

SOCY7708: Hierarchical Linear Modeling
Instructor: Natasha Sarkisian
Class notes: Three Level Models

A 3-level HLM model is a logical extension of the 2-level model. Similar model building strategies apply, although now both level 1 and level 2 slopes may vary.

The data for our example are stored in eg1.dta, eg2.dta, and eg3.dta. These data consist of 7230 observations collected on 1721 children from 60 schools, beginning at the end of grade one and followed up annually thereafter until grade six. The students are the level-2 units, the schools are the level-3 units. The outcome variable of interest is the result of a math test, represented by the variable MATH. On level 1, the information available includes the year of the study (denoted by the variable YEAR), the grade level at each testing occasion (denoted by GRADE) and an indicator that a child is retained in grade for a particular year (denoted by the variable RETAINED, which assumes a value of 1 if retained, 0 otherwise). At level-2, the following variables are available: GENDER, denoting the gender of the child (1=female), BLACK, denoting whether the child is Black, and HISPANIC, denoting whether the child is Hispanic. At level-3, the following variables are available: SIZE, measuring the size of school, LOWINC indicating the percentage of low-income students in the school, and MOBILITY indicating the level of residential mobility on the school level (%).

Let's merge the three files; we will use variables CHILDDID and SCHOOLID to match across levels.

```
. merge m:1 schoolid childid using eg2.dta

Result                                     # of obs.
-----
not matched                                0
matched                                   7,230  (_merge==3)
-----

. drop _merge

. merge m:1 schoolid using eg3.dta

Result                                     # of obs.
-----
not matched                                0
matched                                   7,230  (_merge==3)
-----

. drop _merge
```

Let's start by modeling the linear growth trajectory unconditionally and without random effects at levels 2 and 3:

```
. tab year

   year |      Freq.      Percent      Cum.
-----+-----
   1.00 |         131         1.81         1.81
```



```
-----+-----
      basic3 |      7,230      . -8373.522      5  16757.04  16791.47
-----+-----
```

Note: BIC uses N = number of observations. See [R] BIC note.

```
. mixed math year0 || schoolid:
```

```
Mixed-effects ML regression      Number of obs      =      7,230
Group variable: schoolid         Number of groups   =      60
```

```
Obs per group:
      min =      18
      avg =     120.5
      max =     387
```

```
Log likelihood = -10343.209      Wald chi2(1)      =     7756.87
                                Prob > chi2           =      0.0000
```

```
-----+-----
      math |      Coef.      Std. Err.      z    P>|z|      [95% Conf. Interval]
-----+-----
      year0 |      .751992     .0085383     88.07  0.000     .7352573     .7687267
      _cons |     -2.649882     .0652758    -40.60  0.000    -2.77782    -2.521944
-----+-----
```

```
-----+-----
      Random-effects Parameters |      Estimate      Std. Err.      [95% Conf. Interval]
-----+-----
schoolid: Identity
      var(_cons) |      .207243     .0401772     .1417298     .3030389
-----+-----
      var(Residual) |      .9978508     .0166657     .9657154     1.031056
-----+-----
```

```
LR test vs. linear model: chibar2(01) = 1235.39      Prob >= chibar2 = 0.0000
```

```
. estat ic
```

Akaike's information criterion and Bayesian information criterion

```
-----+-----
      Model |      N      ll(null)      ll(model)      df      AIC      BIC
-----+-----
      . |      7,230      . -10343.21      4  20694.42  20721.96
-----+-----
```

Note: BIC uses N = number of observations. See [R] BIC note.

```
. est store basic_school
```

```
. lrtest basic3 basic_school
```

```
Likelihood-ratio test      LR chi2(1) =     3939.37
(Assumption: basic_school nested in basic3)      Prob > chi2 =      0.0000
```

Note: The reported degrees of freedom assumes the null hypothesis is not on the boundary of the parameter space. If this is not true, then the reported test is conservative.

```
. mixed math year0 || childid:
```

```
Mixed-effects ML regression      Number of obs      =      7,230
Group variable: childid         Number of groups   =     1,721
```

```
Obs per group:
      min =      2
      avg =     4.2
```

```

max = 6
Wald chi2(1) = 19156.93
Prob > chi2 = 0.0000
Log likelihood = -8515.4376
-----
math |      Coef.   Std. Err.      z    P>|z|    [95% Conf. Interval]
-----+-----
year0 |   .7474524   .0054003   138.41  0.000    .736868   .7580369
_cons |  -2.707306   .0281623   -96.13  0.000   -2.762503 -2.652109
-----
Random-effects Parameters |   Estimate  Std. Err.   [95% Conf. Interval]
-----+-----
childid: Identity |
var(_cons) |   .8677144   .0325757    .8061596   .9339692
-----+-----
var(Residual) |   .3469385   .0066102    .3342216   .3601392
-----
LR test vs. linear model: chibar2(01) = 4890.93      Prob >= chibar2 = 0.0000

```

```
. estat ic
```

```

Akaike's information criterion and Bayesian information criterion
-----
Model |      N    ll(null)  ll(model)    df      AIC      BIC
-----+-----
. |      7,230          .  -8515.438     4   17038.88  17066.42
-----

```

Note: BIC uses N = number of observations. See [R] BIC note.

```
. est store basic_child
. lrtest basic3 basic_child
```

```

Likelihood-ratio test          LR chi2(1) = 283.83
(Assumption: basic_child nested in basic3)  Prob > chi2 = 0.0000

```

Note: The reported degrees of freedom assumes the null hypothesis is not on the boundary of the parameter space. If this is not true, then the reported test is conservative.

```
. mixed math year0 || schoolid: year0, cov(unstr) || childid: year0, cov(unstr)
```

```

Mixed-effects ML regression          Number of obs = 7,230
-----
Group Variable |      No. of      Observations per Group
                |      Groups      Minimum   Average   Maximum
-----+-----
schoolid |           60           18   120.5     387
childid |       1,721            2     4.2         6
-----

```

```

Wald chi2(1) = 2499.85
Prob > chi2 = 0.0000
Log likelihood = -8163.1156
-----
math |      Coef.   Std. Err.      z    P>|z|    [95% Conf. Interval]
-----+-----
year0 |   .7630274   .015261    50.00  0.000    .7331163   .7929384
_cons |  -2.686874   .0568029   -47.30  0.000   -2.798205 -2.575542
-----

```

```

Random-effects Parameters |   Estimate  Std. Err.   [95% Conf. Interval]
-----+-----

```

```

schoolid: Unstructured |
      var(year0) |      .0110172   .0025621   .0069843   .017379
      var(_cons) |      .148943   .0348396   .0941702   .2355736
      cov(year0,_cons) |    -.0104966   .0071906   -.0245899   .0035967
-----+-----
childid: Unstructured |
      var(year0) |      .0112561   .001961   .0080001   .0158373
      var(_cons) |      .4768837   .0323937   .4174381   .5447946
      cov(year0,_cons) |    .0186451   .0062279   .0064385   .0308516
-----+-----
      var(Residual) |    .3014383   .0066403   .2887004   .3147382
-----+-----
LR test vs. linear model: chi2(6) = 5595.58          Prob > chi2 = 0.0000

```

Note: LR test is conservative and provided only for reference.

```
. estat recov
```

Random-effects covariance matrix for level schoolid

```

      |      year0      _cons
-----+-----
year0 |      .0110172
_cons |    -.0104966      .148943

```

Random-effects covariance matrix for level childid

```

      |      year0      _cons
-----+-----
year0 |      .0112561
_cons |    .0186451      .4768837

```

```
. est store year_varies
```

```
. estat ic
```

Akaike's information criterion and Bayesian information criterion

```

-----+-----
Model |      N   ll(null)   ll(model)   df      AIC      BIC
-----+-----
year_varies |      7,230      . -8163.116      9  16344.23  16406.21
-----+-----

```

Note: BIC uses N = number of observations. See [R] BIC note.

```
. lrtest basic3 year_varies
```

```

Likelihood-ratio test          LR chi2(4) =      420.81
(Assumption: basic3 nested in year_varies)  Prob > chi2 =      0.0000

```

Note: The reported degrees of freedom assumes the null hypothesis is not on the boundary of the parameter space. If this is not true, then the reported test is conservative.

Let's add level 1 variable, retained, and examine random slope for it:

```
. mixed math year0 retained || schoolid: year0 retained, cov(unstr) || childid:
year0 retained, cov(unstr)
```

```

Mixed-effects ML regression          Number of obs      =      7,230
-----+-----
      |      No. of      Observations per Group

```



```
-----
Log likelihood = -8154.6735          Wald chi2(2)      =    2519.79
                                   Prob > chi2         =     0.0000
-----
```

math	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
year0	.764384	.0152565	50.10	0.000	.7344818	.7942863
retained	.1387952	.0333952	4.16	0.000	.0733417	.2042486
_cons	-2.698604	.0574234	-46.99	0.000	-2.811151	-2.586056

```
-----
```

Random-effects Parameters	Estimate	Std. Err.	[95% Conf. Interval]	
schoolid: Unstructured				
var(year0)	.0110101	.0025559	.0069854	.0173538
var(_cons)	.1521118	.0355767	.0961788	.2405728
cov(year0,_cons)	-.010395	.0072502	-.0246051	.0038151
childid: Unstructured				
var(year0)	.0114371	.0019599	.0081744	.0160021
var(_cons)	.4920312	.0329821	.431454	.5611137
cov(year0,_cons)	.0175375	.0062609	.0052664	.0298087
var(Residual)	.2986372	.0065957	.2859857	.3118484

```
-----
LR test vs. linear model: chi2(6) = 5468.63          Prob > chi2 = 0.0000
-----
```

Note: LR test is conservative and provided only for reference.

```
. lrtest . retained_varies
```

```
Likelihood-ratio test          LR chi2(6) =    14.79
(Assumption: . nested in retained_var~s)  Prob > chi2 =    0.0220
```

Note: The reported degrees of freedom assumes the null hypothesis is not on the boundary of the parameter space. If this is not true, then the reported test is conservative.

```
. estat ic
```

Akaike's information criterion and Bayesian information criterion

```
-----
```

Model	N	ll(null)	ll(model)	df	AIC	BIC
.	7,230	.	-8154.673	10	16329.35	16398.21

```
-----
```

Note: BIC uses N = number of observations. See [R] BIC note.

Now we can add some explanatory variables on level 2 and cross-level interactions. I also attempted to allow slopes of level 2 variables to vary, but only BLACK seems to have any sizeable slope variation, so I included it here. Also, covariances among random effects of slopes on school level had to be constrained to be 0 (they appeared very close to 0, and wouldn't estimate otherwise in this model with black and child-level covariances).

```
. mixed math c.year0##i.female c.year0##i.black c.year0##i.hispanic i.retained
##i.female i.retained##i.black i.retained##i.hispanic || schoolid: year0
retained black || childid: year0 retained, cov(unstr)
```

Mixed-effects ML regression Number of obs = 7,230

Group Variable	No. of Groups	Observations per Group		
		Minimum	Average	Maximum
schoolid	60	18	120.5	387
childid	1,721	2	4.2	6

Log likelihood = -8116.7578 Wald chi2(11) = 3245.78
Prob > chi2 = 0.0000

math	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
year0	.7770095	.0211966	36.66	0.000	.7354649	.8185541
1.female	-.0515563	.0473707	-1.09	0.276	-.1444012	.0412886
female# c.year0						
1	.0150246	.011759	1.28	0.201	-.0080226	.0380718
year0	0 (omitted)					
1.black	-.4544733	.0883343	-5.14	0.000	-.6276054	-.2813412
black# c.year0						
1	-.0492823	.021027	-2.34	0.019	-.0904944	-.0080702
year0	0 (omitted)					
1.hispanic	-.4798107	.0969086	-4.95	0.000	-.669748	-.2898734
hispanic# c.year0						
1	.036259	.0241381	1.50	0.133	-.0110509	.0835688
1.retained	-.0514953	.1233298	-0.42	0.676	-.2932173	.1902267
retained# female						
1 1	.0909304	.0677083	1.34	0.179	-.0417755	.2236362
retained# black						
1 1	.1106674	.1286472	0.86	0.390	-.1414765	.3628113
retained# hispanic						
1 1	.0534891	.1533108	0.35	0.727	-.2469946	.3539729
_cons	-2.307613	.0813587	-28.36	0.000	-2.467073	-2.148153

Random-effects Parameters	Estimate	Std. Err.	[95% Conf. Interval]	
schoolid: Independent				
var(year0)	.0081062	.0019893	.005011	.0131131
var(retained)	.026531	.0158145	.0082484	.0853363
var(black)	.0487207	.0463472	.0075506	.314374
var(_cons)	.0877577	.0349077	.0402438	.1913687
childid: Unstructured				
var(year0)	.0119105	.0020023	.0085671	.0165587


```

          var(retained) |   .0237611   .0366761   .0011535   .4894708
          var(_cons) |   .4877012   .0332189   .4267521   .5573551
    cov(year0,retained) |   .0029577   .0072705   -.0112922   .0172076
          cov(year0,_cons) |   .0148103   .0063953   .0022757   .0273448
          cov(retained,_cons) |  -.087717   .0390159   -.1641867   -.0112473
-----+-----
          var(Residual) |   .296089   .0067807   .2830929   .3096817
-----+-----
LR test vs. linear model: chi2(10) = 5000.43          Prob > chi2 = 0.0000

```

Note: LR test is conservative and provided only for reference.

Next, we can consider exploring level 3 predictors. Let's mean-center and aggregate:

```

. egen tag=tag(schoolid)

. for var size lowinc mobility: sum X if tag==1 \ gen Xm=X-r(mean)

-> sum size if tag==1

  Variable |      Obs      Mean   Std. Dev.   Min     Max
-----+-----
    size |         60   642.5333   317.3676   113    1486

-> gen sizem=size-r(mean)

-> sum lowinc if tag==1

  Variable |      Obs      Mean   Std. Dev.   Min     Max
-----+-----
  lowinc |         60   73.73667   27.27405     0    100

-> gen lowincm=lowinc-r(mean)

-> sum mobility if tag==1

  Variable |      Obs      Mean   Std. Dev.   Min     Max
-----+-----
  mobility |         60   34.745    13.20786     8.8     67

-> gen mobilitym=mobility-r(mean)

. for var retained female black hispanic: bysort schoolid: egen X_sch=mean(X)

-> bysort schoolid: egen retained_sch=mean(retained)

-> bysort schoolid: egen female_sch=mean(female)

-> bysort schoolid: egen black_sch=mean(black)

-> bysort schoolid: egen hispanic_sch=mean(hispanic)

. for var retained_sch- hispanic_sch: sum X if tag==1 \ gen Xm=X-r(mean)

-> sum retained_sch if tag==1

  Variable |      Obs      Mean   Std. Dev.   Min     Max
-----+-----
  retained_sch |         60   .0554006   .0670196     0   .3714286

-> gen retained_schm=retained_sch-r(mean)

```

```
-> sum female_sch if tag==1
```

Variable	Obs	Mean	Std. Dev.	Min	Max
female_sch	60	.4984854	.1116362	.1428571	.744186

```
-> gen female_schm=female_sch-r(mean)
```

```
-> sum black_sch if tag==1
```

Variable	Obs	Mean	Std. Dev.	Min	Max
black_sch	60	.5734986	.4109914	0	1

```
-> gen black_schm=black_sch-r(mean)
```

```
-> sum hispanic_sch if tag==1
```

Variable	Obs	Mean	Std. Dev.	Min	Max
hispanic_sch	60	.2031471	.2850104	0	1

```
-> gen hispanic_schm=hispanic_sch-r(mean)
```

```
. mixed math c.year0##i.female c.year0##i.black c.year0##i.hispanic
i.retained##i.female i.retained##i.black i.retained##i.hispanic sizem lowincm
mobilitym retained_schm female_schm black_schm hispanic_schm || schoolid: year0
retained black || childid: year0 retained, cov(unstr)
```

```
Mixed-effects ML regression                               Number of obs   =       7,230
```

Group Variable	No. of Groups	Observations per Group		
		Minimum	Average	Maximum
schoolid	60	18	120.5	387
childid	1,721	2	4.2	6

```
Log likelihood = -8100.6828                               Wald chi2(18)   =       3435.06
                                                            Prob > chi2     =         0.0000
```

	math	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
year0		.7789677	.0211264	36.87	0.000	.7375608 .8203746
1.female		-.0588736	.0475313	-1.24	0.215	-.1520333 .034286
female#						
c.year0						
1		.0157074	.0117746	1.33	0.182	-.0073703 .0387852
year0		0	(omitted)			
1.black		-.4507678	.0950697	-4.74	0.000	-.637101 -.2644345
black#						
c.year0						
1		-.0514084	.021239	-2.42	0.016	-.0930361 -.0097806
year0		0	(omitted)			
1.hispanic		-.4231151	.1020328	-4.15	0.000	-.6230957 -.2231345
hispanic#						
c.year0						
1		.0316889	.0242593	1.31	0.191	-.0158585 .0792363

1.retained		-.0393028	.1244782	-0.32	0.752	-.2832756	.20467
retained#							
female							
1 1		.089565	.0677085	1.32	0.186	-.0431413	.2222713
retained#							
black							
1 1		.1145005	.1299947	0.88	0.378	-.1402845	.3692855
retained#							
hispanic							
1 1		.0571244	.1542868	0.37	0.711	-.2452722	.359521
size		-.0001194	.0001421	-0.84	0.401	-.0003979	.0001591
lowincm		-.0045189	.0026423	-1.71	0.087	-.0096977	.0006598
mobilitym		-.0046824	.0036244	-1.29	0.196	-.0117861	.0024214
retained_s~m		-3.333575	.7736528	-4.31	0.000	-4.849907	-1.817244
female_schm		.1893789	.4174829	0.45	0.650	-.6288726	1.00763
black_schm		.490364	.2572757	1.91	0.057	-.0138871	.994615
hispanic_s~m		.3489856	.341413	1.02	0.307	-.3201715	1.018143
_cons		-2.318878	.0820417	-28.26	0.000	-2.479677	-2.158079

Random-effects Parameters		Estimate	Std. Err.	[95% Conf. Interval]
schoolid: Independent				
var(year0)		.0076979	.0019595	.0046741 .0126778
var(retained)		.0319476	.0172229	.011106 .0919006
var(black)		5.61e-13	5.79e-10	0
var(_cons)		.0565938	.0162535	.0322335 .0993644
childid: Unstructured				
var(year0)		.0120101	.0020084	.0086537 .0166682
var(retained)		.0177631	.0357394	.0003443 .9165099
var(_cons)		.4877259	.0331133	.4269578 .557143
cov(year0,retained)		.0031477	.0072575	-.0110766 .0173721
cov(year0,_cons)		.0147692	.0064009	.0022238 .0273147
cov(retained,_cons)		-.0765241	.0388023	-.1525753 -.000473
var(Residual)		.2959666	.006776	.2829794 .3095499

LR test vs. linear model: $\chi^2(10) = 4499.42$ Prob > $\chi^2 = 0.0000$

Note: LR test is conservative and provided only for reference.

We could further examine cross-level interactions involving these three level predictors (we can do two-way interactions with level 1 and level 2 variables as well as three-way interactions across all levels). We can also see that variance for black slope should probably be fixed after all.