SIMS Variance Decomposition

The Second International Mathematics Study (SIMS; Muthén, 1991, JEM).

- National probability sample of school districts selected proportional to size; a probability sample of schools selected proportional to size within school district, and two classes randomly drawn within each school
- 3,724 students observed in 197 classes from 113 schools with class sizes varying from 2 to 38; typical class size of around 20
- Eight variables corresponding to various areas of eighth-grade mathematics
- Same set of items administered as a pretest in the Fall of eighth grade and as a posttest in the Spring.

77

SIMS Variance Decomposition (Continued)

Muthén (1991). Multilevel factor analysis of class and student achievement components. Journal of Educational Measurement, 28, 338-354.

- Research questions: "The substantive questions of interest in this article are the variance decomposition of the subscores with respect to within-class student variation and between-class variation and the change of this decomposition from pretest to posttest. In the SIMS … such variance decomposition relates to the effects of tracking and differential curricula in eighth-grade math. On the one hand, one may hypothesize that effects of selection and instruction tend to increase between-class variation relative to within-class variation, assuming that the classes are homogeneous, have different performance levels to begin with, and show faster growth for higher initial performance level. On the other hand, one may hypothesize that eighth-grade exposure to new topics will increase individual differences among students within each class so that posttest within-class variation will be sizable relative to posttest between-class variation."

78

SIMS Variance Decomposition (Continued)

$y_{rij} = v_r + \lambda_{Br}\, \eta_{Bj} + \varepsilon_{Brj} + \lambda_{wr}\, \eta_{wij} + \varepsilon_{wrij}$

$V(y_{rij}) =$ BF + BE + WF + WE

Between reliability: BF / (BF + BE)
– BE often small (can be fixed at 0)

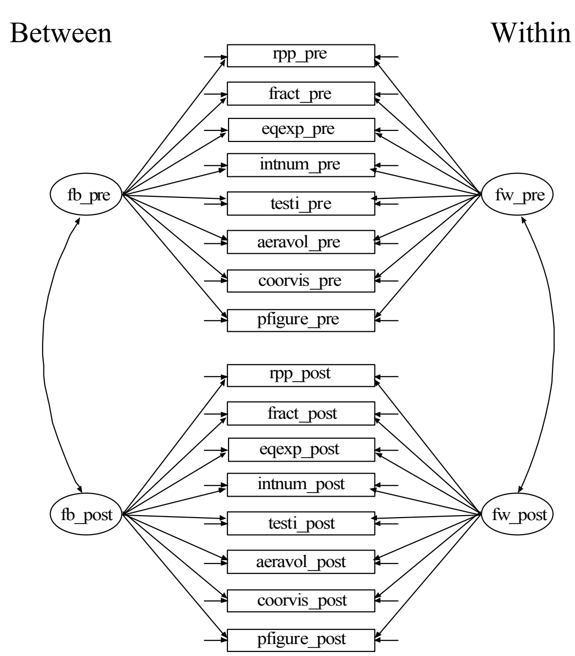
Within reliability: WF / (WF + WE)
– sum of a small number of items gives a large WE

Intraclass correlation:
ICC = (BF + BE) / (BF + BE + WF+ WE)

Large measurement error → large WE → small ICC

True ICC = BF / (BF + WF)

79



80

Table 4: Variance Decomposition of SIMS Achievement Scores (percentages of total variance in parenthesis)

		ANOVA								FACTOR ANALYSIS			
		Pretest			Posttest			% Increase In Variance		Error-free Prop. Between		Error-free % Increase In Variance	
	Number of Items	Between	Within	Prop-Between	Between	Within	Prop-Between	Between	Within	Pre	Post	Between	Within
RPP	8	1.542 (34.0)	2.990 (66.0)	.34	2.084 (38.5)	3.326 (61.5)	.38	35	11	.54	.52	29	41
FRACT	8	1.460 (38.2)	2.366 (61.8)	.38	1.906 (40.8)	2.767 (59.2)	.41	31	17	.60	.58	29	41
EQEXP	6	.543 (26.9)	1.473 (73.1)	.27	1.041 (38.7)	1.646 (61.3)	.39	92	18	.65	.64	113	117
INTNUM	2	.127 (25.2)	.358 (70.9)	.29	.195 (30.6)	.442 (69.4)	.31	54	24	.63	.61	29	41
TESTI	5	.580 (33.3)	1.163 (66.7)	.33	.664 (34.5)	1.258 (65.5)	.34	15	8	.58	.56	29	41
AREAVOL	2	.094 (17.2)	.451 (82.8)	.17	.156 (24.1)	.490 (75.9)	.24	66	9	.54	.52	29	41
COORVIS	3	.173 (20.9)	.656 (79.1)	.21	.275 (28.7)	.680 (68.3)	.32	59	4	.57	.55	29	41
PFIGURE	5	.363 (22.9)	1.224 (77.1)	.23	.711 (42.9)	1.451 (67.1)	.33	96	19	.60	.54	87	136

Second-Generation JHU PIRC Trial Aggression Items

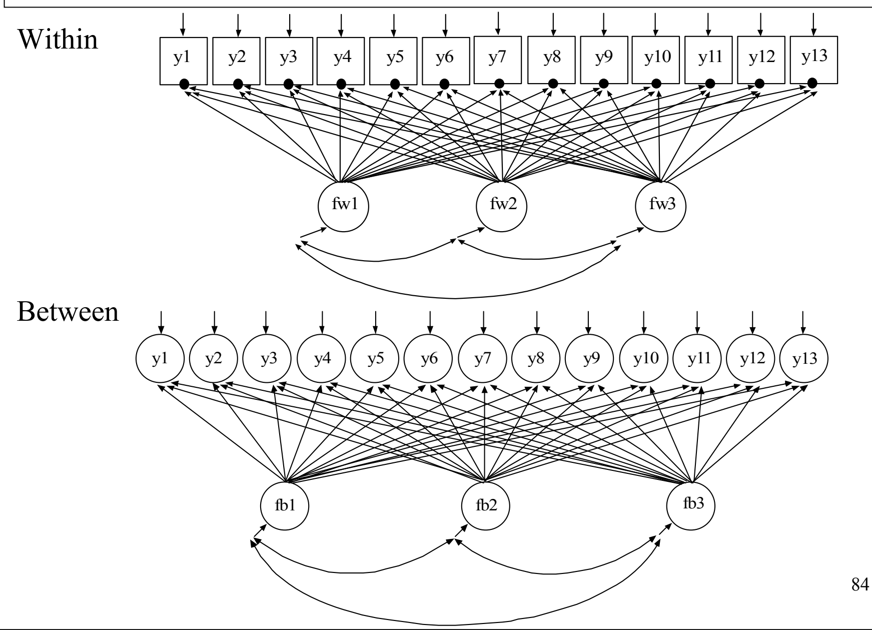
Item Distributions for Cohort 3: Fall 1st Grade (n=362 males in 27 classrooms)

	Almost Never (*scored as 1*)	*Rarely* (*scored as 2*)	*Sometimes* (*scored as 3*)	*Often* (*scored as 4*)	*Very Often* (*scored as 5*)	*Almost Always* (*scored as 6*)
Stubborn	42.5	21.3	18.5	7.2	6.4	4.1
Breaks Rules	37.6	16.0	22.7	7.5	8.3	8.0
Harms Others	69.3	12.4	9.40	3.9	2.5	2.5
Breaks Things	79.8	6.60	5.20	3.9	3.6	0.8
Yells at Others	61.9	14.1	11.9	5.8	4.1	2.2
Takes Others' Property	72.9	9.70	10.8	2.5	2.2	1.9
Fights	60.5	13.8	13.5	5.5	3.0	3.6
Harms Property	74.9	9.90	9.10	2.8	2.8	0.6
Lies	72.4	12.4	8.00	2.8	3.3	1.1
Talks Back to Adults	79.6	9.70	7.80	1.4	0.8	1.4
Teases Classmates	55.0	14.4	17.7	7.2	4.4	1.4
Fights With Classmates	67.4	12.4	10.2	5.0	3.3	1.7
Loses Temper	61.6	15.5	13.8	4.7	3.0	1.4

Hypothesized Aggressiveness Factors

- Verbal aggression
  - Yells at others
  - Talks back to adults
  - Loses temper
  - Stubborn
- Property aggression
  - Breaks things
  - Harms property
  - Takes others' property
  - Harms others
- Person aggression
  - Fights
  - Fights with classmates
  - Teases classmates

Two-Level Factor Analysis

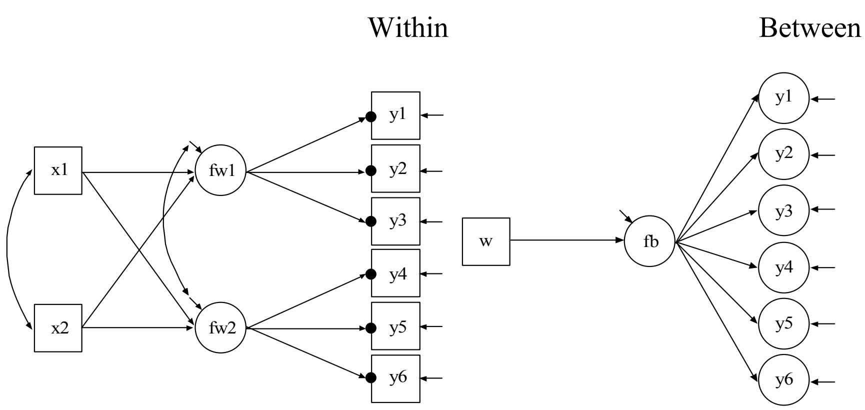


Promax Rotated Loadings

	Within-Level Loadings			Between-Level Loadings		
	1	**2**	**3**	**1**	**2**	**3**
Stubborn	0.07	**0.70**	0.05	-0.19	**1.03**	0.07
Breaks Rules	0.25	0.31	0.37	0.15	0.28	0.31
Harms Others	**0.52**	0.16	0.27	0.35	-0.20	**0.72**
Breaks Things	**0.84**	0.16	-0.01	**0.71**	0.01	0.41
Yells at Others	0.15	**0.64**	0.13	0.38	**0.74**	-0.01
Takes Others' Property	**0.57**	0.00	0.37	**0.86**	-0.04	0.12
Fights	0.20	0.21	**0.63**	0.09	0.03	**0.89**
Harms Property	**0.73**	0.21	0.10	**0.90**	-0.05	0.16
Lies	0.48	0.28	0.24	**0.86**	0.33	-0.21
Talks Back to Adults	0.29	**0.71**	0.23	0.41	0.58	-0.04
Teases Classmates	0.11	0.19	**0.62**	0.37	0.31	0.30
Fights With Classmates	0.10	0.31	**0.63**	-0.19	0.38	**0.88**
Loses Temper	0.12	**0.75**	0.04	0.17	**0.78**	0.12

Two-Level Factor Analysis With Covariates

Two-Level Factor Analysis With Covariates



87

Input For Two-Level Factor Analysis With Covariates

```
TITLE:      this is an example of a two-level CFA with
            continuous factor indicators with two factors on the
            within level and one factor on the between level
DATA:       FILE IS ex9.8.dat;
VARIABLE:   NAMES ARE y1-y6 x1 x2 w clus;
            WITHIN = x1 x2;
            BETWEEN = w;
            CLUSTER IS clus;
ANALYSIS:   TYPE IS TWOLEVEL;
MODEL:      %WITHIN%
            fw1 BY y1-y3;
            fw2 BY y4-y6;
            fw1 ON x1 x2;
            fw2 ON x1 x2;
            %BETWEEN%
            fb BY y1-y6;
            fb ON w;
```

88

Input For Monte Carlo Simulations For Two-Level Factor Analysis With Covariates

```
TITLE:          This is an example of a two-level CFA with
                continuous factor indicators with two
                factors on the within level and one factor
                on the between level
MONTECARLO:
                NAMES ARE y1-y6 x1 x2 w;
                NOBSERVATIONS = 1000;
                NCSIZES = 3;
                CSIZES = 40 (5) 50 (10) 20 (15);
                SEED = 58459;
                NREPS = 1;
                SAVE = ex9.8.dat;
                WITHIN = x1 x2;
                BETWEEN = w;
ANALYSIS:       TYPE = TWOLEVEL;
```


Input For Monte Carlo Simulations For Two-Level Factor Analysis With Covariates (Continued)

```
MODEL POPULATION:
                %WITHIN%
                x1-x2@1;
                fw1 BY y1@1 y2-y3*1;
                fw2 BY y4@1 y5-y6*1;
                fw1-fw2*1;
                y1-y6*1;
                fw1 ON x1*.5 x2*.7;
                fw2 ON x1*.7 x2*.5;

                %BETWEEN%
                [w@0]; w*1;
                fb BY y1@1 y2-y6*1;
                y1-y6*.3;
                fb*.5;
                fb ON w*1;
```

Input For Monte Carlo Simulations For Two-Level Factor Analysis With Covariates (Continued)

```
MODEL:
              %WITHIN%

              fw1 BY y1@1 y2-y3*1;
              fw2 BY y4@1 y5-y6*1;
              fw1-fw2*1;
              y1-y6*1;
              fw1 ON x1*.5 x2*.7;
              fw2 ON x1*.7 x2*.5;

              %BETWEEN%

              fb BY y1@1 y2-y6*1;
              y1-y6*.3;
              fb*.5;
              fb ON w*1;
OUTPUT:
              TECH8 TECH9;
```

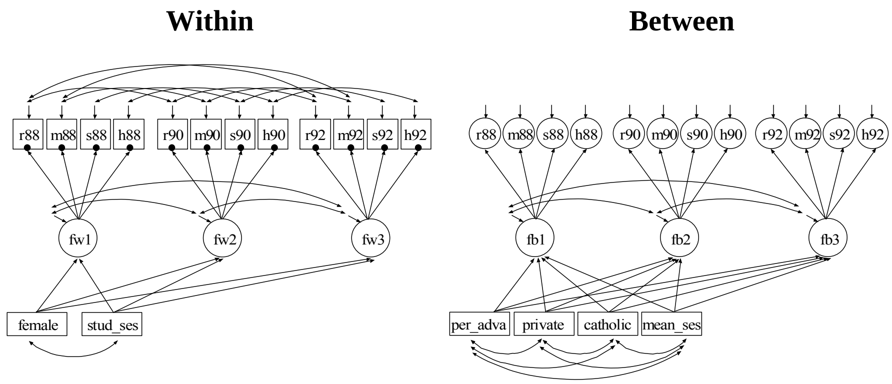
91

NELS Data

- The data—National Education Longitudinal Study (NELS:88)
  - Base year Grade 8—followed up in Grades 10 and 12
  - Students sampled within 1,035 schools—approximately 26 students per school, n = 14,217
  - Variables—reading, math, science, history-citizenship-geography, and background variables
- Data for the analysis—reading, math, science, history-citizenship-geography

92

NELS Two-Level Longitudinal Factor Analysis With Covariates



93

Input For NELS Two-Level Longitudinal Factor Analysis With Covariates

```
TITLE:      two-level factor analysis with covariates using the NELS
            data
DATA:       FILE = NELS.dat;
            FORMAT = 2f7.0 f11.4 12f5.2 11f8.2;
VARIABLE:   NAMES = id school f2pnlwt r88 m88 s88 h88 r90 m90 s90 h90
            r92 m92 s92 h92 stud_ses female per_mino urban size rural
            private mean_ses catholic stu_teac per_adva;
                !Variable Description
                !m88 = math IRT score in 1988
                !m90 = math IRT score in 1990
                !m92 = math IRT score in 1992
                !r88 = reading IRT score in 1988
                !r90 = reading IRT score in 1990
                !r92 = reading IRT score in 1992
```

94

Input For NELS Two-Level Longitudinal Factor Analysis With Covariates (Continued)

```
          !s88 = science IRT score in 1988
          !s90 = science IRT score in 1990
          !s92 = science IRT score in 1992
          !h88 = history IRT score in 1988
          !h90 = history IRT score in 1990
          !h92 = history IRT score in 1992
          !female = scored 1 vs 0
          !stud_ses = student family ses in 1990 (f1ses)
          !per_adva = percent teachers with an MA or higher
          !private = private school (scored 1 vs 0)
          !catholic = catholic school (scored 1 vs 0)
          !private = 0, catholic = 0 implies public school
MISSING = BLANK;
CLUSTER = school;
USEV = r88 m88 s88 h88 r90 m90 s90 h90 r92 m92 s92 h92
female stud_ses per_adva private catholic mean_ses;
WITHIN = female stud_ses;
BETWEEN = per_adva private catholic mean_ses;
```

95

Input For NELS Two-Level Longitudinal Factor Analysis With Covariates (Continued)

```
ANALYSIS:     TYPE = TWOLEVEL MISSING;
MODEL:        %WITHIN%
              fw1 BY r88-h88;
              fw2 BY r90-h90;
              fw3 BY r92-h92;
              r88 WITH r90; r90 WITH r92; r88 WITH r92;
              m88 WITH m90; m90 WITH m92; m88 WITH m92;
              s88 WITH s90; s90 WITH s92;
              h88 WITH h90; h90 WITH h92;
              fw1-fw3 ON female stud_ses;
              %BETWEEN%
              fb1 BY r88-h88;
              fb2 BY r90-h90;
              fb3 BY r92-h92;
              fb1-fb3 ON per_adva private catholic mean_ses;
OUTPUT:       SAMPSTAT STANDARDIZED TECH1 TECH8 MODINDICES;
```

96

Output Excerpts NELS Two-Level Longitudinal Factor Analysis With Covariates

Summary Of Data

```
    Number of patterns      15
    Number of clusters     913

Average cluster size 15.572

    Estimated Intraclass Correlations for the Y Variables
```

Variable	Intraclass Correlation	Variable	Intraclass Correlation	Variable	Intraclass Correlation
R88	0.067	M88	0.129	S88	0.100
H88	0.105	R90	0.076	M90	0.117
S90	0.110	H90	0.106	R92	0.073
M92	0.111	S92	0.099	H92	0.091

97

Output Excerpts NELS Two-Level Longitudinal Factor Analysis With Covariates (Continued)

Tests Of Model Fit

```
Chi-Square Test of Model Fit
              Value                                 4883.539*
              Degrees of Freedom                          146
              P-Value                                  0.0000
              Scaling Correction Factor                 1.046
                 for MLR

Chi-Square Test of Model Fit for the Baseline Model
              Value                               150256.855
              Degrees of Freedom                          202
              P-Value                                  0.0000

CFI/TLI
              CFI                                       0.968
              TLI                                       0.956

Loglikelihood
              H0 Value                            -487323.777
              H1 Value                            -484770.257
```

98

Output Excerpts NELS Two-Level Longitudinal Factor Analysis With Covariates (Continued)

```
Information Criteria

                Number of Free Parameters                   94
                Akaike (AIC)                        974835.554
                Bayesian (BIC)                      975546.400
                Sample-Size Adjusted BIC            975247.676
                  (n* = (n + 2) / 24)

RMSEA (Root Mean Square Error Of Approximation)
                Estimate                                 0.048
SRMR (Standardized Root Mean Square Residual
                Value for Between                        0.041
                Value for Within                         0.027
```

99

Output Excerpts NELS Two-Level Longitudinal Factor Analysis With Covariates (Continued)

Model Results

	Estimates	S.E.	Est./S.E.	Std	StdYX
Within Level					
FW1 BY					
R88	1.000	0.000	0.000	6.528	0.812
M88	0.940	0.010	94.856	6.135	0.804
S88	1.005	0.010	95.778	6.559	0.837
H88	1.041	0.011	97.888	6.796	0.837
FW2 BY					
R90	1.000	0.000	0.000	8.038	0.842
M90	0.911	0.008	109.676	7.321	0.838
S90	1.003	0.010	99.042	8.065	0.859
H90	0.939	0.008	113.603	7.544	0.855

100

Output Excerpts NELS Two-Level Longitudinal Factor Analysis With Covariates (Continued)

```
 FW3        BY
   R92              1.000      0.000      0.000      8.460      0.832
   M92              0.939      0.009    101.473      7.946      0.845
   S92              1.003      0.011     90.276      8.482      0.861
   H92              0.934      0.009    102.825      7.905      0.858
 FW1        ON
   FEMALE          -0.403      0.128     -3.150     -0.062     -0.031
   STUD_SES         3.378      0.096     35.264      0.517      0.418
FW2         ON
   FEMALE          -0.621      0.157     -3.945     -0.077     -0.039
   STUD_SES         4.169      0.110     37.746      0.519      0.420
FW3         ON
   FEMALE          -1.027      0.169     -6.087     -0.121     -0.064
   STUD_SES         4.418      0.122     36.124      0.522      0.422
```

101

Output Excerpts NELS Two-Level Longitudinal Factor Analysis With Covariates (Continued)

```
Residual Variances
   R88             22.021      0.383     57.464     22.021      0.341
   M88             20.618      0.338     61.009     20.618      0.354
   S88             18.383      0.323     56.939     18.383      0.299
   H88             19.805      0.370     53.587     19.805      0.300
   R90             26.546      0.491     54.033     26.546      0.291
   M90             22.756      0.375     60.748     22.756      0.298
   S90             23.150      0.383     60.516     23.150      0.262
   H90             21.002      0.403     52.124     21.002      0.270
   R92             31.821      0.617     51.562     31.821      0.308
   M92             25.213      0.485     52.018     25.213      0.285
   S92             25.155      0.524     47.974     25.155      0.259
   H92             22.479      0.489     46.016     22.479      0.265
   FW1             35.081      0.699     50.201      0.823      0.823
   FW2             53.079      1.005     52.806      0.822      0.822
   FW3             58.438      1.242     47.041      0.817      0.817
```

102

Output Excerpts NELS Two-Level Longitudinal Factor Analysis With Covariates (Continued)

```
Between Level
FB1          BY
   R88               1.000      0.000      0.000      1.952      0.933
   M88               1.553      0.070     22.138      3.031      0.979
   S88               1.061      0.058     18.255      2.071      0.887
   H88               1.065      0.053     19.988      2.078      0.814
FB2          BY
   R90               1.000      0.000      0.000      2.413      0.923
   M90               1.407      0.058     24.407      3.395      1.003
   S90               1.220      0.062     19.697      2.943      0.946
   H90               0.973      0.047     20.496      2.348      0.829
FB3          BY
   R92               1.000      0.000      0.000      2.472      0.947
   M92               1.435      0.065     22.095      3.546      0.997
   S92               1.160      0.065     17.889      2.868      0.938
   H92               0.963      0.041     23.244      2.380      0.871
```

103

Output Excerpts NELS Two-Level Longitudinal Factor Analysis With Covariates (Continued)

```
Between Level
FB1          ON
   PER_ADVA          0.217      0.292      0.742      0.111      0.024
   PRIVATE           0.303      0.344      0.883      0.155      0.042
   CATHOLIC         -0.696      0.277     -2.512     -0.357     -0.088
   MEAN_SES          2.513      0.206     12.185      1.288      0.672
FB2          ON
   PER_ADVA          0.280      0.338      0.828      0.116      0.025
   PRIVATE           0.453      0.392      1.155      0.188      0.051
   CATHOLIC         -0.538      0.334     -1.609     -0.223     -0.055
   MEAN_SES          3.054      0.239     12.805      1.266      0.660
FB3          ON
   PER_ADVA          0.473      0.375      1.261      0.192      0.041
   PRIVATE           0.673      0.435      1.547      0.272      0.074
   CATHOLIC         -0.206      0.372     -0.554     -0.084     -0.021
   MEAN_SES          3.142      0.258     12.169      1.271      0.663
```

104

Output Excerpts NELS Two-Level Longitudinal Factor Analysis With Covariates (Continued)

Residual Variances

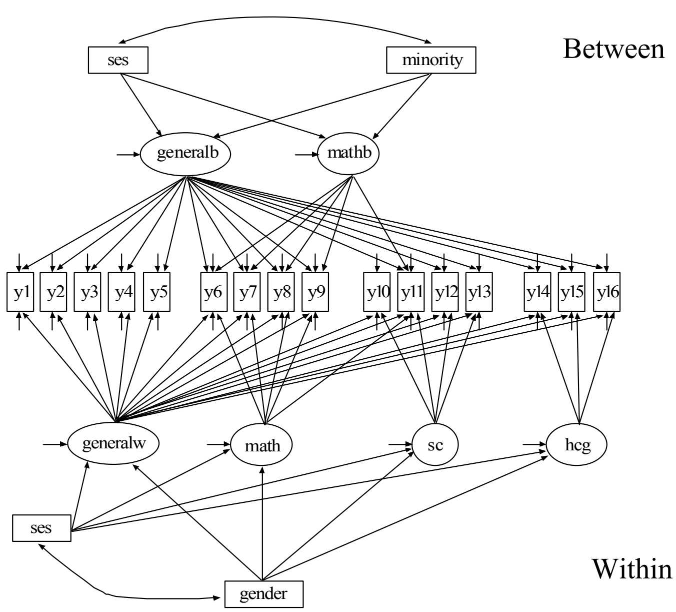
R88	0.564	0.104	5.437	0.564	0.129
M88	0.399	0.093	4.292	0.399	0.042
S88	1.160	0.126	9.170	1.160	0.213
H88	2.203	0.203	10.839	2.203	0.338
R90	1.017	0.160	6.352	1.017	0.149
M90	-0.068	0.055	-1.225	-0.068	-0.006
S90	1.025	0.172	5.945	1.025	0.106
H90	2.518	0.216	11.636	2.518	0.313
R92	0.706	0.182	3.886	0.706	0.104
M92	0.076	0.076	1.000	0.076	0.006
S92	1.120	0.190	5.901	1.120	0.120
H92	1.810	0.211	8.599	1.810	0.242
FB1	1.979	0.245	8.066	0.520	0.520
FB2	3.061	0.345	8.875	0.526	0.526
FB3	3.010	0.409	7.363	0.493	0.493

Multiple-Group, Two-Level Factor Analysis With Covariates

NELS Data

- The data—National Education Longitudinal Study (NELS:88)
  - Base year Grade 8—followed up in Grades 10 and 12
  - Students sampled within 1,035 schools—approximately 26 students per school
  - Variables—reading, math, science, history-citizenship-geography, and background variables
- Data for the analysis—reading, math, science, history-citizenship-geography, gender, individual SES, school SES, and minority status, n = 14,217 with 913 schools (clusters)

107



108

Input For NELS:88 Two-Group, Two-Level Model For Public And Catholic Schools

```
TITLE:     NELS:88 with listwise deletion
           disaggregated model for two groups, public and
           catholic schools
DATA:      FILE IS EX831.DAT;;
VARIABLE:  NAMES = ses y1-y16 gender cluster minority group;
           CLUSTER = cluster;
           WITHIN = gender;
           BETWEEN = minority;
           GROUPING = group(1=public 2=catholic);
DEFINE:    minority = minority/5;
ANALYSIS:  TYPE = TWOLEVEL;
           H1ITER = 2500;
           MITER = 1000;
```


Input For NELS:88 Two-Group, Two-Level Model For Public And Catholic Schools (Continued)

```
MODEL:     %WITHIN%
                 generalw BY y1* y2-y6 y8-y16 y7@1;
                 mathw BY y6* y8* y9* y11 y7@1;
                 scw BY y10 y11*.5 y12*.3 y13*.2;
                 hcgw BY y14*.7 y16*2 y15@1;
                 generalw WITH mathw-hcgw@0;
                 mathw WITH scw-hcgw@0;
                 scw WITH hcgw@0;
                 generalw mathw scw hcgw ON gender ses;
           %BETWEEN%
                 generalb BY y1* y2-y6 y8-y16 y7@1;
                 mathb BY y6* y8 y9 y11 y7@1;
                 y1-y16@0;
                 generalb WITH mathb@0;
                 generalb mathb ON ses minority;
```

Output Excerpts NELS:88 Two-Group, Two-Level Model For Public And Catholic Schools

Summary Of Data

```
Group PUBLIC
   Number of clusters        195
     Size (s) Cluster ID with Size s
        1        68114   68519
        2        72872
        7        72765
        8        45991   72012
        9        68071
       10         7298   72187
       11        72463    7105   72405
       12        24083   68971    7737   68390
       13        45861   72219   72049
       14        68511   72148   72175   72176  25464
       15        68023   25071   68748   45928   7915   78324
       16        45362    7403   72415   77204  77219   72456
       17        45502   68487   45824    7203  24948    7829   72612    7892
                 25835    7591   68155   68295
```


Output Excerpts NELS:88 Two-Group, Two-Level Model For Public And Catholic Schools (Continued)

```
       18        72133   25580   24910   68614  25074   72990   68328   25404
                  7348
       19         7671   68662   68671   45385   7438    7332   25615   72799
                 68340   72956   25642   25658  24856   78283   68030
       20        72617   72715    7211   25422   7330   72292   72060   72993
                  7451   68461   78162   78232  72170   25130
       21        45394    7193   68180   24589   7205   25894   25958   68391
                 77254   77634   68448   45271   7584   25227   78598
       22        68254   68397   68648   72768   7192    7117    7119   68753
                 24813
       23        68456   25361    7157   25702  25804   45620   24858    7658
                 25163   45041   77351   45183  77684   78101   68788   68817
                  7792   78311   68048   68453
       24        77222   24053    7000   77403  24138   68297   78011   25536
                  7778   72042   25360   25977  45747    7616   78886
       25        68906   68720   25354   68427  72833   77268    7269   68520
                 77537   72075
       26        72973   45555   24828   68315  45087   25328   77710   25848
       27        45831   25618   68652   72080  45900   25208   45452    7103
```

Output Excerpts NELS:88 Two-Group, Two-Level Model For Public And Catholic Schools (Continued)

```
        28        25666   68809   25076   25224   68551
        30         7343   45978   25722   45924
        31        77109    7230   68855
        32        25178
        33        45330   25745   25825
        35        25667
        36        72129
        37        25834
        38        45287
        39        45197    7090
        43        45366
```

113

Output Excerpts NELS:88 Two-Group, Two-Level Model For Public And Catholic Schools (Continued)

```
Group PUBLIC
  Number of clusters          195
  Average cluster size 21.292
  Estimated Intraclass Correlations for the Y Variables
             Intraclass                Intraclass               Intraclass
   Variable  Correlation  Variable  Correlation  Variable  Correlation

   Y1           .111      Y7           .100      Y12          .115
   Y2           .105      Y8           .124      Y13          .185
   Y3           .213      Y9           .069      Y14          .094
   Y4           .160      Y10          .147      Y15          .132
   Y5           .081      Y11          .105      Y16          .159
   Y6           .159
```

114

Output Excerpts NELS:88 Two-Group, Two-Level Model For Public And Catholic Schools (Continued)

```
Group CATHOLIC

  Number of clusters           40
  Average cluster size 26.016

  Estimated Intraclass Correlations for the Y Variables
```

Variable	Intraclass Correlation	Variable	Intraclass Correlation	Variable	Intraclass Correlation
Y1	.010	Y7	.029	Y12	.056
Y2	.039	Y8	.061	Y13	.176
Y3	.180	Y9	.056	Y14	.078
Y4	.091	Y10	.079	Y15	.071
Y5	.055	Y11	.056	Y16	.154
Y6	.118				

115

Output Excerpts NELS:88 Two-Group, Two-Level Model For Public And Catholic Schools (Continued)

Tests Of Model Fit

```
Loglikelihood
              Value                              1716.922*
              Degrees of Freedom                      575
              P-Value                              0.0000
              Scaling Correction Factor             0.872
                 for MLR

Chi-Square Test of Model
              Value                             35476.471
              Degrees of Freedom                      608
              P-Value                              0.0000

CFI/TLI
              CFI                                   0.967
              TLI                                   0.965

Loglikelihood
              H0 Value                        -130332.921
              H1 Value                        -129584.053
```

116

Output Excerpts NELS:88 Two-Group, Two-Level Model For Public And Catholic Schools (Continued)

	Estimates	S.E.	Est./S.E.	Std	StdYX
Group Public					
Within Level					
GENERALW ON					
GENDER	-0.193	0.029	-6.559	-0.256	-0.128
SES	0.233	0.016	14.269	0.309	0.279
MATHW ON					
GENDER	0.266	0.025	10.534	0.510	0.255
SES	0.054	0.014	3.879	0.103	0.093
SCW ON					
GENDER	0.452	0.032	14.005	0.961	0.480
SES	0.018	0.015	1.244	0.039	0.035
HCGW ON					
GENDER	0.152	0.023	6.588	0.681	0.341
SES	0.002	0.007	0.239	0.007	0.007

117

Output Excerpts NELS:88 Two-Group, Two-Level Model For Public And Catholic Schools (Continued)

	Estimates	S.E.	Est./S.E.	Std	StdYX
Group Catholic					
Within Level					
GENERALW ON					
GENDER	-0.294	0.059	-5.000	-0.403	-0.201
SES	0.169	0.021	7.892	0.232	0.193
MATHW ON					
GENDER	0.332	0.051	6.478	0.627	0.313
SES	-0.030	0.017	-1.707	-0.056	-0.047
SCW ON					
GENDER	0.555	0.063	8.860	1.226	0.613
SES	-0.022	0.014	-1.592	-0.049	-0.041
HCGW ON					
GENDER	0.160	0.029	5.610	0.785	0.392
SES	0.001	0.007	0.089	0.003	0.002

118

Output Excerpts NELS:88 Two-Group, Two-Level Model For Public And Catholic Schools (Continued)

Group Public
Between Level

	Estimates	S.E.	Est./S.E.	Std	StdYX
GENERALB ON					
SES	0.505	0.079	6.390	1.244	0.726
MINORITY	-0.217	0.088	-2.452	-0.534	-0.188
MATHB ON					
SES	0.198	0.070	2.825	0.984	0.574
MINORITY	-0.031	0.087	-0.354	-0.153	-0.054
GENERALB WITH					
MATHB	0.000	0.000	0.000	0.000	0.000
Intercepts					
GENERALB	0.000	0.000	0.000	0.000	0.000
MATHB	0.000	0.000	0.000	0.000	0.000

119

Output Excerpts NELS:88 Two-Group, Two-Level Model For Public And Catholic Schools (Continued)

Group Catholic
Between Level

	Estimates	S.E.	Est./S.E.	Std	StdYX
GENERALB ON					
SES	0.262	0.067	3.929	0.975	0.538
MINORITY	-0.327	0.069	-4.707	-0.216	-0.573
MATHB ON					
SES	0.205	0.071	2.901	0.746	0.412
MINORITY	-0.213	0.095	-2.241	-0.778	-0.367
GENERALB WITH					
MATHB	0.000	0.000	0.000	0.000	0.000
Intercepts					
GENERALB	0.466	0.163	2.854	1.734	1.734
MATHB	0.573	0.177	3.239	2.087	2.087

120

Further Readings On Two-Level Factor Analysis

Harnqvist, K., Gustafsson, J.E., Muthén, B, & Nelson, G. (1994). Hierarchical models of ability at class and individual levels. Intelligence, 18, 165-187. (#53)

Hox, J. (2002). Multilevel analysis. Techniques and applications. Mahwah, NJ: Lawrence Erlbaum

Longford, N. T., & Muthén, B. (1992). Factor analysis for clustered observations. Psychometrika, 57, 581-597. (#41)

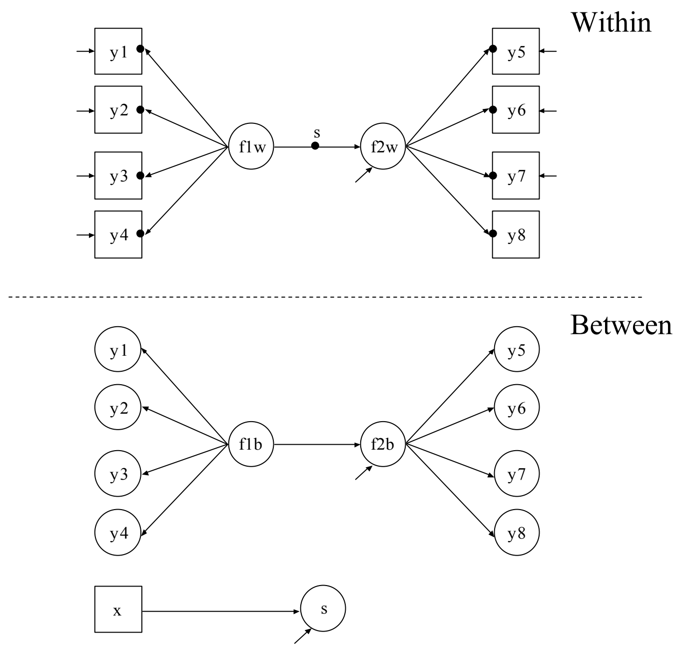
Muthén, B. (1989). Latent variable modeling in heterogeneous populations. Psychometrika, 54, 557-585. (#24)

Muthén, B. (1990). Mean and covariance structure analysis of hierarchical data. Paper presented at the Psychometric Society meeting in Princeton, NJ, June 1990. UCLA Statistics Series 62. (#32)

Muthén, B. (1991). Multilevel factor analysis of class and student achievement components. Journal of Educational Measurement, 28, 338-354. (#37)

Muthén, B. (1994). Multilevel covariance structure analysis. In J. Hox & I. Kreft (eds.), Multilevel Modeling, a special issue of Sociological Methods & Research, 22, 376-398. (#55)

Two-Level SEM: Random Slopes For Regressions Among Factors



123

Input For A Two-Level SEM With A Random Slope

```
TITLE:       a twolevel SEM with a random slope

DATA:        FILE = etaeta3.dat;

VARIABLE:    NAMES ARE y1-y8 x clus;
             CLUSTER = clus;
             BETWEEN = x;

ANALYSIS:    TYPE = TWOLEVEL RANDOM MISSING;
             ALGORITHM = INTEGRATION;
```

124

Input For A Two-Level SEM With A Random Slope (Continued)

```
MODEL:      %WITHIN%
            f1w BY y1@1
            y2 (1)
            y3 (2)
            y4 (3);
            f2w BY y5@1
            y6 (4)
            y7 (5)
            y8 (6);
            s | f2w ON f1w;

            %BETWEEN%
            f1b BY y1@1
            y2 (1)
            y3 (2)
            y4 (3);
            f2b BY y5@1
            y6 (4)
            y7 (5)
            y8 (6);
            f2b ON f1b;
            s ON x;
OUTPUT:     TECH1 TECH8;
```

125

Output Excerpts Two-Level SEM With A Random Slope

Tests Of Model Fit

```
Loglikelihood

             H0 Value                          -12689.557

Information Criteria

             Number of Free Parameters                 30
             Akaike (AIC)                       25439.114
             Bayesian (BIC)                     25585.122
             Sample-Size Adjusted BIC           25489.843
                (n* = (n + 2) / 24)
```

126

Output Excerpts Two-Level SEM With A Random Slope (Continued)

Model Results

```
                    Estimates     S.E.   Est./S.E.
 Within Level

  F1W       BY
     Y1                1.000     0.000      0.000
     Y2                0.992     0.035     28.597
     Y3                0.978     0.041     23.593
     Y4                1.001     0.037     26.884
  F2W       BY
     Y5                1.000     0.000      0.000
     Y6                0.978     0.028     34.417
     Y7                1.049     0.030     35.174
     Y8                1.008     0.026     38.090
  F1W       WITH
     F2W               0.000     0.000      0.000
```

127

Output Excerpts Two-Level SEM With A Random Slope (Continued)

```
                    Estimates     S.E.   Est./S.E.
  Variances
     F1W               1.016     0.082     12.325
     F2W               0.580     0.063      9.144
  Residual Variances
     Y1                0.979     0.063     15.517
     Y2                0.949     0.056     16.854
     Y3                1.052     0.060     17.406
     Y4                0.971     0.053     18.174
     Y5                1.039     0.057     18.187
     Y6                1.062     0.058     18.292
     Y7                0.941     0.058     16.191
     Y8                1.076     0.060     17.835
```

128

Output Excerpts Two-Level SEM With A Random Slope (Continued)

		Estimates	S.E.	Est./S.E.
Between Level				
F1B	BY			
Y1		1.000	0.000	0.000
Y2		0.992	0.035	28.597
Y3		0.978	0.041	23.593
Y4		1.001	0.037	26.884
F2B	BY			
Y5		1.000	0.000	0.000
Y6		0.978	0.028	34.417
Y7		1.049	0.030	35.174
Y8		1.008	0.026	38.090
F2B	ON			
F1B		0.180	0.080	2.248

129

Output Excerpts Two-Level SEM With A Random Slope (Continued)

		Estimates	S.E.	Est./S.E.
S	ON			
X		0.999	0.082	12.150
Intercepts				
Y1		-0.099	0.063	-1.560
Y2		-0.011	0.064	-0.175
Y3		-0.069	0.067	-1.034
Y4		-0.001	0.065	-0.017
Y5		0.030	0.062	0.475
Y6		-0.008	0.064	-0.129
Y7		0.041	0.064	0.635
Y8		0.002	0.071	0.035
S		0.777	0.073	10.604
Variances				
F1B		0.568	0.096	5.900
Residual Variances				
F2B		0.237	0.056	4.211
S		0.420	0.088	4.756

130

Multilevel Estimation, Testing, Modification, And Identification

Estimators

- Muthén's limited information estimator (MUML) – random intercepts
  - ESTIMATOR = MUML
  - Muthén's limited information estimator for unbalanced data
  - Maximum likelihood for balanced data
- Full-information maximum likelihood (FIML) – random intercepts and random slopes
  - ESTIMATOR = ML, **MLR**, MLF
  - Full-information maximum likelihood for balanced and unbalanced data
  - Robust maximum likelihood estimator
  - MAR missing data
  - Asparouhov and Muthén

Multilevel Estimation, Testing, Modification, And Identification (Continued)

Tests of Model Fit

- MUML – chi-square, robust chi-square, CFI, TLI, RMSEA, and SRMR
- FIML – chi-square, robust chi-square, CFI, TLI, RMSEA, and SRMR
- FIML with random slopes – no tests of model fit

Model Modification

- MUML – modification indices not available
- FIML – modification indices available

Model identification is the same as for CFA for both the between and within parts of the model.

Practical Issues Related To The Analysis Of Multilevel Data

Size Of The Intraclass Correlation

- Small intraclass correlations can be ignored but important information about between-level variability may be missed by conventional analysis
- The importance of the size of an intraclass correlation depends on the size of the clusters
- Intraclass correlations are attenuated by individual-level measurement error
- Effects of clustering not always seen in intraclass correlations

133

Practical Issues Related To The Analysis Of Multilevel Data (Continued)

Within-Level And Between-Level Variables

- Variables measured on the individual level can be used in both the between and within parts of the model or only in the within part of the model (WITHIN=)
- Variables measured on the between level can be used only in the between part of the model (BETWEEN=)

Sample Size

- There should be at least 30-50 between-level units (clusters)
- Clusters with only one observation are allowed

134

Steps In SEM Multilevel Analysis For Continuous Outcomes

1) Explore SEM model using the sample covariance matrix from the total sample
2) Estimate the SEM model using the pooled-within sample covariance matrix with sample size n - G
3) Investigate the size of the intraclass correlations and DEFF's
4) Explore the between structure using the estimated between covariance matrix with sample size G
5) Estimate and modify the two-level model suggested by the previous steps

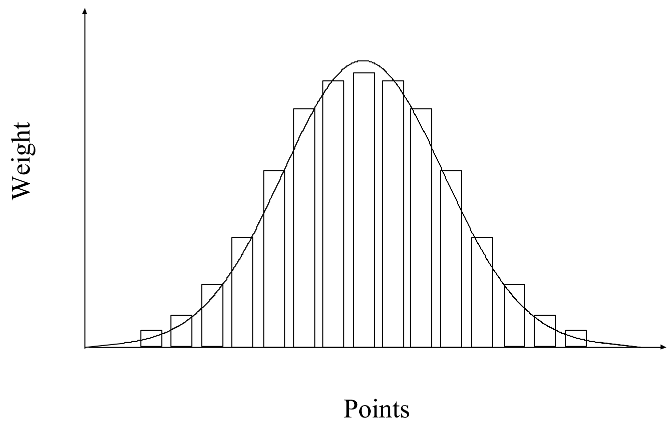
Muthén, B. (1994). Multilevel covariance structure analysis. In J. Hox & I. Kreft (eds.), Multilevel Modeling, a special issue of Sociological Methods & Research, 22, 376-398. (#55)

135

Technical Aspects Of Multilevel Modeling

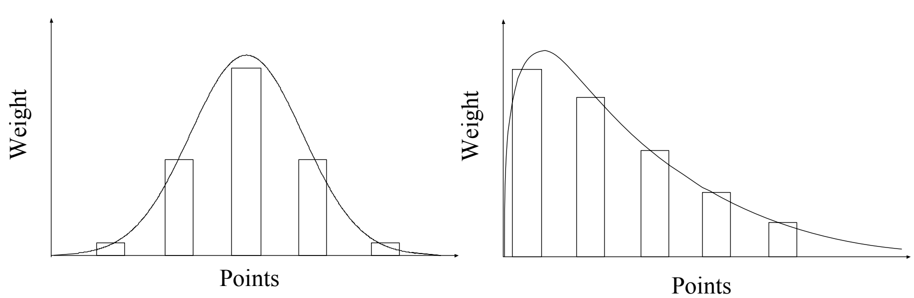
136

Numerical Integration With A Normal Latent Variable Distribution



Fixed weights and points

Non-Parametric Estimation Of The Random Effect Distribution



Estimated weights and points
(class probabilities and class means)

Numerical Integration

Numerical integration is needed with maximum likelihood estimation when the posterior distribution for the latent variables does not have a closed form expression. This occurs for models with categorical outcomes that are influenced by continuous latent variables, for models with interactions involving continuous latent variables, and for certain models with random slopes such as multilevel mixture models.

When the posterior distribution does not have a closed form, it is necessary to integrate over the density of the latent variables multiplied by the conditional distribution of the outcomes given the latent variables. Numerical integration approximates this integration by using a weighted sum over a set of integration points (quadrature nodes) representing values of the latent variable.

Numerical Integration (Continued)

Numerical integration is computationally heavy and thereby time-consuming because the integration must be done at each iteration, both when computing the function value and when computing the derivative values. The computational burden increases as a function of the number of integration points, increases linearly as a function of the number of observations, and increases exponentially as a function of the dimension of integration, that is, the number of latent variables for which numerical integration is needed.

Practical Aspects Of Numerical Integration

- Types of numerical integration available in Mplus with or without adaptive quadrature
  - Standard (rectangular, trapezoid) – default with 15 integration points per dimension
  - Gauss-Hermite
  - Monte Carlo
- Computational burden for latent variables that need numerical integration
  - One or two latent variables — Light
  - Three to five latent variables — Heavy
  - Over five latent variables — Very heavy

141

Practical Aspects Of Numerical Integration (Continued)

- Suggestions for using numerical integration
  - Start with a model with a small number of random effects and add more one at a time
  - Start with an analysis with TECH8 and MITERATIONS=1 to obtain information from the screen printing on the dimensions of integration and the time required for one iteration and with TECH1 to check model specifications
  - With more than 3 dimensions, reduce the number of integration points to 5 or 10 or use Monte Carlo integration with the default of 500 integration points
  - If the TECH8 output shows large negative values in the column labeled ABS CHANGE, increase the number of integration points to improve the precision of the numerical integration and resolve convergence problems

142

Technical Aspects Of Numerical Integration

Maximum likelihood estimation using the EM algorithm computes in each iteration the posterior distribution for normally distributed latent variables f,

$$[\,f\,|\,y\,] = [\,f\,]\,[\,y\,|\,f\,]\,/\,[\,y\,], \tag{97}$$

where the marginal density for $[y]$ is expressed by integration

$$[\,y\,] = \int [\,f\,]\,[\,y\,|\,f\,]\,df. \tag{98}$$

- Numerical integration is not needed for normally distributed y - the posterior distribution is normal

143

Technical Aspects Of Numerical Integration (Continued)

- Numerical integration needed for:
  - Categorical outcomes u influenced by continuous latent variables f, because $[u]$ has no closed form
  - Latent variable interactions f x x, f x y, f_1 x f_2, where $[y]$ has no closed form, for example

$$[\,y\,] = \int [\,f_1, f_2\,]\,[\,y|f_1, f_2, f_1 f_2\,]\,df_1\,df_2 \tag{99}$$

  - Random slopes, e.g. with two-level mixture modeling

Numerical integration approximates the integral by a sum

$$[\,y\,] = \int [\,f\,]\,[\,y\,|\,f\,]\,df = \sum_{k=1}^{K} w_k\,[\,y\,|\,f_k\,] \tag{100}$$

144

Multivariate Approach To Multilevel Modeling

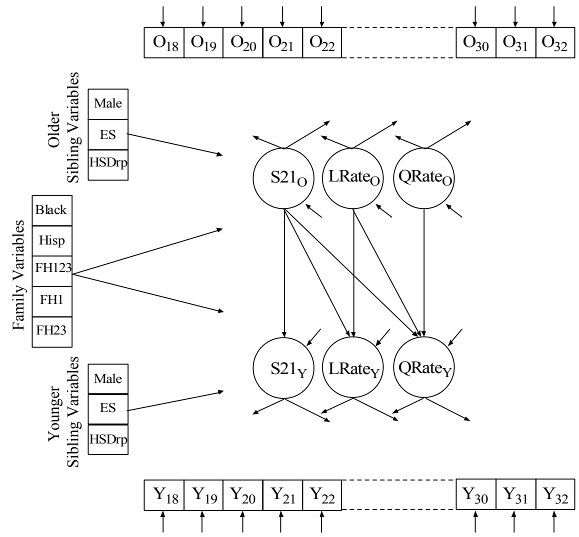
145

Multivariate Modeling Of Family Members

- Multilevel modeling: clusters independent, model for between- and within-cluster variation, units within a cluster statistically equivalent
- Multivariate approach: clusters independent, model for all variables for each cluster unit, different parameters for different cluster units.
  - Used in latent variable growth modeling where the cluster units are the repeated measures over time
  - Allows for different cluster sizes by missing data techniques
  - More flexible than the multilevel approach, but computationally convenient only for applications with small cluster sizes (e.g. twins, spouses)

146

Figure 1. A Longitudinal Growth Model of Heavy Drinking for Two-Sibling Families



Source: Khoo, S.T. & Muthen, B. (2000). Longitudinal data on families: Growth modeling alternatives. Multivariate Applications in Substance Use Research, J. Rose, L. Chassin, C. Presson & J. Sherman (eds.), Hillsdale, N.J.: Erlbaum, pp. 43-78.

147

Three-Level Modeling As Single-Level Analysis

Doubly multivariate:

- Repeated measures in wide, multivariate form
- Siblings in wide, multivariate form

It is possible to do four-level by TYPE = TWOLEVEL, for instance families within geographical segments

148

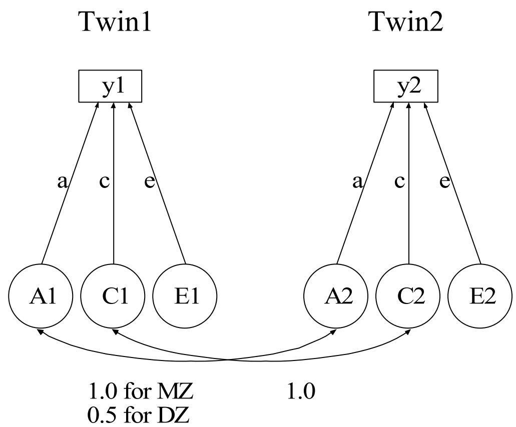
Input For Multivariate Modeling Of Family Data

```
TITLE:      Multivariate modeling of family data
            one observation per family
DATA:       FILE IS multi.dat;
VARIABLE:   NAMES ARE o18-o32 y18-y32 omale oes ohsdrop ymale yoes
            yhsdrop black hisp fh123 fh1 fh23;
MODEL:      s21o lrateo qrateo | o18@0 o19@1 o20@2 o21@3 o22@4
            o23@5 o24@6 o25@7 o26@8 o27@9 o28@10 o29@11 o30@12
            o31@13 o32@14;
            s21y lratey qratey | y18@0 y19@1 y20@2 y21@3 y22@4 y23@5
            y24@6 y25@7 y26@8 y27@9 y28@10 y29@11 y30@12
            y31@13 y32@14;
            s12o ON omale oes ohsdrop black hisp fh123 fh1 fh23;
            221y ON ymale yes yhsdrop black hisp fh123 fh1 fh23;
            s21y ON s21o;
            lratey ON s21o lrateo;
            qratey ON s21o lrateo qrateo;
```

149

Twin Modeling

150



Neale & Cardon (1992)
Prescott (2004)

151

Multilevel Growth Models

152