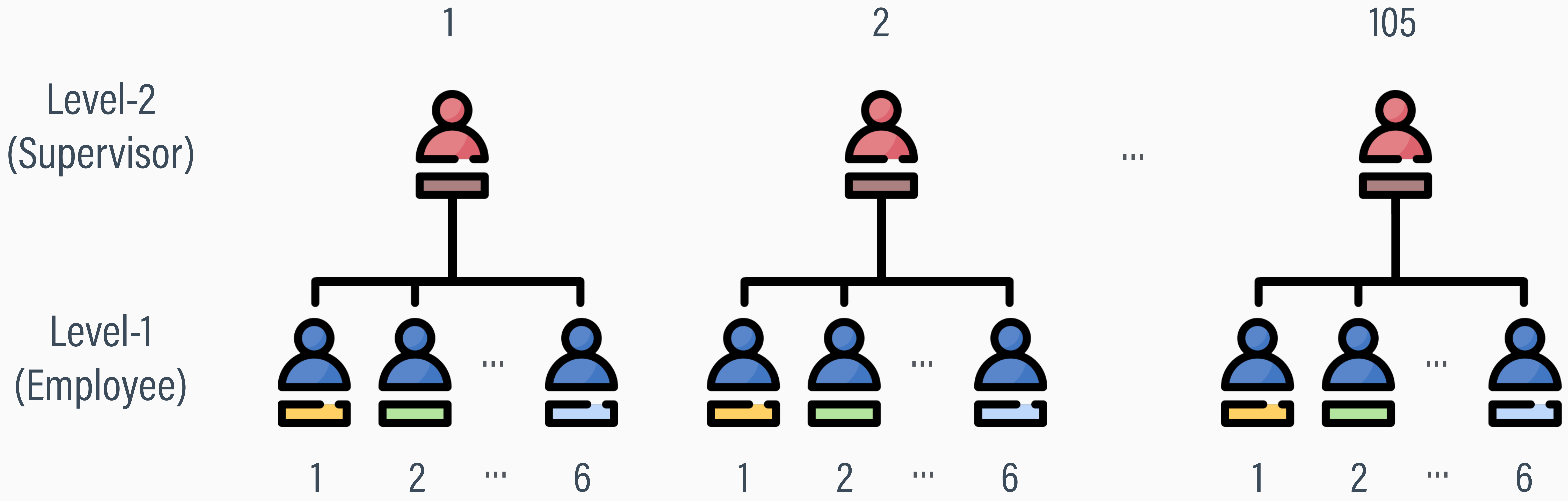
MODELING OVERVIEW





ORGANIZATIONAL APPLICATION

- $n_j = 6$ employees at level-1 nested within $J = 105$ teams or workgroups at level-2 ($N = 630$ data records in total)



VARIABLE INFORMATION

= predictor measured at level-2

= predictor measured at level-1

= outcome measured at level-1

empower = lmx + male + climate

Variable	Definition	Level	Scale
Team	Team-level (level-2) identifier	2	Integers (1 to 105)
Empower	Employee empowerment	1	Numeric (14 to 42)
LMX	Leader–member exchange (supervisor–supervisee relationship quality)	1	Numeric (0 to 17)
Male	Sex dummy code	1	Female = 0, Male = 1
Climate	Leadership climate	2	Numeric (10 to 33)

DATA STRUCTURE

- Stacked (long) data format where each level-2 unit (supervisor) has one row per level-1 (supervisee) observation
- The i subscript indexes level-1 observations, and j indexes level-2 units
- Variables measured at level-2 repeat across all rows within a cluster

Row	i	j	$JobSat_{ij}$	LMX_{ij}	$Climate_j$	
1	1	1	1.0	5.6	7.3	
2	2	1	2.0	4.3	7.3	
…	…	1	…	…	…	Team 1
6	6	1	5.0	7.3	7.3	
7	1	2	7.0	3.9	4.0	
8	2	2	3.0	7.1	4.0	Team 2
…	…	2	…	…	…	
12	6	2	6.0	3.5	4.0	
…	…	…	…	…	…	
625	1	105	2.0	5.4	3.3	
626	2	105	3.0	3.5	3.3	Team 105
…	…	105	…	…		
630	6	105	1.0	7.9	3.3	

OUTLINE

1. Modeling Step 1: Estimate ICCs
2. Modeling Step 2: Random Intercept Model
3. Modeling Step 3: Random Slope Model 1
4. Deciding Whether a Random Slope Is Needed
5. Modeling Step 3: Random Slope Model 2
6. Checking MLM Assumptions
7. Latent Variable Specification

ESTIMATING INTRACLASS CORRELATIONS

- Estimate intraclass correlations (ICCs) before model fitting to assess how much variance lies within and between clusters
- Some predictors have little or no level-2 variation, in which case they are already “pure” within-cluster variables
- Disaggregating a level-1 predictor and modeling its means is only warranted when there is sufficient level-2 variation

BLIMP SCRIPT 5.1

```
DATA: Employee.dat;
VARIABLES: Employee Team Turnover Male Empower LMX JobSat Climate TeamPerf;
CLUSTERID: Team;
NOMINAL: Male;    # invokes logistic regression model
MODEL:
{ Empower LMX Male } ~ intercept | intercept;    # empty model for all variables in { }
BURN: 20000;
ITERATIONS: 10000;
SEED: 90291;
```

RBLIMP SCRIPT 5 (MODEL 1)

```
model1 <- rblimp(
        data = Employee,
        nominal = 'Male',
        clusterid = 'Team',
        model = '{ Empower LMX Male } ~ intercept | intercept',
        seed = 90291,
        burn = 10000,
        iter = 20000)
output(model1)
```

BLIMP OUTPUT

= level-2 estimate

= level-1 estimate

```
Outcome Variable:  Empower

Parameters                          Estimate     StdDev       2.5%      97.5%      ChiSq     PValue      N_Eff
                                ----------------------------------------------------------------------------
Variances:
  L2 : Var(Intercept)                  2.552      0.811      1.268      4.413        ---        ---   1120.636
  Residual Var.                       18.067      1.122     16.076     20.457        ---        ---   5080.598

Coefficients:
  Intercept                           28.609      0.231     28.154     29.065  15296.126      0.000   3350.180

Standard Deviations:
  L2 : SD(Intercept)                   1.598      0.250      1.126      2.101        ---        ---   1061.977
  Residual SD                          4.251      0.131      4.010      4.523        ---        ---   5088.976

Proportion Variance Explained
  by Coefficients                      0.000      0.000      0.000      0.000        ---        ---        nan
  by Level-2 Random Intercepts         0.124      0.036      0.064      0.200        ---        ---   1109.611
  by Level-1 Residual Variation        0.876      0.036      0.800      0.936        ---        ---   1109.611

                                ----------------------------------------------------------------------------
```

BLIMP OUTPUT

= level-2 estimate
= level-1 estimate

Outcome Variable: LMX

Parameters	Estimate	StdDev	2.5%	97.5%	ChiSq	PValue	N_Eff
Variances:							
L2 : Var(Intercept)	1.313	0.391	0.668	2.185	---	---	1038.524
Residual Var.	7.906	0.494	7.014	8.948	---	---	4320.947
Coefficients:							
Intercept	9.594	0.160	9.283	9.910	3602.019	0.000	3114.055
Standard Deviations:							
L2 : SD(Intercept)	1.146	0.170	0.817	1.478	---	---	940.443
Residual SD	2.812	0.088	2.648	2.991	---	---	4350.316
Proportion Variance Explained							
by Coefficients	0.000	0.000	0.000	0.000	---	---	nan
by Level-2 Random Intercepts	0.143	0.038	0.074	0.223	---	---	1023.422
by Level-1 Residual Variation	0.857	0.038	0.777	0.926	---	---	1023.422

BLIMP OUTPUT

= level-2 estimate

= level-1 estimate

```
Outcome Variable:  Male

Parameters                       Estimate    StdDev      2.5%     97.5%     ChiSq    PValue     N_Eff
                               -----------------------------------------------------------------------
Male.1

Variances:
  L2 : Var(Intercept)               0.094     0.102     0.004     0.372       ---       ---   145.999

Coefficients:
  Intercept                        -0.066     0.087    -0.237     0.108     0.575     0.448  7164.489

Standard Deviations:
  L2 : SD(Intercept)                0.307     0.146     0.063     0.610       ---       ---   112.963

Odds Ratio:
  Intercept                         0.936     0.082     0.789     1.114       ---       ---  7166.122

Proportion Variance Explained
  by Coefficients                   0.000     0.000     0.000     0.000       ---       ---       nan
  by Level-2 Random Intercepts      0.028     0.028     0.001     0.102       ---       ---   140.269
  by Level-1 Residual Variation     0.972     0.028     0.898     0.999       ---       ---   140.269

                               -----------------------------------------------------------------------
```

The ICC for sex is ~ .03. Sex is dummy coded such that female = 0 and male = 1. In small groups of two or three, discuss the meaning of the team-level gender averages (random intercepts). What does it mean for gender to have a non-zero ICC?

SUMMARY

- Both the outcome and LMX predictor have salient level-2 variation in the team-specific means ($ICC_{empower} = .12$ and $ICC_{lmx} = .14$)
- This feature indicates the need for MLM (independence violation) with a disaggregated predictor
- The level-1 sex dummy code has very little level-2 variation in its team-specific means ($ICC_{male} = .03$)
- Disaggregating sex is unnecessary because it already functions primarily as a within-team predictor

OUTLINE

1. Modeling Step 1: Estimate ICCs
2. Modeling Step 2: Random Intercept Model
3. Modeling Step 3: Random Slope Model 1
4. Deciding Whether a Random Slope Is Needed
5. Modeling Step 3: Random Slope Model 2
6. Checking MLM Assumptions
7. Latent Variable Specification

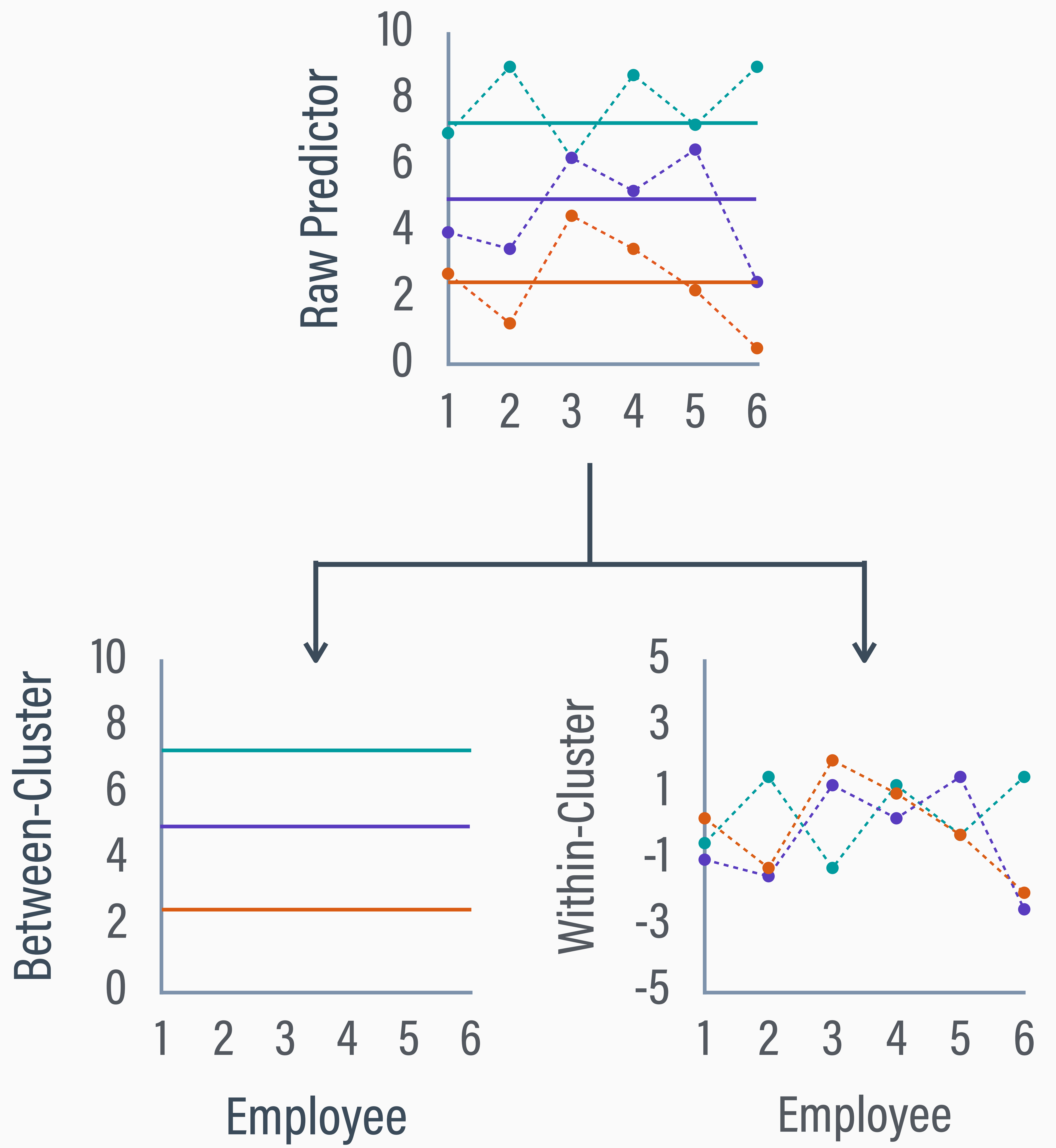
DISAGGREGATED PREDICTOR

- Disaggregation centers each person's score around its level-2 team mean

$$lmx_j^b = \mu_{j(lmx)}$$

$$lmx_{ij}^w = lmx_{ij} - \mu_{j(lmx)}$$

- lmx^w contains only within-team (level-1) variation, and lmx^b reflects only team-level (level-2) mean differences



WITHIN-CLUSTER (LEVEL-1) MODEL

- Empowerment score for employee i in team j is the sum of a level-2 mean (β_{0j}), fixed effects due to within-team predictors (β_1 and β_2), and a within-team residual (ε_{ij})

$$empower_{ij} = \beta_{0j} + \beta_1(lmx_{ij}^{w}) + \beta_2(male_{ij}) + \varepsilon_{ij}$$

- Assumption: residuals are normal with constant variation across all employees (level-1 units) and teams (level-2 units)

$$\varepsilon_{ij} \sim N(0,\sigma_{\varepsilon}^{2})$$

BETWEEN-CLUSTER (LEVEL-2) MODEL

- The empowerment mean for team j (β_{0j}) is the sum of the grand mean (γ_{00}), fixed effects due to team-average LMX and climate, and a between-team residual (u_{0j})

$$\beta_{0j} = \gamma_{00} + \gamma_{01}(lmx_j^b) + \gamma_{02}(climate_j) + u_{0j}$$
$$\beta_1 = \gamma_{10}$$
$$\beta_2 = \gamma_{20}$$

- Assumption: random intercept residuals are normal with constant variation across teams (level-2 units)

$$u_{0j} \sim N(0, \sigma_u^2)$$

In small groups of two or three, perform the substitution demonstrated in the Module 4 slides to get the combined-model notation from the Raudenbush and Bryk (2002) book.

$$\text{empower}_{ij} = \gamma_{00} + \gamma_{10}(\text{lmx}^{w}_{ij}) + \gamma_{20}(\text{male}_{ij})$$
$$+ \gamma_{01}(\text{lmx}^{b}_{j}) + \gamma_{02}(\text{climate}_{j}) + u_{0j} + \varepsilon_{ij}$$

COMMON NOTATIONAL SYSTEMS

Combined-model equation (Raudenbush & Bryk, 2002)

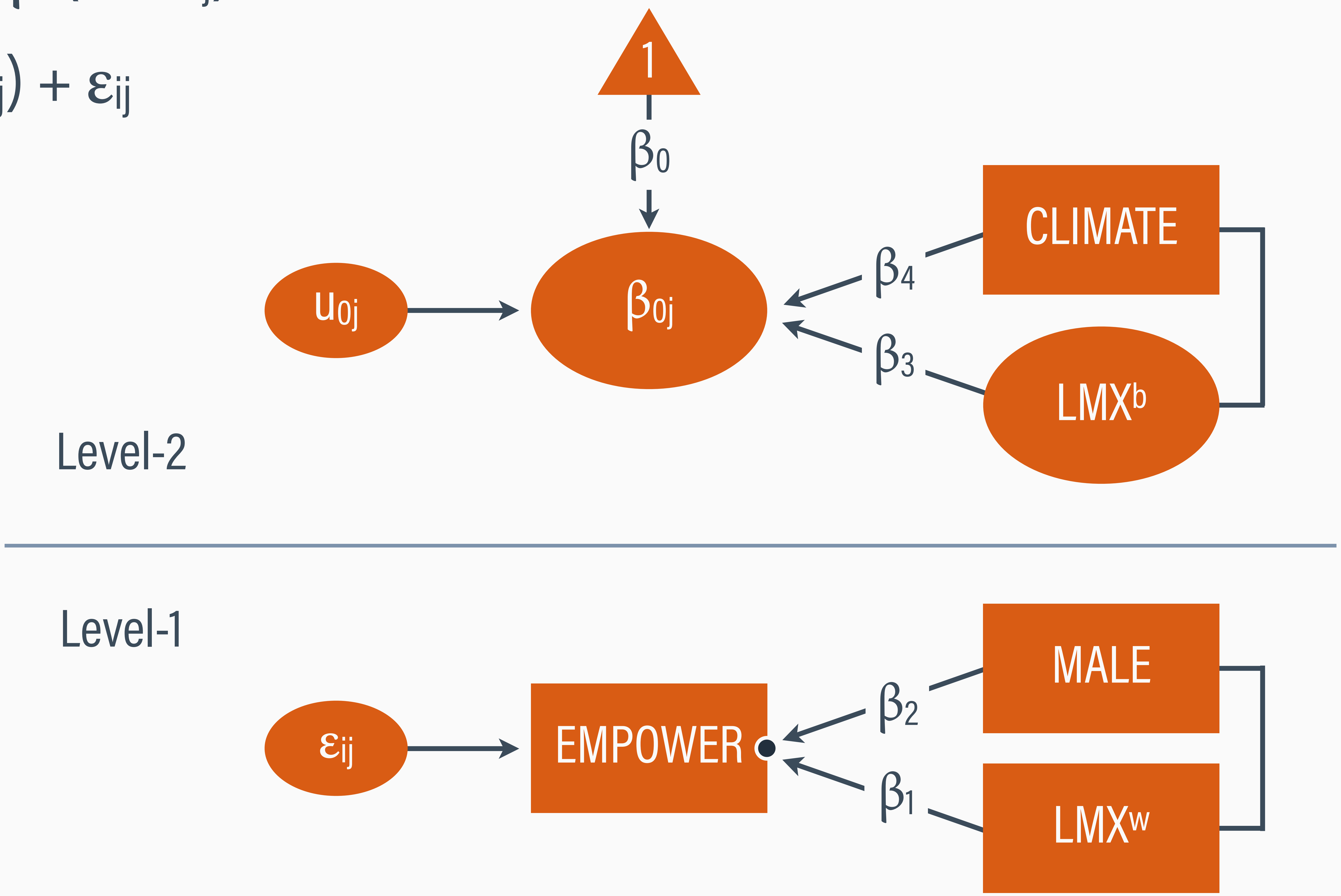
$$\text{empower}_{ij} = \gamma_{00} + \gamma_{10}(\text{lmx}^{w}_{ij}) + \gamma_{20}(\text{male}_{ij}) + \gamma_{01}(\text{lmx}^{b}_{j}) + \gamma_{02}(\text{climate}_{j}) + u_{0j} + \varepsilon_{ij}$$

Standard(ish) regression notation (Scott, Shrout, & Weinberg, 2013)

$$\text{empower}_{ij} = \beta_{0} + \beta_{1}(\text{lmx}^{w}_{ij}) + \beta_{2}(\text{male}_{ij}) + \beta_{3}(\text{lmx}^{b}_{j}) + \beta_{4}(\text{climate}_{j}) + u_{0j} + \varepsilon_{ij}$$

PATH DIAGRAM

$$empower_{ij} = (\beta_0 + u_{0j}) + \beta_1(lmx^{w}_{ij}) + \beta_2(male_{ij}) +$$
$$\beta_3(lmx^{b}_{j}) + \beta_4(climate_{j}) + \varepsilon_{ij}$$



• = random intercept (β_{0j})

BLIMP SCRIPT 5.2

```
DATA: Employee.dat;
VARIABLES: Employee Team Turnover Male Empower LMX JobSat Climate TeamPerf;
NOMINAL: Male;    # invokes categorical variable model
CLUSTERID: Team;
CENTER: groupmean = LMX; grandmean = LMX.mean Climate;
MODEL: Empower ~ intercept LMX Male LMX.mean Climate | intercept;
BURN: 10000;
ITERATIONS: 10000;
SEED: 90291;
```

RBLIMP SCRIPT 5 (MODEL 2)

```
model2 <- rblimp(
    data = Employee,
    nominal = 'Male',
    clusterid = 'Team',
    center = 'groupmean = LMX; grandmean = LMX.mean Climate',
    model = 'Empower ~ intercept LMX Male LMX.mean Climate | intercept',
    seed = 90291,
    burn = 10000,
    iter = 10000)
output(model2)
```

BLIMP OUTPUT

= level-2 estimate
= level-1 estimate
= combined estimate

```
Outcome Variable:  Empower

Grand Mean Centered: Climate LMX.mean[Team]
Group Mean Centered: LMX
```

Parameters	Estimate	StdDev	2.5%	97.5%	ChiSq	PValue	N_Eff
Variances:							
L2 : Var(Intercept)	0.890	0.562	0.097	2.225	---	---	202.846
Residual Var.	13.531	0.838	12.010	15.270	---	---	2690.721
Coefficients:							
Intercept	27.689	0.267	27.170	28.216	10771.431	0.000	2528.639
LMX	0.673	0.056	0.563	0.782	143.910	0.000	4166.700
Male.1	1.894	0.305	1.297	2.494	38.680	0.000	8259.863
LMX.mean[Team]	-0.251	0.311	-0.964	0.268	0.777	0.378	362.172
Climate	0.282	0.052	0.183	0.385	29.835	0.000	1313.674
Standard Deviations:							
L2 : SD(Intercept)	0.943	0.300	0.312	1.492	---	---	166.959
Residual SD	3.678	0.113	3.466	3.908	---	---	2688.662

...

FIXED EFFECT INTERPRETATIONS

- $\beta_0 = 27.69$ is the empowerment mean for the male = 0 group (because all numeric predictors are centered)
- $\beta_1 = 0.67$ is the expected empowerment difference between two LMX scores from the same team that differ by one point, controlling for sex
- $\beta_2 = 1.89$ is the expected empowerment difference for males (relative to females), controlling for within-team LMX

FIXED EFFECTS, CONTINUED

- $\beta_3 = -0.25$ is the expected empowerment difference between two teams whose average LMX ratings differ by one point, controlling for team-level climate

- $\beta_4 = 0.28$ is the expected empowerment difference between two teams whose climate ratings differ by one point, controlling for team-average LMX

RANDOM EFFECT INTERPRETATIONS

- $u_{0j} = \beta_{0j} - (\beta_0 + \beta_3(LMX_j^b) + \beta_4(climate_j))$
- var(u_{0j}) = 0.89 is the average squared distance between the level-2 empowerment means and their predicted values
- sd(u_{0j}) = 0.94 is the average distance between the level-2 empowerment means and their predicted values

- $\varepsilon_{ij} = empower_{ij} - (\beta_{0j} + \beta_1(lmx_{ij}^w) + \beta_2(male_{ij}))$
- var(ε_{ij}) = 13.53 is the average squared distance between the level-1 empowerment scores and their predicted values
- sd(ε_{ij}) = 3.68 is the average distance between the level-1 empowerment scores and their predicted values

BLIMP OUTPUT, CONTINUED

= level-2 estimate
= level-1 estimate
= combined estimate

```
Outcome Variable:  Empower

Grand Mean Centered: Climate LMX.mean[Team]
Group Mean Centered: LMX


Parameters                      Estimate    StdDev     2.5%     97.5%     ChiSq    PValue     N_Eff
                                ---------------------------------------------------------------------

...

Standardized Coefficients:
  LMX                              0.415     0.031    0.351     0.474   174.148     0.000  1976.422
  Male.1                           0.208     0.032    0.143     0.270    40.965     0.000  8210.609
  LMX.mean[Team]                  -0.058     0.063   -0.181     0.068     0.830     0.362   425.266
  Climate                          0.248     0.044    0.163     0.334    31.957     0.000  1311.770

Proportion Variance Explained
  by Coefficients                  0.305     0.032    0.241     0.369       ---       ---  1224.888
  by Level-2 Random Intercepts     0.043     0.026    0.005     0.104       ---       ---   199.719
  by Level-1 Residual Variation    0.649     0.034    0.583     0.716       ---       ---   850.733

                                ---------------------------------------------------------------------
```

EFFECT SIZE INTERPRETATIONS

- $R^2_{(predictors)} = .31$ is the proportion of the total variation explained by all predictors
- $R^2_{(residual\text{-}between)} = .04$ is the proportion of the total variation attributable to the between-cluster residuals (the u_{0j} terms)
- $R^2_{(residual\text{-}within)} = .65$ is the proportion of the total variation attributable to the within-cluster residuals (the ε_{ij} terms)

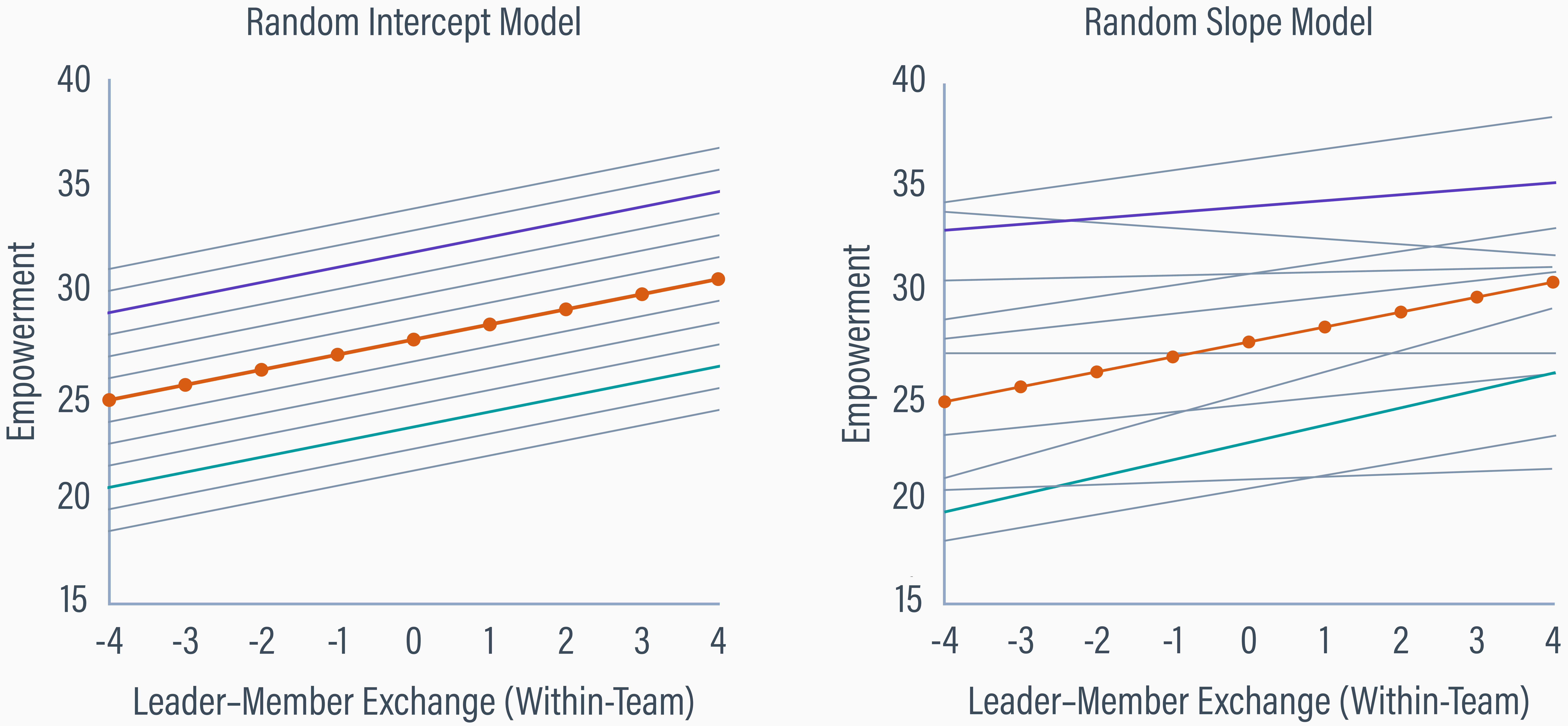
OUTLINE

1. Modeling Step 1: Estimate ICCs
2. Modeling Step 2: Random Intercept Model
3. Modeling Step 3: Random Slope Model 1
4. Deciding Whether a Random Slope Is Needed
5. Modeling Step 3: Random Slope Model 2
6. Checking MLM Assumptions
7. Latent Variable Specification

RANDOM SLOPE COEFFICIENTS

- Random slope or random coefficient models allow the influence of a level-1 predictor to vary across level-2 units
- The influence of LMX or sex could vary across teams
- e.g., In some teams, increases in supervisor–supervisee relationship quality (LMX) may strongly influence empowerment, whereas in others the link may be weaker

MODEL COMPARISON



WITHIN-CLUSTER (LEVEL-1) MODEL

- Both the team-specific empowerment mean (β_{0j}) and the leader-member exchange slope vary across clusters (β_{1j}), but the sex slope (β_2) is constant

$$empower_{ij} = \beta_{0j} + \beta_{1j}(lmx^{w}_{ij}) + \beta_2(male_{ij}) + \varepsilon_{ij}$$

- Assumption: residuals are normal with constant variation across all employees (level-1 units) and teams (level-2 units)

$$\varepsilon_{ij} \sim N(0,\sigma^2_{\varepsilon})$$

BETWEEN-CLUSTER (LEVEL-2) MODEL

- The empowerment mean for team j (β_{0j}) is the sum of the grand mean (γ_{00}), fixed effects (γ_{01} and γ_{02}), and a between-team residual (u_{0j})
- The LMX slope for team j (β_{1j}) is the sum of the mean slope (γ_{10}) and a team-level residual (u_{1j})
- Assumption: random intercept and slope residuals are bivariate normal *and correlated*

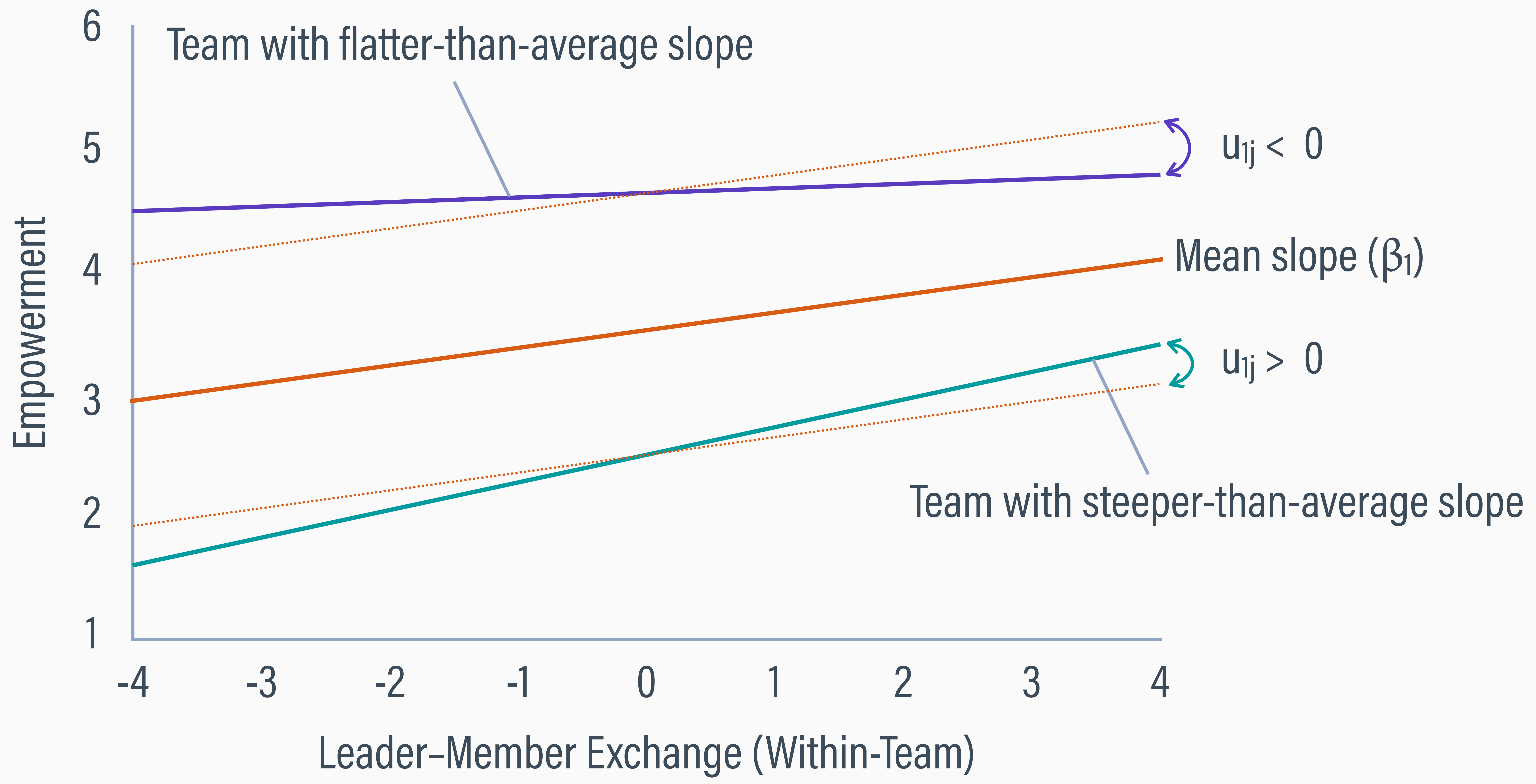
$$\beta_{0j} = \gamma_{00} + \gamma_{01}(lmx_j^b) + \gamma_{02}(climate_j) + u_{0j}$$

$$\beta_{1j} = \gamma_{10} + u_{1j}$$

$$\beta_2 = \gamma_{20}$$

$$\begin{pmatrix} u_{0j} \\ u_{1j} \end{pmatrix} \sim N(0, \boldsymbol{\Sigma}_u) \quad \boldsymbol{\Sigma}_u = \begin{pmatrix} \sigma^2_{u0} & \sigma_{u0u1} \\ \sigma_{u1u0} & \sigma^2_{u1} \end{pmatrix}$$

SLOPE RESIDUAL GRAPHIC



DECODING THE SUBSCRIPTS

The second subscript numbers the coefficients on the right side starting at 0 (the intercept)

$$\beta_{0j} = \gamma_{00} + \gamma_{01}(lmx_j) + \gamma_{02}(climate_j) + u_{0j}$$

$$\beta_{1j} = \gamma_{10} + u_{1j}$$

$$\beta_{2j} = \gamma_{20}$$

The first subscript tells which level-2 outcome variable these terms are attached to (γ_{00} and u_{0j} belong in β_{0j}'s equation, γ_{10} is attached to β_1)

COMBINED-MODEL EQUATION

Substituting the right sides of the level-2 equations ...

$$\beta_{0j} = \gamma_{00} + \gamma_{01}(lmx_j^b) + \gamma_{02}(climate_j) + u_{0j}$$

$$\beta_1 = \gamma_{10} + u_{1j}$$

$$\beta_2 = \gamma_{20}$$

into their coefficients from the level-1 equation

$$jobsat_{ij} = \beta_{0j} + \beta_{1j}(lmx_{ij}^w) + \beta_2(male_{ij}) + \varepsilon_{ij}$$

gives the combined-model regression equation (Raudenbush & Bryk, 2002)

$$empower_{ij} = \overbrace{(\gamma_{00} + u_{0j}) + \gamma_{01}(lmx_j^b) + \gamma_{02}(climate_j)}^{\text{Leading 0 subscript conveys team-level effects}} + \underbrace{(\gamma_{10} + u_{1j})(lmx_{ij}^w) + \gamma_{20}(male_{ij})}_{\text{Non-zero subscript conveys within-team effects}} + \varepsilon_{ij}$$

COMMON NOTATIONAL SYSTEMS

Combined-model equation (Raudenbush & Bryk, 2002)

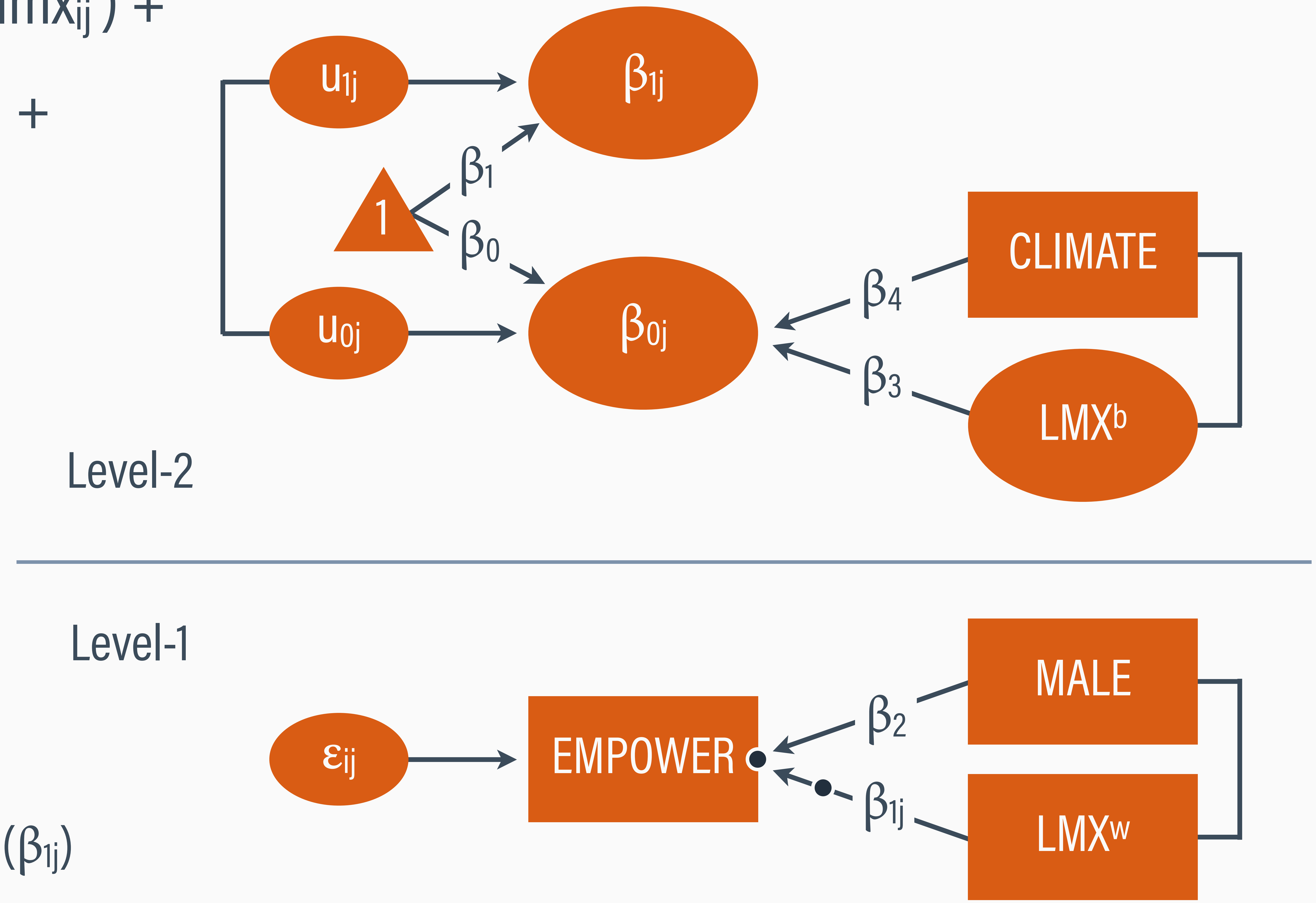
$$empower_{ij} = (\gamma_{00} + u_{0j}) + (\gamma_{10} + u_{1j})(lmx^{w}_{ij}) + \gamma_{20}(male_{ij}) + \gamma_{01}(lmx^{b}_{j}) + \gamma_{02}(climate_{j}) + \varepsilon_{ij}$$

Standard(ish) regression notation (Scott, Shrout, & Weinberg, 2013)

$$empower_{ij} = (\beta_{0} + u_{0j}) + (\beta_{1} + u_{1j})(lmx^{w}_{ij}) + \beta_{2}(male_{ij}) + \beta_{3}(lmx^{b}_{j}) + \beta_{4}(climate_{j}) + \varepsilon_{ij}$$

PATH DIAGRAM

$$\text{empower}_{ij} = (\beta_0 + u_{0j}) + (\beta_1 + u_{1j})(\text{lmx}_{ij}^{w}) + \beta_2(\text{male}_{ij}) + \beta_3(\text{lmx}_j^{b}) + \beta_4(\text{climate}_j) + \varepsilon_{ij}$$



• = random intercept (β_{0j}) and random slope (β_{1j})

BLIMP SCRIPT 5.3

```
DATA: Employee.dat;
VARIABLES: Employee Team Turnover Male Empower LMX JobSat Climate TeamPerf;
NOMINAL: Male;    # automatic dummy coding for nominal predictors
CLUSTERID: Team;
CENTER: groupmean = LMX; grandmean = LMX.mean Climate;
MODEL: Empower ~ intercept LMX Male LMX.mean Climate | intercept LMX;
BURN: 10000;
ITERATIONS: 10000;
SEED: 90291;
```

RBLIMP SCRIPT 5 (MODEL 3)

```
model3 <- rblimp(
    data = Employee,
    nominal = 'Male',
    clusterid = 'Team',
    center = 'groupmean = LMX; grandmean = LMX.mean Climate',
    model = 'Empower ~ intercept LMX Male LMX.mean Climate | intercept LMX',
    seed = 90291,
    burn = 10000,
    iter = 10000)
output(model3)
```

BLIMP OUTPUT

= level-2 estimate
= level-1 estimate
= combined estimate

```
Outcome Variable:  Empower

Grand Mean Centered: Climate LMX.mean[Team]
Group Mean Centered: LMX

Parameters                     Estimate     StdDev       2.5%      97.5%      ChiSq     PValue      N_Eff
                              ---------------------------------------------------------------------------
Variances:
  L2 : Var(Intercept)             1.077      0.563      0.252      2.412        ---        ---    197.283
  L2 : Cov(LMX,Intercept)        -0.172      0.147     -0.489      0.097        ---        ---    324.463
  L2 : Var(LMX)                   0.188      0.076      0.064      0.364        ---        ---    390.430
  Residual Var.                  12.288      0.824     10.809     14.021        ---        ---   1868.918

Coefficients:
  Intercept                      27.687      0.268     27.161     28.222  10663.961      0.000   2099.378
  LMX                             0.638      0.072      0.496      0.775     77.733      0.000   2189.207
  Male.1                          1.893      0.303      1.296      2.488     38.980      0.000   7340.383
  LMX.mean[Team]                 -0.154      0.297     -0.799      0.361      0.330      0.566    360.867
  Climate                         0.291      0.052      0.186      0.392     30.822      0.000   1256.661

Standard Deviations:
  L2 : SD(Intercept)              1.038      0.267      0.502      1.553        ---        ---    162.850
  L2 : Cor(LMX,Intercept)        -0.419      0.307     -0.942      0.245        ---        ---    179.032
  L2 : SD(LMX)                    0.433      0.088      0.252      0.603        ---        ---    345.870
  Residual SD                     3.505      0.117      3.288      3.745        ---        ---   1865.490

...
                              ---------------------------------------------------------------------------
```

FIXED EFFECT INTERPRETATIONS

- β_0 = 27.69 is the empowerment mean for the male = 0 group (because all numeric predictors are centered)

- β_1 = 0.64 is the expected empowerment difference between two LMX scores from the same team that differ by one point, controlling for sex

- β_2 = 1.89 is the expected empowerment difference for males (relative to females), controlling for within-team LMX

FIXED EFFECTS, CONTINUED

- $\beta_3 = -0.15$ is the expected empowerment difference between two teams whose average LMX ratings differ by one point, controlling for team-level climate

- $\beta_4 = 0.29$ is the expected empowerment difference between two teams whose climate ratings differ by one point, controlling for team-average LMX

RANDOM EFFECT INTERPRETATIONS

- $u_{0j} = \beta_{0j} - (\beta_0 + \beta_3(LMX_j^b) + \beta_4(climate_j))$
- $var(u_{0j}) = 1.08$ is the average squared distance between the level-2 empowerment means and their predicted values
- $sd(u_{0j}) = 1.04$ is the average distance between the level-2 empowerment means and their predicted values

- $u_{1j} = \beta_{1j} - \beta_1$
- $var(u_{1j}) = 0.19$ is the average squared distance between the team-specific LMX slopes and the average slope
- $sd(u_{1j}) = 0.43$ is the average distance between the team-specific LMX slopes and the average slope

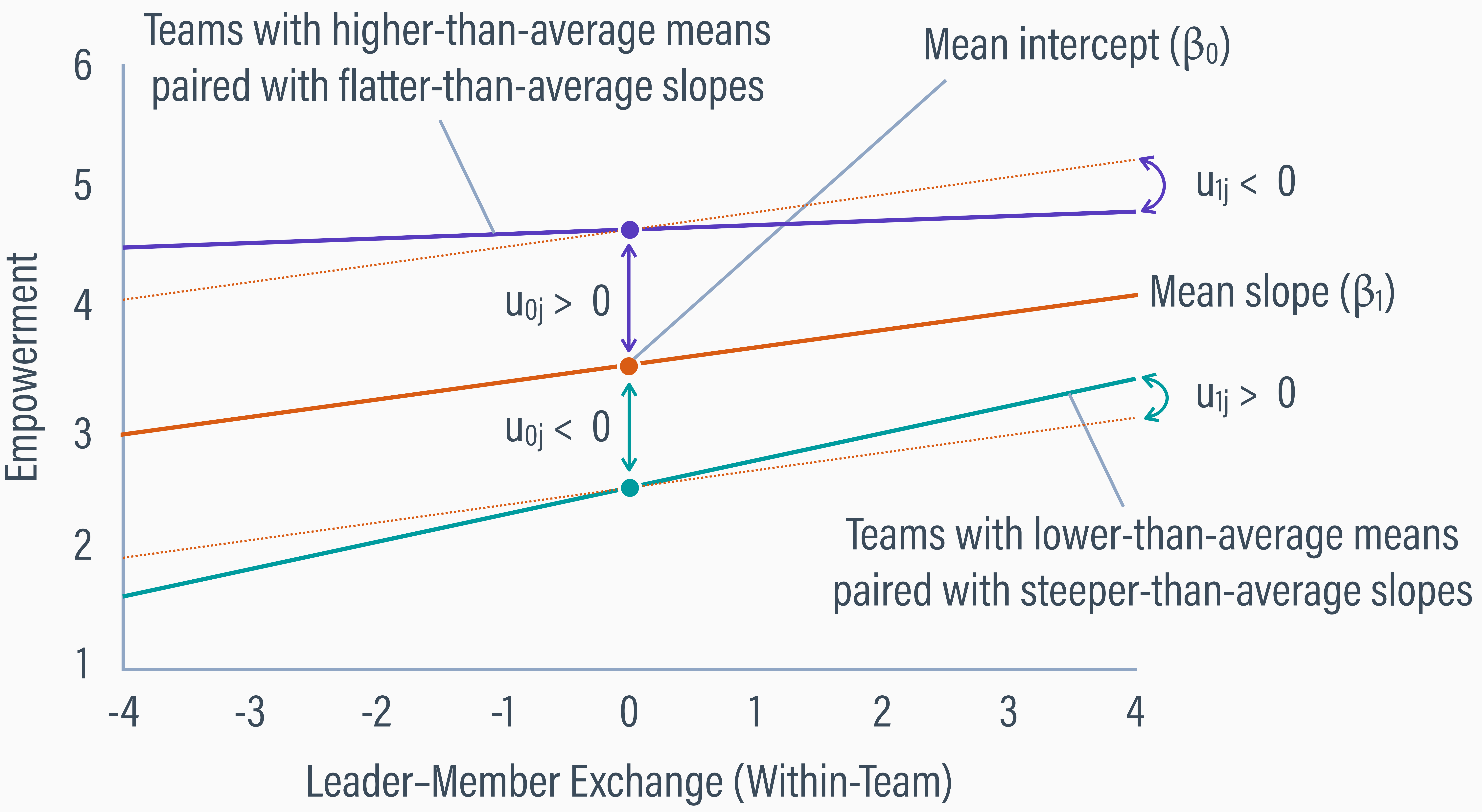
RANDOM EFFECT CORRELATION

- The random intercepts (team-specific means) and random slopes (team-specific LMX effects) are correlated

- The unstandardized association is $cov(u_{0j}, u_{1j}) = -0.17$, and the linear correlation is $cor(u_{0j}, u_{1j}) = -.42$

- A positive correlation implies that teams with higher-than-average means (high u_{0j}) tend to have flatter-than-average slopes (low u_{1j}) and vice versa (low u_{0j} paired with high u_{1j})

NEGATIVE CORRELATION GRAPHIC



WITHIN-CLUSTER RESIDUAL VARIANCE

- $\varepsilon_{ij} = \text{empower}_{ij} - (\ \beta_{0j} + \beta_1(\text{lmx}_{ij}^{W}) + \beta_2(\text{male}_{ij})\ )$
- $var(\varepsilon_{ij}) = 12.29$ is the average squared distance between the level-1 empowerment scores and their predicted values
- $sd(\varepsilon_{ij}) = 3.51$ is the average distance between the level-1 empowerment scores and their predicted values

OUTLINE

1. Modeling Step 1: Estimate ICCs
2. Modeling Step 2: Random Intercept Model
3. Modeling Step 3: Random Slope Model 1
4. Deciding Whether a Random Slope Is Needed
5. Modeling Step 3: Random Slope Model 2
6. Checking MLM Assumptions
7. Latent Variable Specification

ARE RANDOM SLOPES NEEDED?

- Ignoring importance sources of random slope variation can bias the fixed effect coefficients
- However, include random slopes judiciously because overfitting impacts convergence, precision, and power
- Auditioning random slopes one predictor at a time is often an effective strategy, keeping those with support from the data

THREE IMPERFECT CRITERIA

- MCMC diagnostics (subpar values imply overfitting)
- Rights and Sterba (2019) R^2 effect size for random slopes
- Specialized χ^2 test (called a chi-bar test) for nested models involving variance parameters

PSR DIAGNOSTIC OUTPUT

Quality control check: PSRF diagnostics all < 1.05 well before the end of the burn-in period

```
BURN-IN POTENTIAL SCALE REDUCTION (PSR) OUTPUT:

  NOTE: Split chain PSR is being used. This splits each chain's
        iterations to create twice as many chains.

   Comparing iterations across 2 chains     Highest PSR   Parameter #
                          251 to 500           2.805            10
                          501 to 1000          1.183            11
                          751 to 1500          1.090            11
                         1001 to 2000          1.017             5
                                  ...            ...            ..
                         2501 to 5000          1.054            11
                         2751 to 5500          1.018            11
                         3001 to 6000          1.029            11
                         3251 to 6500          1.020            11
                         3501 to 7000          1.026            11
                         3751 to 7500          1.007            29
                         4001 to 8000          1.009            10
                         4251 to 8500          1.027            11
                         4501 to 9000          1.027            11
                         4751 to 9500          1.018            11
                         5001 to 10000         1.020            11
```

BLIMP OUTPUT

Quality control check: All effective sample size (N_Eff) diagnostics > 100

```
Outcome Variable:  Empower

Grand Mean Centered: Climate LMX.mean[Team]
Group Mean Centered: LMX

Parameters                  Estimate    StdDev      2.5%     97.5%      ChiSq    PValue     N_Eff
                            -------------------------------------------------------------------
Variances:
  L2 : Var(Intercept)          1.077     0.563     0.252     2.412        ---       ---   197.283
  L2 : Cov(LMX,Intercept)     -0.172     0.147    -0.489     0.097        ---       ---   324.463
  L2 : Var(LMX)                0.188     0.076     0.064     0.364        ---       ---   390.430
  Residual Var.               12.288     0.824    10.809    14.021        ---       ---  1868.918

Coefficients:
  Intercept                   27.687     0.268    27.161    28.222  10663.961     0.000  2099.378
  LMX                          0.638     0.072     0.496     0.775     77.733     0.000  2189.207
  Male.1                       1.893     0.303     1.296     2.488     38.980     0.000  7340.383
  LMX.mean[Team]              -0.154     0.297    -0.799     0.361      0.330     0.566   360.867
  Climate                      0.291     0.052     0.186     0.392     30.822     0.000  1256.661

Standard Deviations:
  L2 : SD(Intercept)           1.038     0.267     0.502     1.553        ---       ---   162.850
  L2 : Cor(LMX,Intercept)     -0.419     0.307    -0.942     0.245        ---       ---   179.032
  L2 : SD(LMX)                 0.433     0.088     0.252     0.603        ---       ---   345.870
  Residual SD                  3.505     0.117     3.288     3.745        ---       ---  1865.490

...
                            -------------------------------------------------------------------
```

MULTILEVEL EFFECT SIZES

- Numerous effect sizes exist for multilevel models
- Rights and Sterba (2019, Psychological Methods) provides a unifying framework that subsumes different approaches
- They assign an R^2 effect size to the random slopes, which explain within-cluster variation by reducing residual distances to the observed scores

PREDICTED OUTCOME VARIATION

- The model-predicted total variance is the sum of (a) variance explained by predictors at both levels, (b) residual random-intercept variation, (c) within-cluster variation due to random slopes, and (d) residual within-cluster variation

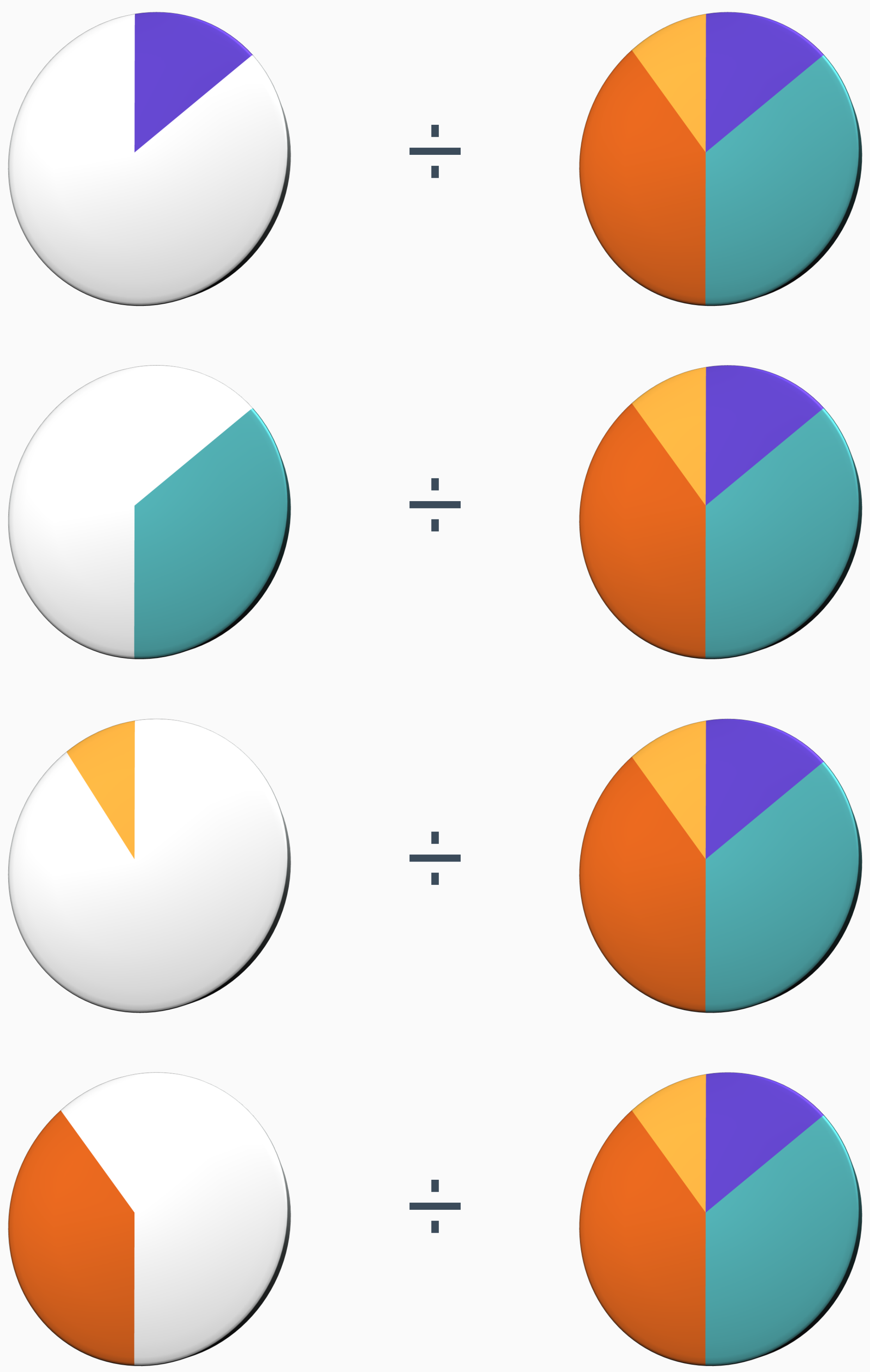
$$\sigma_Y^2 = \boldsymbol{\beta}^T \boldsymbol{\Sigma}_X \boldsymbol{\beta} + \mathrm{var}\left(u_{0j}\right) + \mathrm{tr}(\boldsymbol{\Sigma}_u \boldsymbol{\Sigma}_X) + \sigma_\varepsilon^2$$

total variation = explained by predictors + intercept residuals + slope residuals + level-1 residuals

VARIANCE EXPLAINED MEASURES

Explained ÷ Total

Fixed effects of predictors	$R^2_{predictors} = \dfrac{\boldsymbol{\beta}^T \boldsymbol{\Sigma}_X \boldsymbol{\beta}}{\sigma^2_Y}$	
Level-2 random intercept residuals	$R^2_{residual\text{-}between} = \dfrac{\sigma^2_{u_0}}{\sigma^2_Y}$	
Level-1 via random slope residuals	$R^2_{slopes} = \dfrac{tr(\boldsymbol{\Sigma}_u \boldsymbol{\Sigma}_X)}{\sigma^2_Y}$	
Level-1 within-cluster residuals	$R^2_{residual\text{-}within} = \dfrac{\sigma^2_\varepsilon}{\sigma^2_Y}$	



BLIMP OUTPUT

= level-2 estimate
= level-1 estimate
= combined estimate

```
Outcome Variable:  Empower

Grand Mean Centered: Climate LMX.mean[Team]
Group Mean Centered: LMX
```

Parameters	Estimate	StdDev	2.5%	97.5%	ChiSq	PValue	N_Eff
...							
Standardized Coefficients:							
LMX	0.393	0.040	0.310	0.466	95.779	0.000	1718.136
Male.1	0.206	0.032	0.142	0.269	40.841	0.000	6764.759
LMX.mean[Team]	-0.034	0.060	-0.151	0.083	0.323	0.570	437.813
Climate	0.255	0.044	0.164	0.339	33.112	0.000	1199.975
Proportion Variance Explained							
by Coefficients	0.289	0.035	0.220	0.359	---	---	1349.110
by Level-2 Random Intercepts	0.051	0.025	0.012	0.109	---	---	184.323
by Level-2 Random Slopes	0.071	0.028	0.024	0.133	---	---	379.856
by Level-1 Residual Variation	0.584	0.039	0.508	0.660	---	---	872.097

EFFECT SIZE INTERPRETATIONS

- $R^2_{(predictors)}$ = .29 is the proportion of the total variation explained by predictors at both levels
- $R^2_{(residual\text{-}between)}$ = .05 is the proportion of the total variation attributable to the random intercept residuals (the u_{0j} terms)
- $R^2_{(slopes)}$ = .07 is the proportion of the total variation attributable to the random slope residuals (the u_{1j} terms)
- $R^2_{(residual\text{-}within)}$ = .58 is the proportion of the total variation attributable to the within-cluster residuals (the ε_{ij} terms)

RANDOM SLOPE R^2 BENCHMARKS

- Random slope R^2 values tend to be quite small (values > .10 appear to be uncommon)
- Across data sets from numerous MLM and longitudinal books, random slope R^2 values ranged from .001 to .094 (Enders, Woller, & Keller, 2023)
- The mean R^2 from cross-section applications was about .04, and the mean from longitudinal applications was about .05

WALD χ^2 STATISTIC

- The Wald χ^2 test equals the sum of squared standardized differences between a set of point estimates and their null values (like a squared z-statistic for multiple parameters)

$$\chi^2_{MCMC} = (\boldsymbol{\theta} - \boldsymbol{\theta}_0)\text{cov}(\boldsymbol{\theta})^{-1}(\boldsymbol{\theta} - \boldsymbol{\theta}_0)$$

- $\boldsymbol{\theta}$ contains multiple point estimates, $\boldsymbol{\theta}_0$ contains the correspond null values (typically 0s), and cov($\boldsymbol{\theta}$) contains the variances and covariances of the estimates across MCMC cycles

SIGNIFICANCE TESTING WITH VARIANCES

- The random slope adds two parameters to the model (a variance and correlation), and both must be tested in $\boldsymbol{\theta}$
- When the true variance is zero, estimates accumulate at the boundary rather than fluctuating around it, producing a nonregular sampling distribution
- Significance testing requires a $\bar{\chi}^2$ (chi-bar) reference distribution, defined as a weighted mixture of chi-square distributions with different degrees of freedom

RBLIMP SCRIPT 5 (MODEL 3)

```
model3 <- rblimp(
        data = Employee,
        nominal = 'Male',
        clusterid = 'Team',
        center = 'groupmean = LMX; grandmean = LMX.mean Climate',
        model = 'Empower ~ intercept LMX Male LMX.mean Climate | intercept LMX',
        seed = 90291,
        burn = 10000,
        iter = 10000)
output(model3)
chibar_test(model = model3, DV = 'Empower', IV = 'LMX')
```

OUTPUT

```
-----------------------------------------------------------------
Wald Chi-Square Statistic:
-----------------------------------------------------------------

  χ² = 6.618

-----------------------------------------------------------------
Chi-Bar Test Results:
________________________________

Method                        ChiSq    dfWeights             PValue    Notes
Standard chi-square           6.618    df = 2                0.037     Ignores boundary issue
Stram & Lee (1994)            6.618    (0.50, 0.25, 0.25)    0.012     Assumes independence (rho=0)
Self & Liang (1987) Case 6    6.618    (0.50, 0.12, 0.38)    0.015     Uses rho=-0.362, theta=0.382

-----------------------------------------------------------------
```

CHI-BAR TEST INTERPRETATION

- The chi-bar test refutes the null hypothesis that the population slope variance and its covariance equal zero
- $\bar{\chi}^2 = 6.62$, $p = .015$
- Adding the LMX random slope improves fit

SUMMARY

- Including a random slope for LMX appears justified based on all three criteria (which won't always agree)
- MCMC diagnostics revealed no estimation issues
- The random slope R^2 effect size was salient
- The chi-bar χ^2 test was significant

OUTLINE

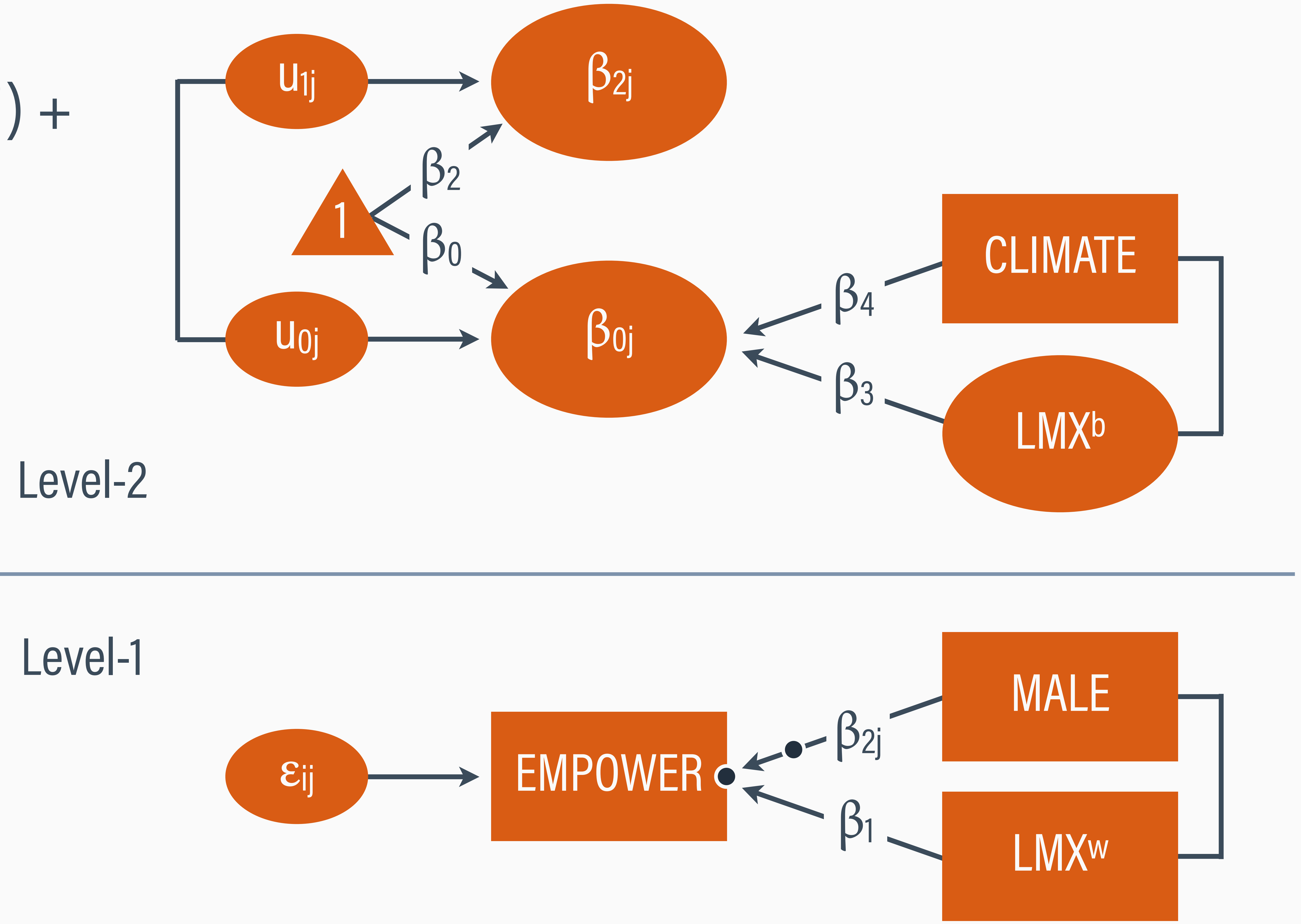
1. Modeling Step 1: Estimate ICCs
2. Modeling Step 2: Random Intercept Model
3. Modeling Step 3: Random Slope Model 1
4. Deciding Whether a Random Slope Is Needed
5. Modeling Step 3: Random Slope Model 2
6. Checking MLM Assumptions
7. Latent Variable Specification

EVALUATING ADDITIONAL SLOPES

- Any level-1 predictor could require a random slope
- Auditioning random slopes one predictor at a time is often an effective strategy, keeping those with support from the data
- Is a random slope for the sex dummy code warranted?

PATH DIAGRAM

$$empower_{ij} = (\beta_0 + u_{0j}) + \beta_1(lmx^{w}_{ij}) + (\beta_2 + u_{2j})(male_{ij}) + \beta_3(lmx^{b}_{j}) + \beta_4(climate_j) + \varepsilon_{ij}$$



• = random intercept (β_{0j}) and random slope (β_{2j})

In the previous model, the fixed effect for sex was $\beta_2 = 1.89$, representing the expected difference in empowerment for males relative to females. In small groups of two or three, discuss the meaning of a random slope for a binary predictor. Provide a conceptual explanation of this parameter that a colleague with no prior exposure to multilevel modeling could understand.

BLIMP SCRIPT 5.4

```
DATA: Employee.dat;
VARIABLES: Employee Team Turnover Male Empower LMX JobSat Climate TeamPerf;
NOMINAL: Male;    # automatic dummy coding for nominal predictors
CLUSTERID: Team;
CENTER: groupmean = LMX; grandmean = LMX.mean Climate;
MODEL: Empower ~ intercept LMX Male LMX.mean Climate | intercept Male;
BURN: 10000;
ITERATIONS: 10000;
SEED: 90291;
```

RBLIMP SCRIPT 5 (MODEL 4)

```
model4 <- rblimp(
    data = Employee,
    nominal = 'Male',
    clusterid = 'Team',
    center = 'groupmean = LMX; grandmean = LMX.mean Climate',
    model = 'Empower ~ intercept LMX Male LMX.mean Climate | intercept Male',
    seed = 90291,
    burn = 10000,
    iter = 10000)
output(model4)
chibar_test(model = model4, DV = 'Empower', IV = ‘Male')
```

PSR DIAGNOSTIC OUTPUT

Quality control check: PSRF diagnostics all < 1.05 well before the end of the burn-in period

```
BURN-IN POTENTIAL SCALE REDUCTION (PSR) OUTPUT:

  NOTE:  Split chain PSR is being used. This splits each chain's
         iterations to create twice as many chains.

    Comparing iterations across 2 chains     Highest PSR   Parameter #
                             251  to  500          1.420           10
                             501  to  1000         1.168           11
                             751  to  1500         1.194           12
                            1001  to  2000         1.131           12
                                      ...            ...           ..
                            2501  to  5000         1.026           12
                            2751  to  5500         1.026           12
                            3001  to  6000         1.022           10
                            3251  to  6500         1.021           12
                            3501  to  7000         1.020           12
                            3751  to  7500         1.021           10
                            4001  to  8000         1.013           19
                            4251  to  8500         1.016           10
                            4501  to  9000         1.010           29
                            4751  to  9500         1.008           11
                            5001  to  10000        1.021           12
```

BLIMP OUTPUT

Quality control check: One effective sample size (N_Eff) diagnostic is a bit low (< 100), suggesting that more iterations are needed

```
Outcome Variable:  Empower

Grand Mean Centered: Climate LMX.mean[Team]
Group Mean Centered: LMX

Parameters                         Estimate     StdDev       2.5%      97.5%      ChiSq     PValue      N_Eff
                               -------------------------------------------------------------------------------
Variances:
  L2 : Var(Intercept)                 1.417      0.833      0.291      3.471        ---        ---    234.195
  L2 : Cov(Male.1,Intercept)         -0.884      0.957     -3.275      0.376        ---        ---    169.972
  L2 : Var(Male.1)                    2.345      1.661      0.232      6.448        ---        ---    163.161
  Residual Var.                      13.006      0.848     11.458     14.778        ---        ---   1438.068

Coefficients:
  Intercept                          27.685      0.275     27.140     28.230  10101.708      0.000   1818.877
  LMX                                 0.676      0.057      0.564      0.787    141.650      0.000   2433.306
  Male.1                              1.889      0.342      1.218      2.544     30.458      0.000   5525.504
  LMX.mean[Team]                     -0.230      0.314     -0.953      0.302      0.645      0.422    286.900
  Climate                             0.281      0.052      0.179      0.384     28.725      0.000   1545.496

Standard Deviations:
  L2 : SD(Intercept)                  1.191      0.339      0.540      1.863        ---        ---    183.668
  L2 : Cor(Male.1,Intercept)         -0.553      0.376     -0.937      0.573        ---        ---     79.993
  L2 : SD(Male.1)                     1.531      0.532      0.482      2.539        ---        ---    127.703
  Residual SD                         3.606      0.117      3.385      3.844        ---        ---   1424.495

...
                               -------------------------------------------------------------------------------
```

BLIMP OUTPUT

= level-2 estimate
= level-1 estimate
= combined estimate

```
Outcome Variable:  Empower

Grand Mean Centered: Climate LMX.mean[Team]
Group Mean Centered: LMX

Parameters                       Estimate     StdDev       2.5%      97.5%       ChiSq      PValue      N_Eff
                              -------------------------------------------------------------------------------
Variances:
  L2 : Var(Intercept)               1.417      0.833      0.291      3.471         ---         ---    234.195
  L2 : Cov(Male.1,Intercept)       -0.884      0.957     -3.275      0.376         ---         ---    169.972
  L2 : Var(Male.1)                  2.345      1.661      0.232      6.448         ---         ---    163.161
  Residual Var.                    13.006      0.848     11.458     14.778         ---         ---   1438.068

Coefficients:
  Intercept                        27.685      0.275     27.140     28.230   10101.708       0.000   1818.877
  LMX                               0.676      0.057      0.564      0.787     141.650       0.000   2433.306
  Male.1                            1.889      0.342      1.218      2.544      30.458       0.000   5525.504
  LMX.mean[Team]                   -0.230      0.314     -0.953      0.302       0.645       0.422    286.900
  Climate                           0.281      0.052      0.179      0.384      28.725       0.000   1545.496

Standard Deviations:
  L2 : SD(Intercept)                1.191      0.339      0.540      1.863         ---         ---    183.668
  L2 : Cor(Male.1,Intercept)       -0.553      0.376     -0.937      0.573         ---         ---     79.993
  L2 : SD(Male.1)                   1.531      0.532      0.482      2.539         ---         ---    127.703
  Residual SD                       3.606      0.117      3.385      3.844         ---         ---   1424.495

...
                              -------------------------------------------------------------------------------
```

BLIMP OUTPUT, CONTINUED

= level-2 estimate
= level-1 estimate
= combined estimate

```
Outcome Variable:  Empower

Grand Mean Centered: Climate LMX.mean[Team]
Group Mean Centered: LMX


Parameters                        Estimate    StdDev      2.5%     97.5%     ChiSq    PValue     N_Eff
                                -------------------------------------------------------------------------------
...

Standardized Coefficients:
  LMX                                0.414     0.032     0.350     0.473   171.102     0.000  1381.297
  Male.1                             0.205     0.036     0.133     0.274    32.230     0.000  5423.973
  LMX.mean[Team]                    -0.053     0.064    -0.178     0.075     0.672     0.412   340.039
  Climate                            0.245     0.044     0.158     0.330    30.727     0.000  1479.362

Proportion Variance Explained
  by Coefficients                    0.300     0.032     0.238     0.364       ---       ---  1166.171
  by Level-2 Random Intercepts       0.051     0.026     0.012     0.112       ---       ---   231.330
  by Level-2 Random Slopes           0.028     0.019     0.003     0.075       ---       ---   158.762
  by Level-1 Residual Variation      0.615     0.036     0.543     0.684       ---       ---   634.902

                                -------------------------------------------------------------------------------
```

OUTPUT

```
----------------------------------------------------------------
Wald Chi-Square Statistic:
----------------------------------------------------------------

  χ² = 2.482

----------------------------------------------------------------
Chi-Bar Test Results:
----------------------------------------------------------------

Method                         ChiSq    dfWeights            PValue    Notes
Standard chi-square            2.482    df = 2               0.289     Ignores boundary issue
Stram & Lee (1994)             2.482    (0.50, 0.25, 0.25)   0.101     Assumes independence (rho=0)
Self & Liang (1987) Case 6     2.482    (0.50, 0.27, 0.23)   0.098     Uses rho=-0.740, theta=0.235

----------------------------------------------------------------
```

In small groups of two or three, discuss whether adding a random slope for sex is warranted. Consider the MCMC diagnostics, the R^2 effect size, and the chi-bar significance test.

OUTLINE

1. Modeling Step 1: Estimate ICCs
2. Modeling Step 2: Random Intercept Model
3. Modeling Step 3: Random Slope Model 1
4. Deciding Whether a Random Slope Is Needed
5. Modeling Step 3: Random Slope Model 2
6. Checking MLM Assumptions
7. Latent Variable Specification

MLM ASSUMPTIONS

- Associations are linear
- Level-1 residuals are normal with constant variation across level-1 units (days) and level-2 units (persons)
- Level-2 random intercept residuals are normal with constant variation across level-2 units (persons)

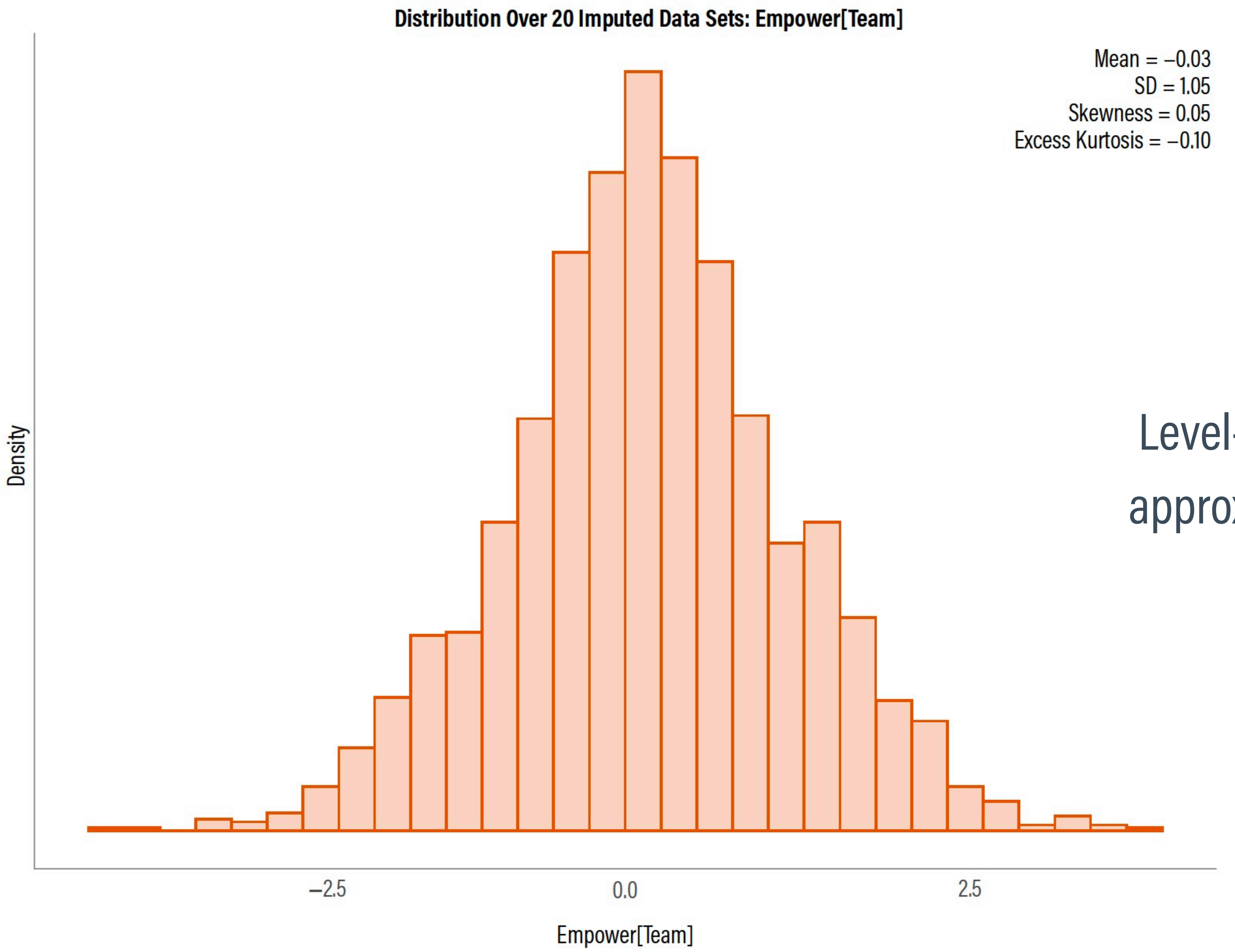
RESIDUAL DIAGNOSTICS

- Each MCMC iteration produces different parameter values—and thus different predicted values and residuals
- Residuals from different MCMC steps can be saved to create multiple datasets for plotting
- The univariate_plot function (imported at the top of the rblimp scripts) graphs residuals from multiple rblimp data sets
- See Module 3 slides for additional examples

RBLIMP SCRIPT 5 (MODEL 6)

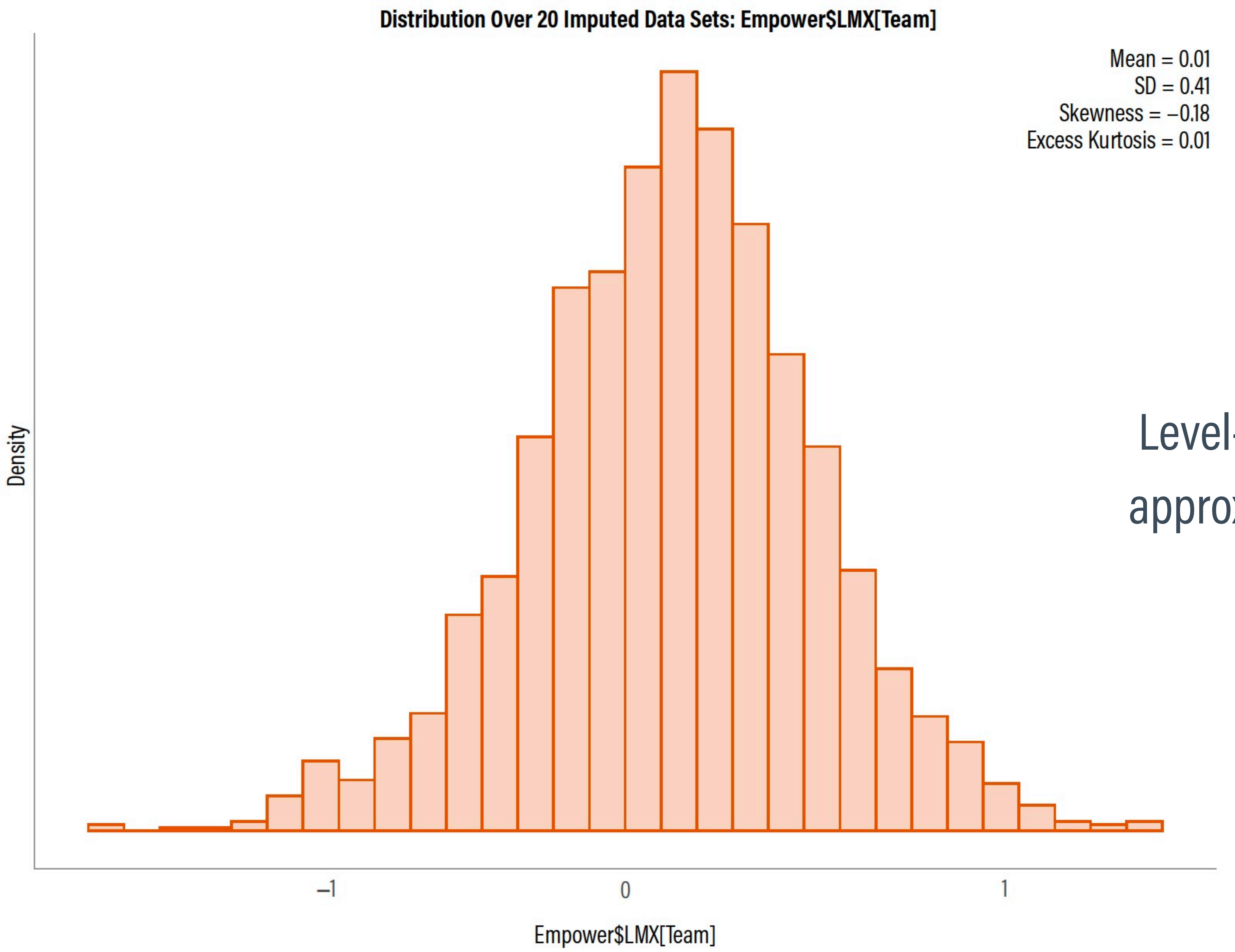
```
model6 <- rblimp(
    data = Employee,
    nominal = 'Male',
    clusterid = 'Team',
    center = 'groupmean = LMX; grandmean = LMX.mean Climate',
    model = 'Empower ~ intercept LMX Male LMX.mean Climate | intercept LMX',
    seed = 90291,
    burn = 10000,
    iter = 10000,
    nimps = 20)
output(model6)
univariate_plot(vars = c('Empower[Team]','Empower$LMX[Team]','Empower.residual'),
  model = model6, stats = T)
```

LEVEL-2 RANDOM INTERCEPT RESIDUALS



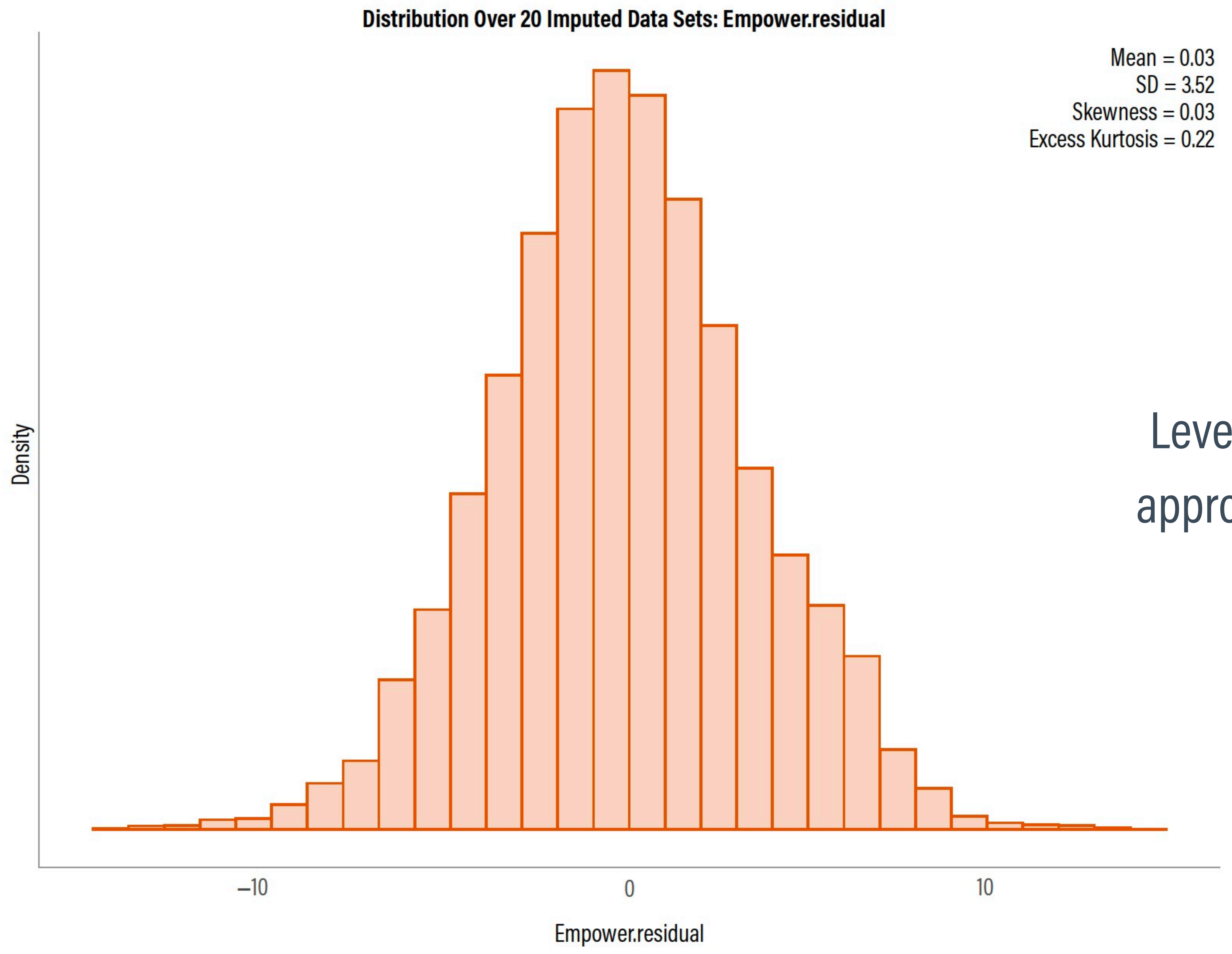
Level-2 residuals are approximately normal

LEVEL-2 RANDOM SLOPE RESIDUALS



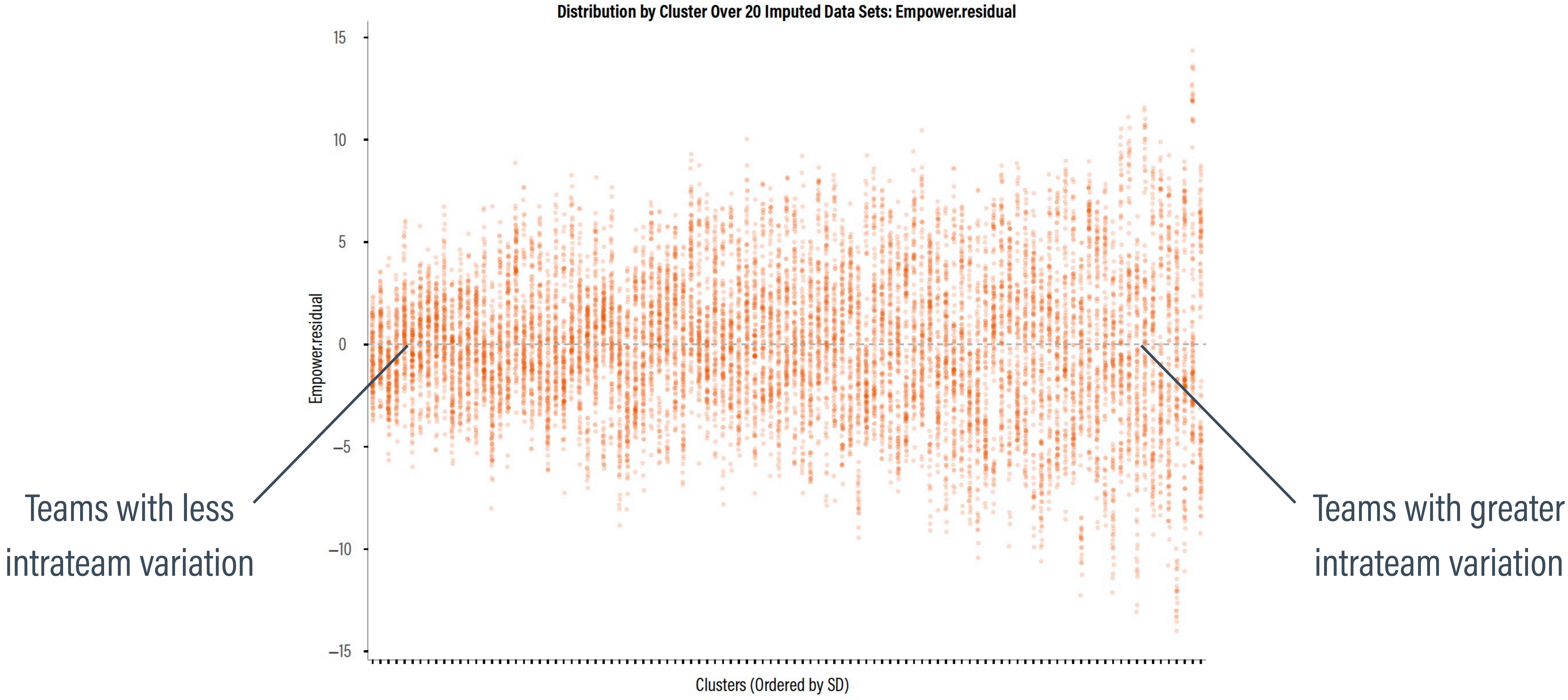
Level-2 residuals are approximately normal

LEVEL-1 RESIDUALS



Level-1 residuals are approximately normal

LEVEL-1 RESIDUALS BY CLUSTER



OUTLINE

1 Modeling Step 1: Estimate ICCs
2 Modeling Step 2: Random Intercept Model
3 Modeling Step 3: Random Slope Model 1
4 Deciding Whether a Random Slope Is Needed
5 Modeling Step 3: Random Slope Model 2
6 Checking MLM Assumptions
7 Latent Variable Specification

BLIMP SCRIPT 5.5

```
DATA: Employee.dat;
VARIABLES: Employee Team Turnover Male Empower LMX JobSat Climate TeamPerf;
NOMINAL: Male;
CLUSTERID: Team;
LATENT: Team = beta0j beta1j;    # define random intercept and slope
CENTER: groupmean = LMX; grandmean = LMX.mean Climate;
MODEL: beta0j ~ intercept LMX.mean Climate;
beta1j ~ intercept;
beta0j ~~ beta1j;    # correlate random intercepts and slopes
Empower ~ intercept@beta0j LMX@beta1j Male;
BURN: 10000;
ITERATIONS: 10000;
SEED: 90291;
```

RBLIMP SCRIPT 5 (MODEL 7)

```
model7 <- rblimp(
        data = Employee,
        nominal = 'Male',
        clusterid = 'Team',
        latent = 'Team = beta0j beta1j',
        center = 'groupmean = LMX; grandmean = LMX.mean Climate;',
        model = '
          beta0j ~ intercept LMX.mean Climate;
          beta1j ~ intercept;
          beta0j ~~ beta1j;    # correlate random intercepts and slopes
          Empower ~ intercept@beta0j LMX@beta1j Male;',
        seed = 90291,
        burn = 10000,
        iter = 10000)
output(model7)
```