Between and Within-Subject Analysis

Kathy Welch
University of Michigan
Department of Biostatistics
For SAS Users Group in Michigan, 2019

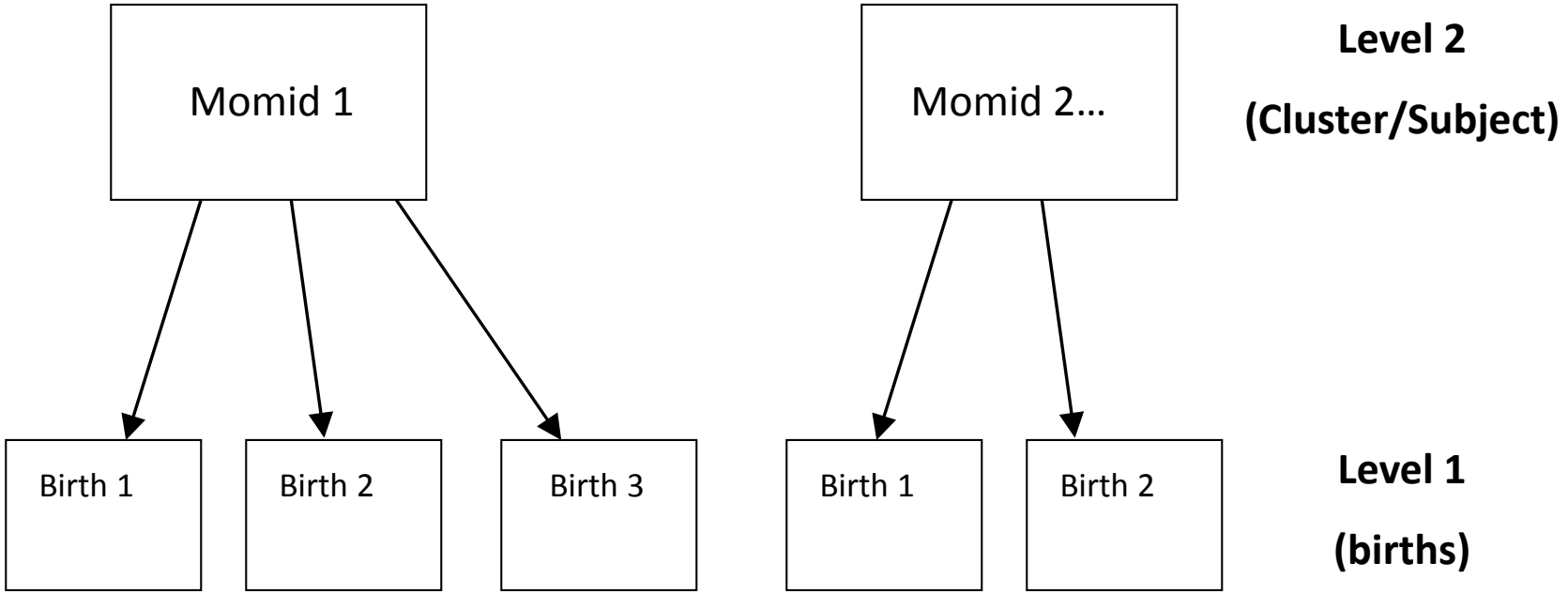
Introduction

- The analysis of **multilevel data** (clustered and longitudinal data) present challenges in accurately estimating the between-subject and within-subject effects of covariates.
- We illustrate an approach to getting consistent estimates of both the between and within-subjects effects of covariates.

Example: Birth Weight Data (Abrevaya)

- Data collected between 1990 and 1998
- 8604 singleton births
- 3978 moms
  - 3330 mom (84%) had two births
  - 648 moms (16%) had three births
- Want to estimate the impact of mother smoking on baby birth weights after adjusting for covariates.
- Example based on Rabe-Hesketh and Skrondal (2012)

Abrevaya Birthweight Data is Multilevel



- **Between-Subject (Level 2) Variables:** married, education level, black
- **Within-Subject (Level 1) Variables:** birthweight, mother's age, smoking status, sex, level of prenatal care, trimester of first prenatal visit

Between-Subject (Level 2) Variables

Do not vary within a subject/cluster

- **Married** indicator for marital status at baseline
- Mom education indicators

  (Less than HS reference)

  – **HSGrad**
  – **SomeColl**
  – **CollGrad**
- **Black** indicator for race/ethnicity

Descriptives for Between-Subject Variables

Variable	N	Mean	Std Dev	Minimum	Maximum
momid	3978	35878.89	23524.87	14.00	109039.00
married	3978	0.89	0.32	0.00	1.00
hsgrad	3978	0.29	0.45	0.00	1.00
somecoll	3978	0.24	0.42	0.00	1.00
collgrad	3978	0.36	0.48	0.00	1.00
black	3978	0.07	0.26	0.00	1.00

Within-Subject (Level 1) Variables

Can Vary within a subject/cluster

- **Birwt**: Birth weight in grams (outcome)
- **Mage**: Mother's age at birth of child
- **Smoke**: Smoked during pregnancy indicator
- **Male**: baby sex indicator
- Kessner score indicators

  (Kessner1: reference Level)
  - **Kessner2:** intermediate prenatal care
  - **Kessner3**: inadequate prenatal care
- Trimester of first prenatal visit Indicators
  - **Novisit**: no prenatal visits

    (Pretri1: reference level)
  - **Pretri2**: first visit in 2nd trimester
  - **Pretri3**: first visit in 3rd trimester

Descriptives for Within-Subject Variables

Variable	N	Mean	Std Dev	Minimum	Maximum
smoke	8604	0.14	0.35	0.00	1.00
male	8604	0.51	0.50	0.00	1.00
mage	8604	28.59	5.49	14.00	45.00
kessner2	8604	0.18	0.39	0.00	1.00
kessner3	8604	0.05	0.22	0.00	1.00
novisit	8604	0.01	0.09	0.00	1.00
pretri2	8604	0.13	0.34	0.00	1.00
pretri3	8604	0.02	0.15	0.00	1.00

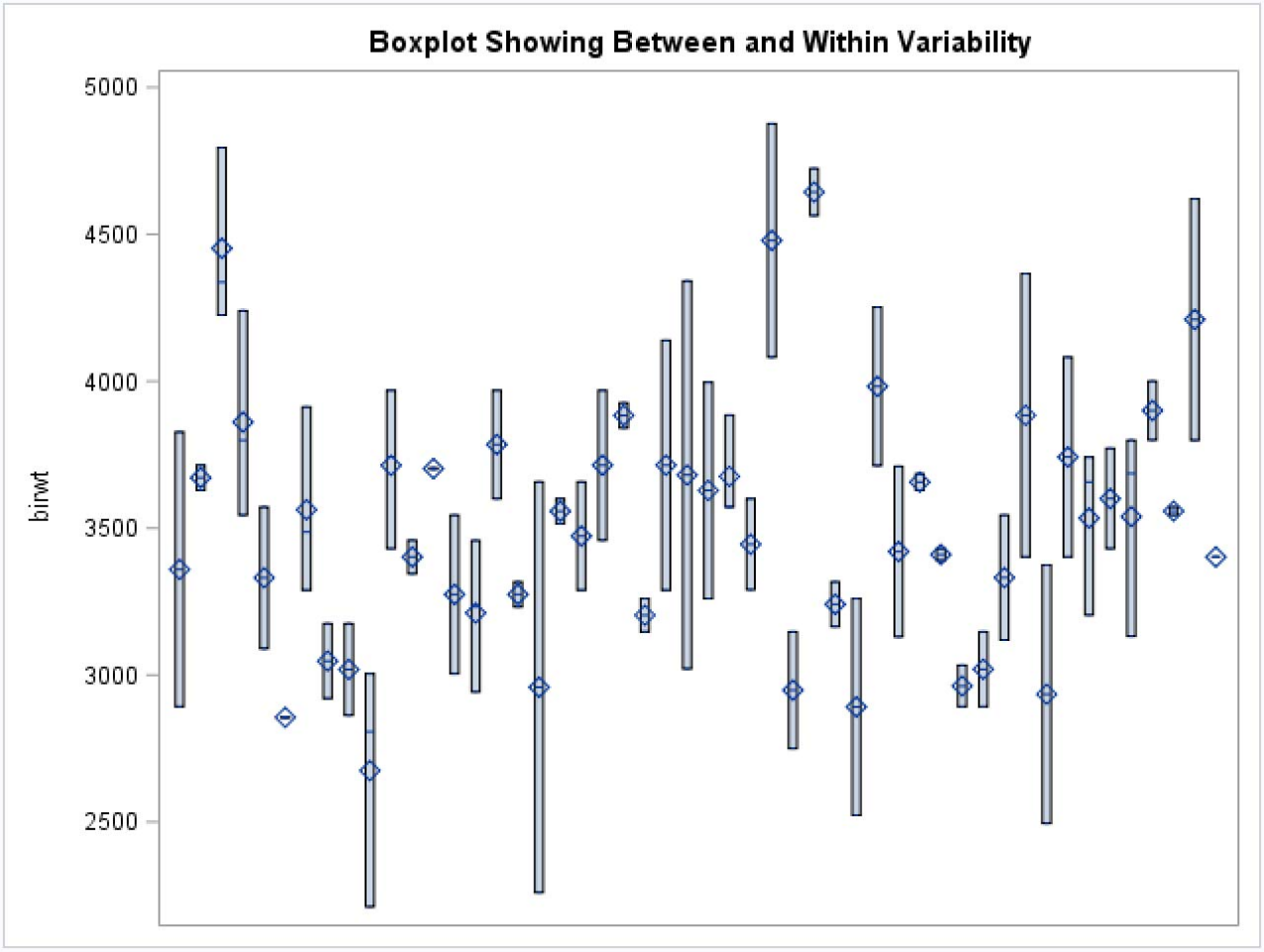
Note on Effects of Smoking (a Within-Subject Variable)

- **Smoking is a within-subject variable** because it can change within the same mom from pregnancy to pregnancy.

But…

- There are different possible effects of smoking
  - Effect of smoking on birthweights for babies from the same mother—comparing birthweights for pregnancies when she smoked to pregnancies when she didn't
  - Effect of smoking on birthweights for babies of different mothers who smoked vs. those who did not smoke

Graph Showing Between and Within-Subject Variability in Birthweights



SAS Code to Generate Graph

```
/*Generate Between-Subjects Data Set*/
proc sort data=smoking;
  by momid;
run;
title "Between-Subjects Variables";
data smoking_level2;
  set smoking;
  by momid;
  if first.momid;
  keep momid hsgrad somecoll collgrad married black;
run;
proc means data=smoking_level2;
run;
```

SAS Code to Generate Graph (Cont)

```
/*Select a Random Sample of 50 Subjects*/
proc surveyselect data=smoking_level2
      out=sampledat
      method=srs
      sampsize=50
      outall
      seed=442291000;
run;

data smoking_graph;
  merge smoking sampledat;
  by momid;
run;
```

SAS Code to Generate Graph (Cont)

```
/*Generate Boxplot*/
title "Boxplot Showing Between and Within Variability";
proc sgplot data=smoking_graph noautolegend;
   where selected=1;
   vbox birwt / category=momid;
   xaxis display=none;
run;
```

First Fit a Linear Mixed Model (LMM): Random Intercept for Each Subject

$$\text{Birwt}_{ij} = \beta_1 \text{smoke}_{ij} + \beta_2 \text{male}_{ij} + \beta_3 \text{mage}_{ij} + \beta_4 \text{kessner2}_{ij}$$
$$+ \beta_5 \text{kessner3}_{ij} + \beta_6 \text{novisit}_{ij} + \beta_7 \text{pretri2}_{ij} + \beta_8 \text{pretri3}_{ij} + \varepsilon_{ij}$$

within-subjects effects

$$+ \beta_9 \text{hsgrad}_j + \beta_{10} \text{somecoll}_j + \beta_{11} \text{collgrad}_j + \beta_{12} \text{married}_j$$
$$+ \beta_{13} \text{black}_j + \beta_0 + b_{0j}$$

between-subjects effects

Where i indexes an individual birth, j indexes a subject (momid)

- $\beta's$ represent the fixed effect of each predictor.
  - Do not vary
- ε_{ij} represent the random error within a subject
  - Vary from birth-to-birth within a subject
- b_{oj} represent the random variability in intercepts
  - Deviation from fixed intercept for each subject
  - Vary from subject-to-subject

LMM Assumptions

1. **Random Error:** $\varepsilon_{ij} \sim N(0, \sigma^2_{error})$
2. **Random Intercept:** $b_{0j} \sim N(0, \sigma^2_{intercept})$
3. ε_{ij} and b_{0j} are independent
4. b_{0j} are independent of the predictor variables

LMM SAS Code

```
/*Linear Mixed Model with Random Intercept
  for each Subject*/
title "Linear Mixed Model";
title2 "Within-Subject Effects are not Strictly 'Within'";
proc mixed data=smoking;
  class momid;
  model birwt = smoke male mage kessner2 kessner3
                novisit pretri2 pretri3
                hsgrad somecoll collgrad
                married black  / solution;
  random int / subject=momid;
run;
```

LMM Estimates

Solution for Fixed Effects					
Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	3117.08	40.9314	3972	76.15	<.0001
smoke	-218.27	18.1760	4618	-12.01	<.0001
male	120.94	9.5644	4618	12.64	<.0001
mage	8.1041	1.3460	4618	6.02	<.0001
kessner2	-92.9166	19.9402	4618	-4.66	<.0001
kessner3	-150.85	40.8586	4618	-3.69	0.0002
novisit	-30.0200	65.7379	4618	-0.46	0.6479
pretri2	92.8456	23.2068	4618	4.00	<.0001
pretri3	178.70	51.6725	4618	3.46	0.0005
hsgrad	56.8481	25.0641	4618	2.27	0.0234
somecoll	80.6825	27.3401	4618	2.95	0.0032
collgrad	90.8226	28.0264	4618	3.24	0.0012
married	49.9240	25.5322	4618	1.96	0.0506
black	-211.41	28.3103	4618	-7.47	<.0001

LMM Estimates Discussion

- Based on LMM, mom smoking is associated with 218 gram lower birth weight on average, after controlling for other covariates
- But, effect of smoking variable incorporates **both** within and between-subjects effects
  - **Within-Mom effect of smoking:** comparison of birth weights within the same mom at times when she was smoking vs. when she was not

  AND

  - **Between-Mom effect of smoking:** comparison of birth weights of moms who were smokers vs. those who were not

Next Fit a Strictly Between-Subjects Model

- Calculate mean of all within-subject variables, including birthweight, for each mom and then fit a linear regression model

Between-Subjects Model:

$$\text{Mbirwt}_j = \beta_0 + \beta_1 \text{msmoke}_j + \beta_2 \text{mmale}_j + \beta_3 \text{mmage}_j + \beta_4 \text{mkessner2}_j$$
$$+ \beta_5 \text{mkessner3}_j + \beta_6 \text{mnovisit}_j + \beta_7 \text{mpretri2}_j + \beta_8 \text{mpretri3}_j$$
$$+ \beta_9 \text{hsgrad}_j + \beta_{10} \text{somecoll}_j + \beta_{11} \text{collgrad}_j + \beta_{12} \text{married}_j$$
$$+ \beta_{13} \text{black}_j + \varepsilon_j$$

SAS Code for Between-Subjects Model

```
/*Between-Subjects Effect Only model
  First aggregate within-subjects variables
  across subjects. Copy Within-Subjects Variables*/
proc sort data=smoking;
 by momid;
run;
proc means data=smoking noprint;
 by momid;
 output out=meandat
        mean(birwt)=mbirwt
        mean(smoke)=msmoke
        mean(male)=mmale
        mean(mage)=mmage
        mean(kessner2)=mkessner2
        mean(kessner3)=mkessner3
        mean(novisit)=mnovisit
        mean(pretri2)=mpretri2
        mean(pretri3)=mpretri3;
 id hsgrad somecoll collgrad married black;
run;
```

SAS Code for Between-Subjects Model (cont)

```
title "Between-Subjects Effects Only";
proc reg data=meandat ;
  model mbirwt = msmoke mmale mmage
                 mkessner2 mkessner3
                 mnovisit mpretri2 mpretri3
                 hsgrad somecoll collgrad
                 married black ;
run;quit;
```

Estimates from Between-Subjects Model

Parameter Estimates						
Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	Intercept	1	3241.44962	46.15955	70.22	<.0001
msmoke	smoke	1	-286.14763	23.22554	-12.32	<.0001
mmale	male	1	104.94319	19.49531	5.38	<.0001
mmage	mage	1	4.39870	1.50545	2.92	0.0035
mkessner2	kessner2	1	-101.49312	37.65605	-2.70	0.0071
mkessner3	kessner3	1	-201.95993	79.28821	-2.55	0.0109
mnovisit	novisit	1	-51.02733	124.20730	-0.41	0.6812
mpretri2	pretri2	1	125.47762	44.72006	2.81	0.0050
mpretri3	pretri3	1	241.12005	100.65665	2.40	0.0166
hsgrad	hsgrad	1	58.80977	25.51424	2.30	0.0212
somecoll	somecoll	1	85.07129	28.13480	3.02	0.0025
collgrad	collgrad	1	99.87509	29.35324	3.40	0.0007
married	married	1	41.91268	26.10719	1.61	0.1085
black	black	1	-218.40454	28.57844	-7.64	<.0001

Between-Subjects Model Discussion

- One observation per subject in the between-subjects model
- Effect of smoking now estimated to reduce birth weight by 286 grams, after adjusting for other covariates
  - This is the strictly between-subjects effect of smoking,
  - A one-unit change in msmoke represents the difference between a subject who always smoked (msmoke=1.0) vs. a subject who never smoked (msmoke=0.0).
  - Note: subjects who smoked during some pregnancies and not during others, would have fractional values of msmoke (e.g. msmoke=0.5 for a mom with 2 pregnancies who smoked during one pregnancy, but not during the other).

Next Fit a Fixed Effect (Strictly Within-Subjects) Model

- Include a dummy variable for each mom minus 1
  - Subject (momid) is now a **fixed effect**
  - *No between-subjects variables can be included*
  - Entire effect of all between-subject variables is taken up by the dummy variables for momid
- There are no random effects in this model
- Warning: This model takes a long time to fit!
  - But there is another way to do it, using absorption

Fixed Effects (Strictly Within-Subjects) Model

Within-Subjects (Fixed Effects) Model:

$$\text{Birwt}_{ij} = \beta_0 + \beta_1 \text{smoke}_{ij} + \beta_2 \text{male}_{ij} + \beta_3 \text{mage}_{ij} + \beta_4 \text{kessner2}_{ij}$$
$$+ \beta_5 \text{kessner3}_{ij} + \beta_6 \text{novisit}_{ij} + \beta_7 \text{pretri2}_{ij} + \beta_8 \text{pretri3}_{ij}$$
$$+ \beta_9 \text{thru} \beta_{3985} \text{momid}_j + \varepsilon_{\text{ij}}$$

Fixed Effects (Within-Subjects) Model SAS Code

```
/*This takes a long time to run*/
title "Within-Subjects Analysis";
proc mixed data=smoking noclprint;
  class momid;
  model birwt =  smoke male mage kessner2 kessner3
                 novisit pretri2 pretri3 momid /solution;
run;
```

Estimates from Fixed-Effects (Within-Subjects) Model

Solution for Fixed Effects						
Effect	momid	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept		2603.79	285.62	4618	9.12	<.0001
smoke		-104.55	29.1008	4618	-3.59	0.0003
male		125.64	10.9227	4618	11.50	<.0001
mage		23.1583	3.0067	4618	7.70	<.0001
kessner2		-91.4948	23.4891	4618	-3.90	<.0001
kessner3		-128.09	47.7964	4618	-2.68	0.0074
novisit		-4.8059	77.7721	4618	-0.06	0.9507
pretri2		81.2904	27.0497	4618	3.01	0.0027
pretri3		153.06	60.0845	4618	2.55	0.0109
momid	14	-1.9036	344.41	4618	-0.01	0.9956
momid	25	-277.74	340.91	4618	-0.81	0.4153
momid	39	-341.73	338.07	4618	-1.01	0.3122
momid	48	-265.29	369.48	4618	-0.72	0.4728

Fixed Effects (Within-Subjects) Model Discussion

- Effect of smoking now estimated to reduce birth weight by 105 grams, after adjusting for other covariates
  - This is the strictly within-subjects effect of smoking
  - Gives a comparison of birth weights within the **same** mom for births when she smoked vs. births when she did not smoke
  - Effect is smaller than the between-subjects effect
  - But is this difference significant?

Alternative Method for Within-Subjects Model Using Absorption: SAS Code

```
/*This runs very quickly*/
title "Alternate Within-Subjects Effects";
proc glm data=smoking ;
  model birwt = smoke male mage kessner2 kessner3 novisit
                pretri2 pretri3 / solution ss3;
 absorb momid;
run; quit;
```

Estimates from Absorption Model

Parameter	Estimate	Standard Error	t Value	Pr > |t|
smoke	-104.5493595	29.10075118	-3.59	0.0003
male	125.6354628	10.92271717	11.50	<.0001
mage	23.1583198	3.00666724	7.70	<.0001
kessner2	-91.4948309	23.48913613	-3.90	<.0001
kessner3	-128.0910069	47.79635759	-2.68	0.0074
novisit	-4.8058981	77.77209979	-0.06	0.9507
pretri2	81.2903885	27.04973758	3.01	0.0027
pretri3	153.0590143	60.08453085	2.55	0.0109

Fit a Between-Within (BW) aka: Hybrid Model

Include both the within- and between-subjects effects of smoking in the same model:

1. Center each within-subject variable by creating the **deviations** at each time point from the mean for *that subject*
   - o Deviations capture the strictly within-subjects effect of smoking and other within-subjects covariates
   - o Label these variables with d prefix

Fit a BW Model (cont.)

2. Also include the Subject-specific means of all within-subjects variables
   - Means capture the between-subjects effects of these variables
   - Label these variables with m prefix

Ex: Mom who had two pregnancies and smoked in one, but not in the other.

Smoke	Dsmoke	Msmoke
1	.5	.5
0	-.5	.5

Between-Within Model (cont.)

- This model allows us to simultaneously estimate both between-subjects and within-subjects effects
- Within- and between-subject effects are independent
- We can compare the within- and between-subjects effects of variables using contrasts (post-hoc tests)

Between-Within Model

$$\left.\begin{aligned} \text{Birwt}_{ij} &= \beta_1 \text{dsmoke}_{ij} + \beta_2 \text{dmale}_{ij} + \beta_3 \text{dmage}_{ij} + \beta_4 \text{dkessner2}_{ij} \\ &+ \beta_5 \text{dkessner3}_{ij} + \beta_6 \text{dnovisit}_{ij} + \beta_7 \text{dpretri2}_{ij} + \beta_8 \text{dpretri3}_{ij} + \varepsilon_{ij} \end{aligned}\right\} \textbf{within-subjects}$$

$$\left.\begin{aligned} &+ \beta_9 \text{msmoke}_j + \beta_{10} \text{mmale}_j + \beta_{11} \text{mmage}_j + \beta_{12} \text{mkessner2}_j \\ &+ \beta_{13} \text{mkessner3}_j + \beta_{14} \text{mnovisit}_j + \beta_{15} \text{mpretri2}_j + \beta_{16} \text{mpretri3}_j \\ &+ \beta_{17} \text{hsgrad}_j + \beta_{18} \text{somecoll}_j + \beta_{19} \text{collgrad}_j + \beta_{20} \text{married}_j \\ &+ \beta_{21} \text{black}_j + \beta_0 + \text{b}_{0j} \end{aligned}\right\} \textbf{between-subjects}$$

Between-Within Model Assumptions

1. **Random Error:** $\varepsilon_{ij} \sim N(0, \sigma^2_{error})$
2. **Random Intercept:** $b_{0j} \sim N(0, \sigma^2_{intercept})$
3. ε_{ij} and b_{0j} are independent
4. b_{0j} are independent of the predictor variables

Between-Within Model SAS Code

```
/*Setup for Between-Within Model Analysis*/
data smoking2;
   set  smoking;
        dbirwt=birwt;
        dsmoke=smoke;
        dmale=male;
        dmage=mage;
        dkessner2=kessner2;
        dkessner3=kessner3;
        dnovisit=novisit;
        dpretri2=pretri2;
        dpretri3=pretri3;
run;
```

Between-Within Model SAS Code (Cont)

```
title "Create Deviations from Mom Mean";
title2 "For Within-Subjects Variables";

proc standard data=smoking2 mean=0 out=stddata;
  by momid;
  var  dbirwt
       dsmoke
       dmale
       dmage
       dkessner2
       dkessner3
       dnovisit
       dpretri2
       dpretri3;
run;
```

Between-Within Model SAS Code (Cont)

```
/*Merge the Mean Data with the Deviations Data*/
data smoking3;
  merge stddata meandat;
  by momid;
run;
```

Between-Within Model SAS Code (Cont)

```
title "Hybrid Model: Between and Within Effects";
title2 "Plus Random Intercept Per Subject";

proc mixed data=smoking3 noclprint;
  class momid;
  model birwt = dsmoke dmale dmage dkessner2 dkessner3 dnovisit dpretri2 dpretri3
                msmoke mmale mmage mkessner2 mkessner3 mnovisit mpretri2 mpretri3
                hsgrad somecoll collgrad married black/ solution;
  random intercept / subject=momid;
  contrast "smoking" dsmoke 1 msmoke -1;
  contrast "male" dmale 1 mmale -1;
  contrast "mage" dmage 1 mmage -1;
  contrast "kessner" dkessner2 1 mkessner2 -1,
                     dkessner3 1 mkessner3 -1;
  contrast "pretri" dnovisit 1 mnovisit -1,
                    dpretri2 1 mpretri2 -1,
                    dpretri3 1 mpretri3 -1;
run;
```

BW Model Results

Solution for Fixed Effects					
Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	3238.41	46.0702	3966	70.29	<.0001
dsmoke	-104.55	29.1314	4616	-3.59	0.0003
dmale	125.64	10.9342	4616	11.49	<.0001
dmage	23.1583	3.0098	4616	7.69	<.0001
dkessner2	-91.4948	23.5139	4616	-3.89	0.0001
dkessner3	-128.09	47.8467	4616	-2.68	0.0075
dnovisit	-4.8059	77.8541	4616	-0.06	0.9508
dpretri2	81.2904	27.0783	4616	3.00	0.0027
dpretri3	153.06	60.1479	4616	2.54	0.0110
msmoke	-287.71	23.2013	4616	-12.40	<.0001
mmale	105.41	19.4867	4616	5.41	<.0001
mmage	4.5639	1.5019	4616	3.04	0.0024
mkessner2	-100.55	37.5315	4616	-2.68	0.0074
mkessner3	-207.50	79.0589	4616	-2.62	0.0087
mnovisit	-42.9381	124.15	4616	-0.35	0.7295
mpretri2	126.06	44.5571	4616	2.83	0.0047
mpretri3	249.11	100.22	4616	2.49	0.0130
hsgrad	56.3019	25.4312	4616	2.21	0.0269
somecoll	83.0741	28.0403	4616	2.96	0.0031
collgrad	98.1793	29.2387	4616	3.36	0.0008
married	42.4602	26.0775	4616	1.63	0.1035
black	-219.00	28.4680	4616	-7.69	<.0001

BW Model: Comparison of Between vs. Within Effects

Contrasts				
Label	Num DF	Den DF	F Value	Pr > F
smoking	1	4616	24.19	<.0001
male	1	4616	0.82	0.3655
mage	1	4616	30.56	<.0001
kessner	2	4616	0.37	0.6901
pretri	3	4616	0.61	0.6091

- The between-subjects effect of smoking (-287.71 gm) is significantly greater than the within-subjects effect (-104.56 gm)
- There is also a significantly greater effect of within-mom age (125.64 gm) than between-mom age (4.56 gm)
- There may be other unmeasured factors that differ between smokers and non-smokers besides smoking alone.

Final BW Model

- Include only between and within effects that differ significantly
- Include a random effect for each subject (mom)

$$\left.\begin{aligned} \text{Birwt}_{ij} &= \beta_1 \text{dsmoke}_{ij} + \beta_2 \text{male}_{ij} + \beta_3 \text{dmage}_{ij} + \beta_4 \text{kessner2}_{ij} \\ &+ \beta_5 \text{kessner3}_{ij} + \beta_6 \text{novisit}_{ij} + \beta_7 \text{pretri1}_{ij} + \beta_8 \text{pretri2}_{ij} + \varepsilon_{ij} \end{aligned}\right\} \textbf{within-subjects}$$

$$\left.\begin{aligned} &+ \beta_9 \text{msmoke}_j + \beta_{10} \text{mmage}_j + \beta_{11} \text{hsgrad}_j + \beta_{12} \text{somecoll}_j \\ &+ \beta_{13} \text{collgrad}_j + \beta_{14} \text{married}_j + \beta_{15} \text{black}_j + \beta_0 + \text{b}_{0\text{j}} \end{aligned}\right\} \textbf{between-subjects}$$

Final BW Model SAS Code

```
ods exclude solutionr;
title "Final BW Model";
title2 "Include Individual Effects Plus Means";
title3 "For variables that have different between and within effects";
proc mixed data=smoking3 noclprint plots=residualpanel plots(maxpoints=9000);
  class momid;
  model birwt = dsmoke
                male
                dmage
                kessner2 kessner3
                novisit pretri2 pretri3
                msmoke mmage
                hsgrad somecoll collgrad
                married black
                / solution;
  random intercept / subject=momid solution;
  ods output solutionr=re_data;
  run;
```

Final BW Model Results

Solution for Fixed Effects					
Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	3236.10	44.0284	3970	73.50	<.0001
dsmoke	-105.18	29.1214	4618	-3.61	0.0003
male	120.51	9.5301	4618	12.65	<.0001
dmage	23.2590	3.0087	4618	7.73	<.0001
kessner2	-94.0274	19.8730	4618	-4.73	<.0001
kessner3	-151.85	40.7157	4618	-3.73	0.0002
novisit	-22.3702	65.5172	4618	-0.34	0.7328
pretri2	94.1728	23.1259	4618	4.07	<.0001
pretri3	180.73	51.4926	4618	3.51	0.0005
msmoke	-288.22	23.1427	4618	-12.45	<.0001
mmage	4.4700	1.5001	4618	2.98	0.0029
hsgrad	55.7221	25.2709	4618	2.20	0.0275
somecoll	82.1824	27.7454	4618	2.96	0.0031
collgrad	96.6840	28.8811	4618	3.35	0.0008
married	42.4676	25.6878	4618	1.65	0.0984
black	-220.64	28.3006	4618	-7.80	<.0001

Comments on Between-Within Model

The within-subjects effects of smoking:

- The average birthweight of babies born to mothers in a pregnancy when they smoked is 105 grams less than in a pregnancy when they didn't smoke, after controlling for other covariates.

The between-subjects effect of smoking:

- The average birthweight of babies born to mothers who smoked during all their pregnancies is 288 grams less than for mothers who never smoked, controlling for other covariates.

Some Thoughts

- It isn't necessary to go through the entire process of fitting a between- and within-model separately, which was done here for illustration purposes.
- You can start with the between-within model.
- For linear models, i.e. with normally distributed outcomes, the BW method produces the same results for the within-subjects effects as the standard fixed effects method
- But for logistic regression, the BW estimates are not identical to those produced by conditional likelihood (the gold standard), fitted using strata in proc logistic in SAS.
  - Some caution, as well as some modifications are suggested for binary (logistic) regression and other non-linear models.
  - See Allison, "Problems with the Hybrid Method".

References

- The data analysis and code for this presentation was generated using SAS 9.4 (TS1M5) software. Copyright © 2016 SAS Institute Inc. SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc., Cary, NC, USA.
- Abrevaya, J. 2006. Estimating the effect of smoking on birth outcomes using a matched panel data approach. lJournal of Applied Econometrics 21: 489-519.
- Rabe-Hesketh, S. and Skrondal, A. (2012). Multilevel and Longitudinal Modeling Using Stata (3rd Edition). College Station, TX: Stata Press.
- Allison, PD (2005) Fixed Effects Regression Methods for Longitudinal Data Using SAS. Cary, NC: The SAS Institute.
- Problems with the Hybrid Method, September 2, 2014 By Paul Allison, Statistical Horizons: https://statisticalhorizons.com/problems-with-the-hybrid-method