Running MLwiN from within Stata: the `runmlwin` command

e-Stat meeting
University of Bristol
7th April 2011

George Leckie and Chris Charlton
Centre for Multilevel Modelling
University of Bristol

INTRODUCTION

Existing multilevel modelling commands in Stata

- Stata provide the `xtmixed`, `xtmelogit` and `xtmepoisson` commands to fit multilevel models
  - Limited range of models can be specified
  - Computationally quite slow to fit models
- Sophia Rabe-Hesketh and Anders Skrondal provide the `gllamm` command
  - Wide range of models can be specified
  - Computationally slow to fit models
- Other user-written multilevel modelling commands include: `hlm`, `realcomimpute`, `runmplus`, `sabre`, `winbugs`

Multilevel modelling in MLwiN

1. Estimation of multilevel models for continuous, binary, ordered categorical, unordered categorical and count data

2. Fast estimation via classical and Bayesian methods

3. Estimation of multilevel models for cross-classified and multiple membership non-hierarchical data structures

4. Estimation of multilevel multivariate response models, multilevel spatial models, multilevel measurement error models, multilevel multiple imputation models and multilevel factor models

RAUDENBUSH (1993) CROSS-CLASSIFED MODELLING EXAMPLE

Scottish neighbourhood study on child educational attainment

- Scottish neighbourhood study on child educational attainment
- 2310 students nested within 17 schools and 524 neighbourhoods
- First analysed by Garner and Raudenbush (1991)
- Re-analysed by Rabe-Hesketh and Skrondal (2008), Raudenbush (1993), Raudenbush and Bryk (2002) and others

Journal of Educational Statistics
Winter 1993, Vol. 18, No. 4, pp. 321–349

A Crossed Random Effects Model for Unbalanced Data With Applications in Cross-Sectional and Longitudinal Research

Stephen W. Raudenbush
Michigan State University

Key words: *hierarchical models, maximum likelihood, covariance components*

Hierarchical linear models have found widespread application when the data have a nested structure—for example, when students are nested within classrooms (a two-level nested structure) or students are nested within classrooms and classrooms are nested within schools (a three-level nested structure). Often, however, the data will have a more complex nested structure. In Example 1, students are nested within both neighborhoods and schools; however, a school can draw students from multiple neighborhoods, and a neighborhood can send students to multiple schools. In Example 2, children are nested within classrooms during the first year of the study; however, each child finds himself or herself with a new teacher and a new set of classmates

TABLE 1

Organization of data in the Scotland neighborhood study: Numbers of observations in each neighborhood by school cell

	School																	
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Totals
Neighborhood																		
26																	5	5
27																	1	1
29															1		8	9
30																	2	2
31															1		1	2
32															1		5	6
33															1		2	3
35																	3	3
36																	2	2
38															1		4	5
.																		.
.																		.
.																		.
251														4		1		5
252													1	3			1	5
253														3				3
256																	2	2
258												4						4
259											6	1					2	9
260											6							6
261											3						3	6
263											14					1	1	16
.																		.
.																		.
.																		.
793									6									6
794		1				1			11									13
795			2						1									3
796			9						1									10
797			3					1										4
798								10										10
799			1							1								2
800			3	1														4
801			1	1														2
803										3								3
Totals	146	22	146	159	155	101	286	112	136	133	92	190	111	154	91	102	174	2310

TABLE 3
Descriptive statistics from the Scotland neighborhood study

Variable	*n*	*m*	*SD*
(a) Neighborhood level			
Social deprivation	524	0.037	0.622
(b) Student level			
Total attainment	2310	0.093	1.00
Primary 7 VRQ	2310	0.506	10.65
Primary 7 reading	2310	−0.044	13.89
Dad occupation	2310	−0.464	11.78
Dad education (1 = Schooling past 15; 0 = No)	2310	0.215	0.41
Mom education (1 = Schooling past 15; 0 = No)	2310	0.248	0.43
Dad unemployment (1 = Yes; 0 = No)	2310	0.109	0.31
Sex (1 = M; 0 = F)	2310	0.480	0.50

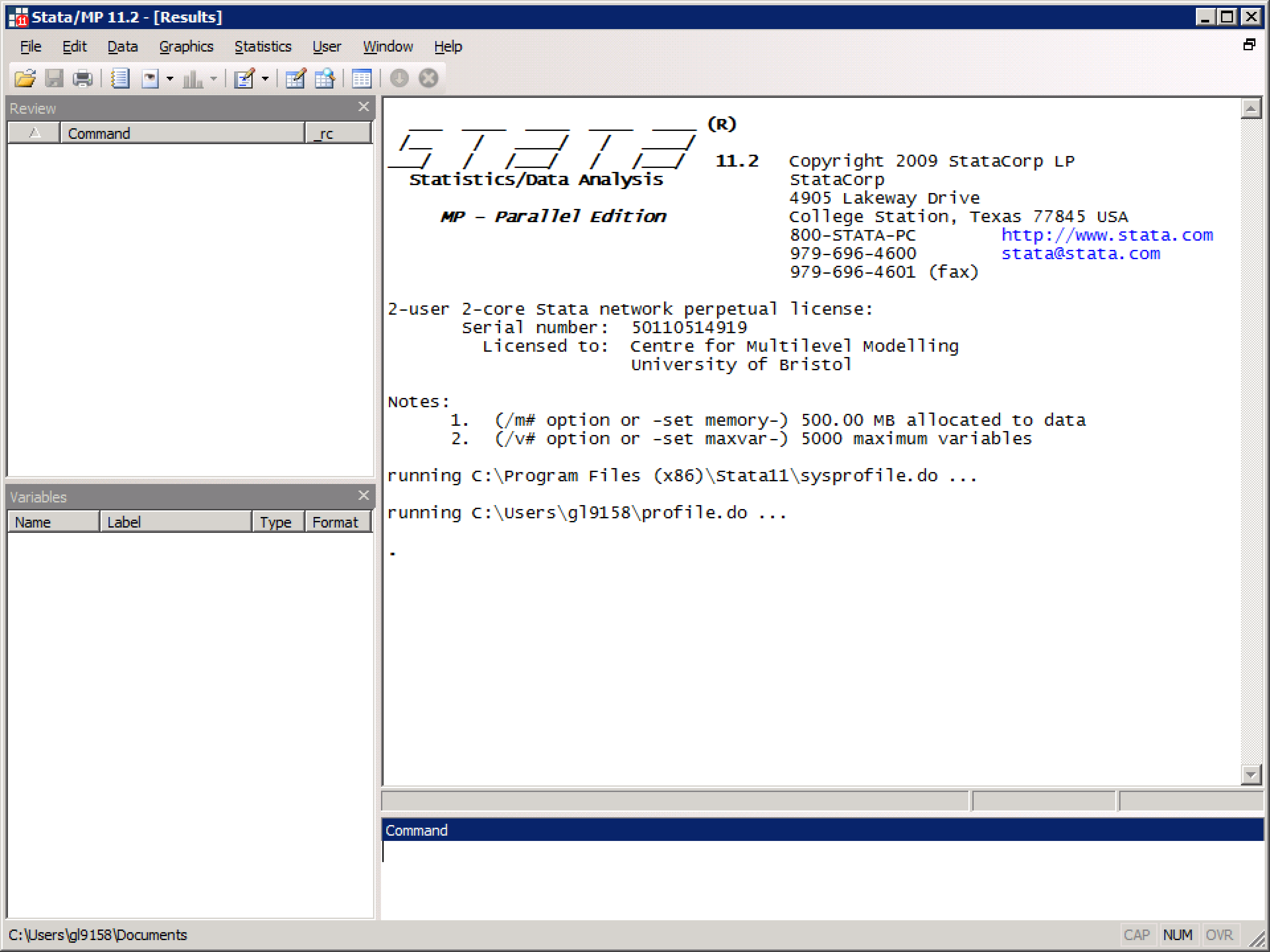
TABLE 4

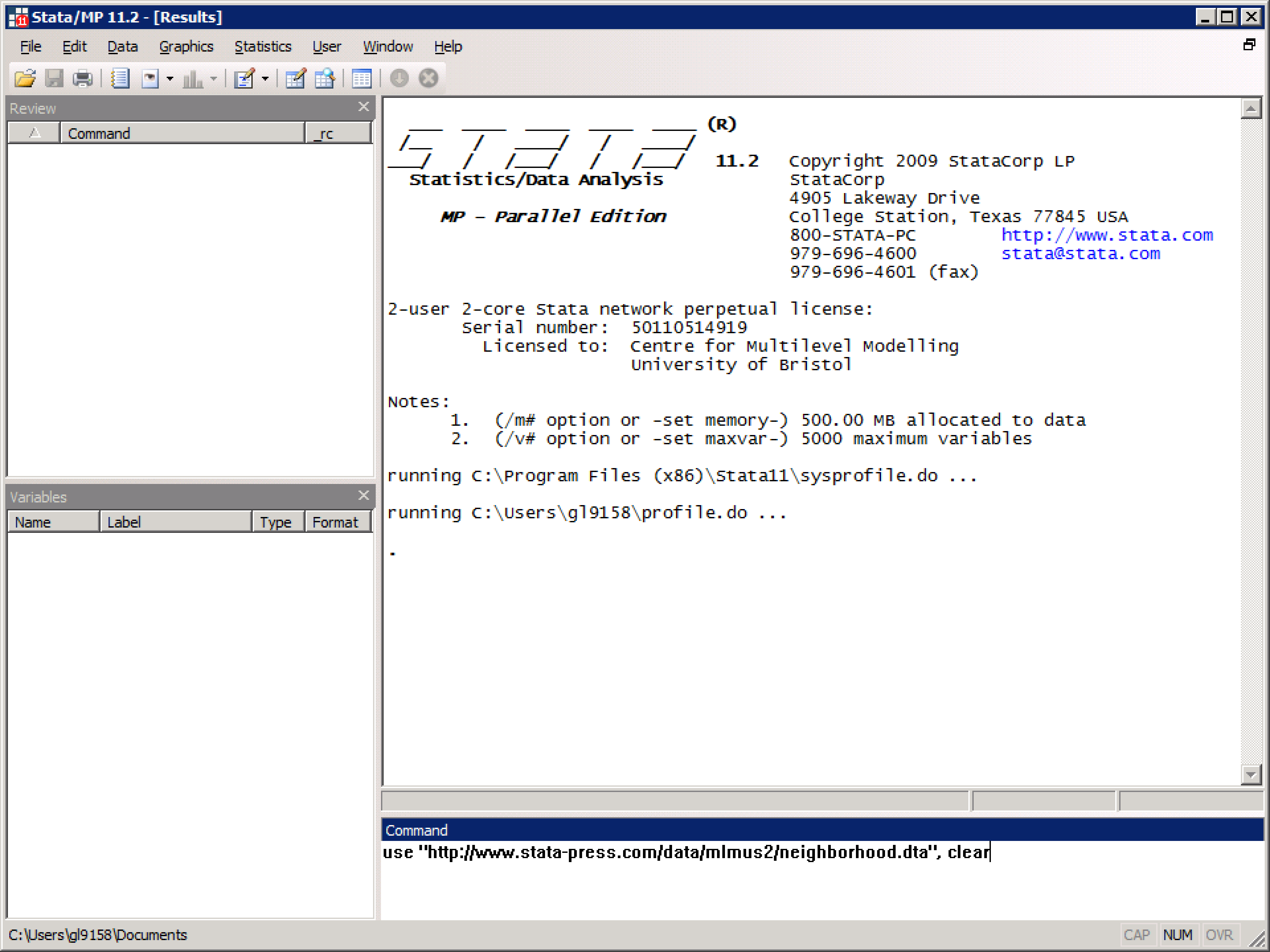
Modeling results for Scotland neighborhood data

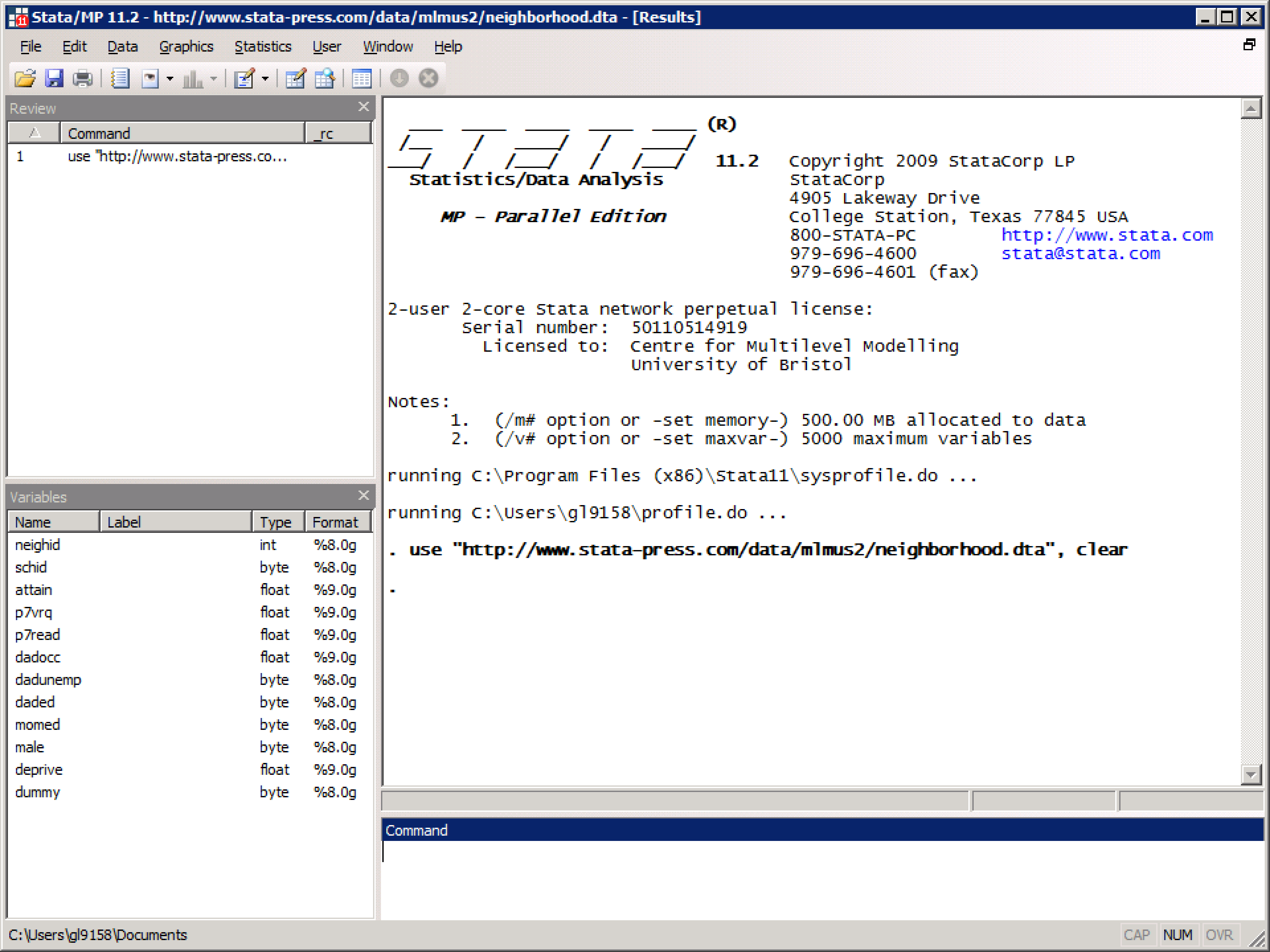
(a) Fixed effects	Model 1: School variance ignored			Model 2: School variance estimated			Model 3: Fitted model		
Predictor	Coefficient	*SE*	*t* ratio	Coefficient	*SE*	*t* ratio	Coefficient	*SE*	*t* ratio
Grand mean	0.080	0.028	—	0.078	0.027	—	0.086	0.028	—
Primary 7 VRQ							0.028	0.002	12.18
Primary 7 reading							0.026	0.002	14.98
Dad occupation							0.008	0.001	5.94
Dad education							.0143	0.041	3.51
Mom education							0.059	0.037	1.59
Dad unemployment							−0.121	0.047	−2.56
Sex							0.056	0.028	1.97
Neighborhood social deprivation							−0.156	0.026	−6.03

(b) Variance components Parameter	Estimate	Estimate	Estimate
Neighborhoods			
$\mathrm{Var}(\pi_{0r}^{(1)}) = \tau_{00}$	0.200	0.141	0.006
Schools			
$\mathrm{Var}(\pi_{c}^{(2)}) = \delta^2$		0.075	0.004
Students			
$\mathrm{Var}(e_{it}) = \sigma^2$	0.804	0.800	0.454

(c) Model fit			
Deviance	—	3074.26	1487.00


Stata/MP 11.2 - [Results]
File Edit Data Graphics Statistics User Window Help
Review
Command
_rc
(R)
11.2
Copyright 2009 StataCorp LP
Statistics/Data Analysis
StataCorp
4905 Lakeway Drive
MP - Parallel Edition
College Station, Texas 77845 USA
800-STATA-PC http://www.stata.com
979-696-4600 stata@stata.com
979-696-4601 (fax)
2-user 2-core Stata network perpetual license:
Serial number: 50110514919
Licensed to: Centre for Multilevel Modelling
University of Bristol
Notes:
1. (/m# option or -set memory-) 500.00 MB allocated to data
2. (/v# option or -set maxvar-) 5000 maximum variables
running C:\Program Files (x86)\Stata11\sysprofile.do ...
running C:\Users\gl9158\profile.do ...
.
Variables
Name
Label
Type
Format
Command
C:\Users\gl9158\Documents
CAP
NUM
OVR


Stata/MP 11.2 - [Results]
File Edit Data Graphics Statistics User Window Help
Review
Command _rc
Variables
Name Label Type Format
(R)
11.2 Copyright 2009 StataCorp LP
Statistics/Data Analysis
StataCorp
4905 Lakeway Drive
MP - Parallel Edition
College Station, Texas 77845 USA
800-STATA-PC http://www.stata.com
979-696-4600 stata@stata.com
979-696-4601 (fax)
2-user 2-core Stata network perpetual license:
Serial number: 50110514919
Licensed to: Centre for Multilevel Modelling
University of Bristol
Notes:
1. (/m# option or -set memory-) 500.00 MB allocated to data
2. (/v# option or -set maxvar-) 5000 maximum variables
running C:\Program Files (x86)\Stata11\sysprofile.do ...
running C:\Users\gl9158\profile.do ...
.
Command
use "http://www.stata-press.com/data/mlmus2/neighborhood.dta", clear
C:\Users\gl9158\Documents
CAP NUM OVR


Stata/MP 11.2 - http://www.stata-press.com/data/mlmus2/neighborhood.dta - [Results]
File Edit Data Graphics Statistics User Window Help
Review
Command _rc
1 use "http://www.stata-press.co...
(R)
11.2 Copyright 2009 StataCorp LP
StataCorp
4905 Lakeway Drive
College Station, Texas 77845 USA
800-STATA-PC http://www.stata.com
979-696-4600 stata@stata.com
979-696-4601 (fax)
Statistics/Data Analysis
MP - Parallel Edition
2-user 2-core Stata network perpetual license:
Serial number: 50110514919
Licensed to: Centre for Multilevel Modelling
University of Bristol
Notes:
1. (/m# option or -set memory-) 500.00 MB allocated to data
2. (/v# option or -set maxvar-) 5000 maximum variables
running C:\Program Files (x86)\Stata11\sysprofile.do ...
running C:\Users\gl9158\profile.do ...
. use "http://www.stata-press.com/data/mlmus2/neighborhood.dta", clear
.
Variables
Name Label Type Format
neighid int %8.0g
schid byte %8.0g
attain float %9.0g
p7vrq float %9.0g
p7read float %9.0g
dadocc float %9.0g
dadunemp byte %8.0g
daded byte %8.0g
momed byte %8.0g
male byte %8.0g
deprive float %9.0g
dummy byte %8.0g
Command
C:\Users\gl9158\Documents
CAP NUM OVR

Stata/MP 11.2 - http://www.stata-press.com/data/mlmus2/neighborhood.dta - [Results]

File Edit Data Graphics Statistics User Window Help

```
  ___  ____  ____  ____  ____ (R)
 /__    /   ____/   /   ____/
___/   /   /___/   /   /___/   11.2   Copyright 2009 StataCorp LP
  Statistics/Data Analysis            StataCorp
                                      4905 Lakeway Drive
     MP - Parallel Edition            College Station, Texas 77845 USA
                                      800-STATA-PC        http://www.stata.com
                                      979-696-4600        stata@stata.com
                                      979-696-4601 (fax)

2-user 2-core Stata network perpetual license:
       Serial number:  50110514919
         Licensed to:  Centre for Multilevel Modelling
                       University of Bristol

Notes:
      1.  (/m# option or -set memory-) 500.00 MB allocated to data
      2.  (/v# option or -set maxvar-) 5000 maximum variables

running C:\Program Files (x86)\Stata11\sysprofile.do ...

running C:\Users\gl9158\profile.do ...

. use "http://www.stata-press.com/data/mlmus2/neighborhood.dta", clear

.
```

C:\Users\gl9158\Documents   CAP NUM OVR

TABLE 1

Organization of data in the Scotland neighborhood study: Numbers of observations in each neighborhood by school cell

	School																	
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Totals
Neighborhood																		
26																	5	5
27																	1	1
29															1		8	9
30																	2	2
31															1		1	2
32															1		5	6
33															1		2	3
35																	3	3
36																	2	2
38															1		4	5
.																		.
.																		.
.																		.
251														4		1		5
252													1	3			1	5
253														3				3
256																	2	2
258												4						4
259											6	1					2	9
260											6							6
261											3						3	6
263											14					1	1	16
.																		.
.																		.
.																		.
793									6									6
794		1				1			11									13
795			2						1									3
796			9						1									10
797			3					1										4
798								10										10
799			1							1								2
800			3	1														4
801			1	1														2
803										3								3
Totals	146	22	146	159	155	101	286	112	136	133	92	190	111	154	91	102	174	2310

Stata/MP 11.2 - http://www.stata-press.com/data/mlmus2/neighborhood.dta - [Results]

File Edit Data Graphics Statistics User Window Help

```
. table neighid schid if inrange(neighid,26,38) | inrange(neighid,251,263) | inrange(neighid
> ,793,800)
```

	schid								
neighid	2	8	10	15	16	17	18	19	20
26									5
27									1
29							1		8
30									2
31							1		1
32							1		5
33							2		2
35									3
36									2
37									1
38							1		4
251						4		1	
252					1	3			1
253						3			
256									2
258				5					
259				6	1		2		
260				7					
261				4			3		
262				5		1	1		
263				14		1	1		
793		1	7						
794	1	1	12						
795	1		1						
796	9								
797	4	1							
798	9	1							
799	1		1						

C:\Users\gl9158\Documents CAP NUM OVR

TABLE 3
Descriptive statistics from the Scotland neighborhood study

Variable	*n*	*m*	*SD*
(a) Neighborhood level			
Social deprivation	524	0.037	0.622
(b) Student level			
Total attainment	2310	0.093	1.00
Primary 7 VRQ	2310	0.506	10.65
Primary 7 reading	2310	−0.044	13.89
Dad occupation	2310	−0.464	11.78
Dad education (1 = Schooling past 15; 0 = No)	2310	0.215	0.41
Mom education (1 = Schooling past 15; 0 = No)	2310	0.248	0.43
Dad unemployment (1 = Yes; 0 = No)	2310	0.109	0.31
Sex (1 = M; 0 = F)	2310	0.480	0.50

Stata/MP 11.2 - http://www.stata-press.com/data/mlmus2/neighborhood.dta - [Results]

File Edit Data Graphics Statistics User Window Help

```
. egen pickone = tag(neighid)

. tabstat deprive if pickone==1, stat(mean sd) format(%4.3f) columns(stats)

    variable |      mean        sd
-------------+--------------------
     deprive |     0.037     0.622
----------------------------------

. tabstat attain p7vrq p7read dadocc daded momed dadunemp male, stat(mean sd) format(%4.3f)
> columns(stats)

    variable |      mean        sd
-------------+--------------------
      attain |     0.093     1.002
       p7vrq |     0.506    10.648
      p7read |    -0.044    13.888
      dadocc |    -0.464    11.782
       daded |     0.215     0.411
       momed |     0.248     0.432
    dadunemp |     0.109     0.312
        male |     0.480     0.500
----------------------------------

.
.
.
.
.
.
.
.
.
.
```

C:\Users\gl9158\Documents CAP NUM OVR

IGLS ESTIMATION

Model 1 fitted using IGLS

$$attain_i = \beta_0 + u^{(2)}_{\text{neighid}(i)} + e_i$$

$$u_j^{(2)} \sim \mathrm{N}\left(0, \sigma^2_{u(2)}\right), \qquad e_i \sim \mathrm{N}(0, \sigma^2_e)$$

```
. runmlwin attain cons, ///
      level2(neighid: cons) ///
      level1(studentid: cons)
```

Stata/MP 11.2 - http://www.stata-press.com/data/mlmus2/neighborhood.dta - [Results]

File Edit Data Graphics Statistics User Window Help

```
. runmlwin attain cons, level2(neighid: cons) level1(studentid: cons) nopause

MLwiN 2.23 multilevel model                     Number of obs      =      2310
Normal response model
Estimation algorithm:  IGLS

-----------------------------------------------------------
                |   No. of       Observations per Group
 Level Variable |   Groups    Minimum    Average    Maximum
----------------+------------------------------------------
        neighid |      524          1        4.4         16
-----------------------------------------------------------

Run time (seconds)   =       1.59
Number of iterations =          3
Log likelihood       = -3207.9849
Deviance             =  6415.9697
------------------------------------------------------------------------------
      attain |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        cons |   .0820249   .0284355     2.88   0.004     .0262923    .1377574
------------------------------------------------------------------------------

------------------------------------------------------------------------------
   Random-effects Parameters |   Estimate   Std. Err.     [95% Conf. Interval]
-----------------------------+------------------------------------------------
Level 2:                     |
                   var(cons) |   .2015358   .0255229      .1515119    .2515597
-----------------------------+------------------------------------------------
Level 1:                     |
                   var(cons) |   .8043717   .0265334      .7523673    .8563761
------------------------------------------------------------------------------

.
```

C:\Users\gl9158\Documents CAP NUM OVR

Model 2 fitted using IGLS

$$attain_i = \beta_0 + u^{(3)}_{\text{schid}(i)} + u^{(2)}_{\text{neighid}(i)} + e_i$$

$$u^{(3)}_j \sim \mathrm{N}\left(0, \sigma^2_{u(3)}\right), \qquad u^{(2)}_j \sim \mathrm{N}\left(0, \sigma^2_{u(2)}\right), \qquad e_i \sim \mathrm{N}(0, \sigma^2_e)$$

```
. tabulate schid, gen(s)
. forvalues i = 1/16 {
.     constraint define `i' [RP3]var(s`i') = [RP3]var(s17)
. }
. runmlwin attain cons, ///
      level3(cons: s1-s17, diagonal) ///
      level2(neighid: cons) ///
      level1(studentid: cons) constraints(1/16)
```

Stata/MP 11.2 - http://www.stata-press.com/data/mlmus2/neighborhood.dta - [Results]

File Edit Data Graphics Statistics User Window Help

```
. quietly tabulate schid, generate(s)

. forvalues s = 1/16 {
  2.
. constraint define `s' [RP3]var(s`s') = [RP3]var(s17)
  3.
. }

. runmlwin attain cons, level3(cons: s1-s17, diagonal) level2(neighid: cons) level1(studenti
> d: cons) constraints(1/16) nopause

 ( 1)   [RP3]var(s1) - [RP3]var(s17) = 0
 ( 2)   [RP3]var(s2) - [RP3]var(s17) = 0
 ( 3)   [RP3]var(s3) - [RP3]var(s17) = 0
 ( 4)   [RP3]var(s4) - [RP3]var(s17) = 0
 ( 5)   [RP3]var(s5) - [RP3]var(s17) = 0
 ( 6)   [RP3]var(s6) - [RP3]var(s17) = 0
 ( 7)   [RP3]var(s7) - [RP3]var(s17) = 0
 ( 8)   [RP3]var(s8) - [RP3]var(s17) = 0
 ( 9)   [RP3]var(s9) - [RP3]var(s17) = 0
 (10)   [RP3]var(s10) - [RP3]var(s17) = 0
 (11)   [RP3]var(s11) - [RP3]var(s17) = 0
 (12)   [RP3]var(s12) - [RP3]var(s17) = 0
 (13)   [RP3]var(s13) - [RP3]var(s17) = 0
 (14)   [RP3]var(s14) - [RP3]var(s17) = 0
 (15)   [RP3]var(s15) - [RP3]var(s17) = 0
 (16)   [RP3]var(s16) - [RP3]var(s17) = 0
Warning:  variance matrix is nonsymmetric or highly singular
MLwiN 2.23 multilevel model                    Number of obs      =       2310
Normal response model
Estimation algorithm: IGLS

                 No. of       Observations per Group
```

C:\Users\gl9158\Documents CAP NUM OVR

Model 3 fitted using IGLS

$$attain_i = \beta_0 + \beta_1 p7vrq_i + \beta_2 p7read_i + \beta_3 dadocc_i + \beta_4 daded_i$$

$$+\beta_5 momed_i + \beta_6 dadunemp_i + \beta_7 male_i + \beta_8 deprive_i$$

$$+u^{(3)}_{\text{schid}(i)} + u^{(2)}_{\text{neighid}(i)} + e_i$$

$$u^{(3)}_j \sim \mathrm{N}\left(0, \sigma^2_{u(3)}\right), \qquad u^{(2)}_j \sim \mathrm{N}\left(0, \sigma^2_{u(2)}\right), \qquad e_i \sim \mathrm{N}(0, \sigma^2_e)$$

```
. runmlwin attain cons p7vrq p7read dadocc daded ///
      momed dadunemp male deprive, ///
      level3(cons: s1-s17, diagonal) ///
      level2(neighid: cons) ///
      level1(studentid: cons) constraints(1/16)
```

Stata/MP 11.2 - http://www.stata-press.com/data/mlmus2/neighborhood.dta - [Results]

File Edit Data Graphics Statistics User Window Help

```
. runmlwin attain cons p7vrq p7read dadocc daded momed dadunemp female deprive, level3(cons:
> s1-s17, diagonal) level2(neighid: cons) level1(studentid: cons) constraints(1/16) nopause

 ( 1)  [RP3]var(s1) - [RP3]var(s17) = 0
 ( 2)  [RP3]var(s2) - [RP3]var(s17) = 0
 ( 3)  [RP3]var(s3) - [RP3]var(s17) = 0
 ( 4)  [RP3]var(s4) - [RP3]var(s17) = 0
 ( 5)  [RP3]var(s5) - [RP3]var(s17) = 0
 ( 6)  [RP3]var(s6) - [RP3]var(s17) = 0
 ( 7)  [RP3]var(s7) - [RP3]var(s17) = 0
 ( 8)  [RP3]var(s8) - [RP3]var(s17) = 0
 ( 9)  [RP3]var(s9) - [RP3]var(s17) = 0
 (10)  [RP3]var(s10) - [RP3]var(s17) = 0
 (11)  [RP3]var(s11) - [RP3]var(s17) = 0
 (12)  [RP3]var(s12) - [RP3]var(s17) = 0
 (13)  [RP3]var(s13) - [RP3]var(s17) = 0
 (14)  [RP3]var(s14) - [RP3]var(s17) = 0
 (15)  [RP3]var(s15) - [RP3]var(s17) = 0
 (16)  [RP3]var(s16) - [RP3]var(s17) = 0
Warning:  variance matrix is nonsymmetric or highly singular
MLwiN 2.23 multilevel model                      Number of obs      =       2310
Normal response model
Estimation algorithm: IGLS

-----------------------------------------------------------
                |   No. of       Observations per Group
 Level Variable |   Groups    Minimum    Average    Maximum
----------------+------------------------------------------
           cons |        1       2310     2310.0       2310
        neighid |      524          1        4.4         16
-----------------------------------------------------------

Run time (seconds)   =       2.54
Number of iterations =          4
Log likelihood       =  2384.6677
```

C:\Users\gl9158\Documents   CAP NUM OVR

TABLE 4
Modeling results for Scotland neighborhood data

(a) Fixed effects	Model 1: School variance ignored			Model 2: School variance estimated			Model 3: Fitted model		
Predictor	Coefficient	*SE*	*t* ratio	Coefficient	*SE*	*t* ratio	Coefficient	*SE*	*t* ratio
Grand mean	0.080	0.028	—	0.078	0.027	—	0.086	0.028	—
Primary 7 VRQ							0.028	0.002	12.18
Primary 7 reading							0.026	0.002	14.98
Dad occupation							0.008	0.001	5.94
Dad education							.0143	0.041	3.51
Mom education							0.059	0.037	1.59
Dad unemployment							−0.121	0.047	−2.56
Sex							0.056	0.028	1.97
Neighborhood social deprivation							−0.156	0.026	−6.03
(b) Variance components									
Parameter	Estimate			Estimate			Estimate		
Neighborhoods $\mathrm{Var}(\pi_{0r}^{(1)}) = \tau_{00}$	0.200			0.141			0.006		
Schools $\mathrm{Var}(\pi_{c}^{(2)}) = \delta^2$				0.075			0.004		
Students $\mathrm{Var}(e_{it}) = \sigma^2$	0.804			0.800			0.454		
(c) Model fit									
Deviance	—			3074.26			1487.00		

Stata/MP 11.2 - http://www.stata-press.com/data/mlmus2/neighborhood.dta - [Results]

File Edit Data Graphics Statistics User Window Help

```
. estimates table model1a model2a model3a, stats(deviance) keep("FP1:" "RP2:" "RP3:var(s1)"
> "RP3:var(s2)" "RP3:var(s17)" "RP1:") b(%9.3f) style(oneline)
```

Variable	model1a	model2a	model3a
FP1			
cons	0.082	0.075	0.030
p7vrq			0.028
p7read			0.026
dadocc			0.008
daded			0.144
momed			0.059
dadunemp			-0.121
female			0.056
deprive			-0.156
RP2			
var(cons)	0.202	0.141	0.004
RP3			
var(s1)		0.075	0.004
var(s2)		0.075	0.004
var(s17)		0.075	0.004
RP1			
var(cons)	0.804	0.799	0.456
Statistics			
deviance	6415.970	6356.711	4769.335

```
.
.
```

C:\Users\gl9158\Documents CAP NUM OVR

MCMC ESTIMATION

Model 1 fitted using MCMC

$$attain_i = \beta_0 + u^{(2)}_{\text{neighid}(i)} + e_i$$

$$u^{(2)}_j \sim \mathrm{N}\left(0, \sigma^2_{u(2)}\right), \qquad e_i \sim \mathrm{N}(0, \sigma^2_e)$$

```
. runmlwin attain cons, ///
      level2(neighid: cons) ///
      level1(studentid: cons) ///
      mcmc(on) initsprevious
```

Stata/MP 11.2 - http://www.stata-press.com/data/mlmus2/neighborhood.dta - [Results]

File Edit Data Graphics Statistics User Window Help

```
. runmlwin attain cons, level2(neighid: cons) level1(studentid: cons) mcmc(on) initsprevious
> nopause

MLwiN 2.23 multilevel model                       Number of obs      =      2310
Normal response model
Estimation algorithm: MCMC

----------------------------------------------------------
                |   No. of       Observations per Group
 Level Variable |   Groups    Minimum    Average    Maximum
----------------+-----------------------------------------
        neighid |      524          1        4.4         16
----------------------------------------------------------

Burnin                     =        500
Chain                      =       5000
Run time (seconds)         =        5.8
Deviance (dbar)            =    6055.62
Deviance (thetabar)        =    5804.62
Effective no. of pars (pd) =     250.99
Bayesian DIC               =    6306.61
------------------------------------------------------------------------------
      attain |      Mean    Std. Dev.       z      ESS     [95% Cred. Interval]
-------------+----------------------------------------------------------------
        cons |  .0831759    .0285483     2.91     1513     .0261265    .1383685
------------------------------------------------------------------------------

------------------------------------------------------------------------------
   Random-effects Parameters |    Mean    Std. Dev.    ESS     [95% Cred. Int]
-----------------------------+------------------------------------------------
Level 2:                     |
                   var(cons) |  .2025886   .0257499     738   .1563706  .2566977
-----------------------------+------------------------------------------------
Level 1:                     |
                   var(cons) |  .8060307   .0265735    2678   .7557447  .8587067
```

C:\Users\gl9158\Documents CAP NUM OVR

Model 2 fitted using MCMC

$$attain_i = \beta_0 + u^{(3)}_{\text{schid}(i)} + u^{(2)}_{\text{neighid}(i)} + e_i$$

$$u^{(3)}_j \sim \mathrm{N}\left(0, \sigma^2_{u(3)}\right), \qquad u^{(2)}_j \sim \mathrm{N}\left(0, \sigma^2_{u(2)}\right), \qquad e_i \sim \mathrm{N}(0, \sigma^2_e)$$

```
. matrix b = (0,.075,.15,.8)

. runmlwin attain cons, ///

      level3(schid: cons) ///

      level2(neighid: cons) ///

      level1(studentid: cons) ///

      mcmc(cc) initsb(b)
```

Stata/MP 11.2 - http://www.stata-press.com/data/mlmus2/neighborhood.dta - [Results]

File Edit Data Graphics Statistics User Window Help

```
. matrix b = (0,.075,.15,.8)

. runmlwin attain cons, level3(schid: cons) level2(neighid: cons) level1(studentid: cons) mc
> mc(cc) initsb(b) nopause

MLwiN 2.23 multilevel model                     Number of obs      =      2310
Normal response model
Estimation algorithm: MCMC

-----------------------------------------------------------
                |   No. of       Observations per Group
 Level Variable |   Groups    Minimum    Average    Maximum
----------------+------------------------------------------
          schid |       17         22      135.9        286
        neighid |      524          1        4.4         16
-----------------------------------------------------------

Burnin                     =        500
Chain                      =       5000
Run time (seconds)         =       6.55
Deviance (dbar)            =    6039.42
Deviance (thetabar)        =    5818.77
Effective no. of pars (pd) =     220.65
Bayesian DIC               =    6260.07
------------------------------------------------------------------------------
      attain |      Mean    Std. Dev.       z      ESS     [95% Cred. Interval]
-------------+----------------------------------------------------------------
        cons |  .0962458    .0651659 |    1.48      228    -.0349854    .2170355
------------------------------------------------------------------------------

------------------------------------------------------------------------------
  Random-effects Parameters |     Mean    Std. Dev.    ESS     [95% Cred. Int]
----------------------------+-------------------------------------------------
Level 3:                    |
                  var(cons) |  .0005634   .0402008    2175    .0387661   .2261023
```

C:\Users\gl9158\Documents   CAP NUM OVR

Model 3 fitted using MCMC

$$attain_i = \beta_0 + \beta_1 p7vrq_i + \beta_2 p7read_i + \beta_3 dadocc_i + \beta_4 daded_i$$
$$+\beta_5 momed_i + \beta_6 dadunemp_i + \beta_7 male_i + \beta_8 deprive_i$$
$$+u^{(3)}_{\mathrm{schid}(i)} + u^{(2)}_{\mathrm{neighid}(i)} + e_i$$
$$u^{(3)}_j \sim \mathrm{N}\left(0, \sigma^2_{u(3)}\right), \qquad u^{(2)}_j \sim \mathrm{N}\left(0, \sigma^2_{u(2)}\right), \qquad e_i \sim \mathrm{N}(0, \sigma^2_e)$$

```
. matrix b = (0,0,0,0,0,0,0,0,0,1,1,1)
. runmlwin attain cons p7vrq p7read dadocc daded ///
        momed dadunemp male deprive, ///
        level3(schid: cons) ///
        level2(neighid: cons) ///
        level1(studentid: cons) mcmc(cc) initsb(b)
```

Stata/MP 11.2 - http://www.stata-press.com/data/mlmus2/neighborhood.dta - [Results]

File Edit Data Graphics Statistics User Window Help

```
. runmlwin attain cons p7vrq p7read dadocc daded momed dadunemp male deprive, level3(schid:
> cons) level2(neighid: cons) level1(studentid: cons) mcmc(cc) initsb(b) nopause

MLwiN 2.23 multilevel model                     Number of obs      =      2310
Normal response model
Estimation algorithm: MCMC

-----------------------------------------------------------
                |   No. of       Observations per Group
 Level Variable |   Groups    Minimum    Average    Maximum
----------------+------------------------------------------
          schid |       17         22      135.9        286
        neighid |      524          1        4.4         16
-----------------------------------------------------------

Burnin                     =        500
Chain                      =       5000
Run time (seconds)         =       11.4
Deviance (dbar)            =    4744.81
Deviance (thetabar)        =    4704.18
Effective no. of pars (pd) =      40.63
Bayesian DIC               =    4785.44
------------------------------------------------------------------------------
      attain |      Mean    Std. Dev.       z      ESS     [95% Cred. Interval]
-------------+----------------------------------------------------------------
        cons |   .0908623    .028849     3.15     1642     .0341885    .1490847
       p7vrq |   .0275759   .0022769    12.11     4555     .0232042    .0320144
      p7read |   .0262263   .0017901    14.65     4683     .0226469    .0297636
      dadocc |   .0080741   .0013762     5.87     4678     .0053839    .0107383
       daded |   .1427545   .0411568     3.47     5455     .0616242    .2235425
       momed |    .060508   .0379714     1.59     4707    -.0130428    .1339119
    dadunemp |  -.1224706   .0468447    -2.61     4503    -.2136019   -.0290561
        male |  -.0556571   .0281836    -1.97     4665    -.1103975   -.0002866
     deprive |  -.1562529   .0260999    -5.99     3698    -.2078412   -.1057263
```

C:\Users\gl9158\Documents   CAP NUM OVR

TABLE 4

Modeling results for Scotland neighborhood data

(a) Fixed effects	Model 1: School variance ignored			Model 2: School variance estimated			Model 3: Fitted model		
Predictor	Coefficient	*SE*	*t* ratio	Coefficient	*SE*	*t* ratio	Coefficient	*SE*	*t* ratio
Grand mean	0.080	0.028	—	0.078	0.027	—	0.086	0.028	—
Primary 7 VRQ							0.028	0.002	12.18
Primary 7 reading							0.026	0.002	14.98
Dad occupation							0.008	0.001	5.94
Dad education							.0143	0.041	3.51
Mom education							0.059	0.037	1.59
Dad unemployment							−0.121	0.047	−2.56
Sex							0.056	0.028	1.97
Neighborhood social deprivation							−0.156	0.026	−6.03

(b) Variance components Parameter	Model 1 Estimate	Model 2 Estimate	Model 3 Estimate
Neighborhoods $\mathrm{Var}(\pi^{(1)}_{0r}) = \tau_{00}$	0.200	0.141	0.006
Schools $\mathrm{Var}(\pi^{(2)}_{c}) = \delta^2$		0.075	0.004
Students $\mathrm{Var}(e_{it}) = \sigma^2$	0.804	0.800	0.454

(c) Model fit	Model 1	Model 2	Model 3
Deviance	—	3074.26	1487.00

Stata/MP 11.2 - http://www.stata-press.com/data/mlmus2/neighborhood.dta - [Results]

File Edit Data Graphics Statistics User Window Help

```
. estimates table model1b model2b model3b, stats(deviance) keep("FP1:" "RP2:" "RP3:" "RP1:")
>  b(%9.3f) style(oneline)
```

Variable	model1b	model2b	model3b
FP1			
cons	0.083	0.096	0.035
p7vrq			0.028
p7read			0.026
dadocc			0.008
daded			0.143
momed			0.061
dadunemp			-0.122
female			0.056
deprive			-0.156
RP2			
var(cons)	0.203	0.142	0.005
RP3			
var(cons)		0.100	0.006
RP1			
var(cons)	0.806	0.800	0.457
Statistics			
deviance			

```
.
.
.
.
```

C:\Users\gl9158\Documents

CAP NUM OVR

WORK EFFICIENTLY

Do-file Editor - Bristol_e-Stat.do

File Edit Tools View

Bristol_e-Stat.do

```

* Open the Raudenbush (1993) Scotish neighbourhood study data
use "http://www.stata-press.com/data/mlmus2/neighborhood.dta", clear

* Replicate Table 1 of Raudenbush (1993). The table shows that the data are
* cross-classified. Students are said to be nested within the cross-
* classification of schools by neighbourhoods. To account for both school
* and neighbourhood effects in our models of student attainment, we will
* need to fit cross-classified multilevel model using runmlwin.
table neighid schid if inrange(neighid,26,38) | inrange(neighid,251,263) ///
    | inrange(neighid,793,800)

* Replicate Table 3 of Raudenbush (1993). The table gives summary statistics.
egen pickone = tag(neighid)

tabstat deprive if pickone==1, stat(mean sd) format(%4.3f) columns(stats)

tabstat attain p7vrq p7read dadocc daded momed dadunemp male, ///
    stat(mean sd) format(%4.3f) columns(stats)


* Generate a unique student identifer variable which will be the level 1
* unit identifier variable in the runmlwin command
gen studentid = _n

* Generate a variable cons to act as the constant or intercept variable in
* the runmlwin models
gen cons = 1

* Sort the data by students within neighbourhoods (otherwise runmlwin will
* complain when we try to fit the following model)
sort neighid studentid

* Fit a two-level (students within neighbourhoods) variances components
* model to attain. This model is refered to as model 1 in Table 4 of
* Raudenbush (1993). Note, you will need to click the "Resume Macro" button
```

Ready   Line: 12, Col: 29   CAP   NUM   OVR

RESOURCES TO HELP YOU LEARN RUNMLWIN

Viewer (#1) [help runmlwin]

help runmlwin

Advice Contents What's New News

help runmlwin

Title

runmlwin - Run the MLwiN multilevel modelling software from within Stata

Syntax

```
runmlwin responses_and_fixed_part, random_part [discrete(discrete_options)] [mcmc(mcmc_options)]
        [general_options]
```

where the syntax of *responses_and_fixed_part* is one of the following

for univariate continuous, binary, proportion and count response models

```
depvar indepvars [if] [in]
```

for univariate ordered and unordered categorical response models

```
depvar indepvars1 [(indepvars2, contrast(numlist)) ... ] [if] [in]
```

where *indepvars1* are those independent variables which appear with separate coefficients in each of every log-odds contrast, while *indepvars2* are those independent variables which appear with common coefficients for those log-odds contrasts specified in **contrast(***numlist***)**. Contrasts can be thought of as the separate "subequations" or "arms" of a multinomial response model. These contrasts are indexed 1,2,... up to the total number of contrasts included in the model. The total number of contrasts will be one less than the number of response categories.

for multivariate response models

```
(depvar1 indepvars1, equation(numlist))
      (depvar2 indepvars2, equation(numlist))
      [(depvar3 indepvars3, equation(numlist))]
      [ ... ]
      [if] [in]
```

where **equation(***numlist***)** specifies equation numbers. Equation numbers are indexed 1,2,... up to the total number of equations (i.e. response variables) included in the model.

and the syntax of *random_part* is

```
[ ... ] [level2(levelvar: [varlist] [, random_part_options])]
         level1(levelvar: [varlist] [, random_part_options])
```

Viewer (#1) [help runmlwin]

help runmlwin

Advice | Contents | What's New | News

Examples

IMPORTANT. The following examples will only work on your computer once you have installed MLwiN and once you have told **runmlwin** what the mlwin.exe file address is. See *Remarks on installation instructions* above for more information.

(a) Continuous response models

Two-level models

Setup

```
. use http://www.bristol.ac.uk/cmm/media/runmlwin/tutorial, clear
```

Two-level random-intercept model, analogous to xtreg (fitted using IGLS)
(See page 28 of the MLwiN User Manual)
(You will need to click the "Resume macro" button twice in MLwiN to fit the model.)

```
. runmlwin normexam cons standlrt, level2(school: cons) level1(student: cons)
```

Two-level random-intercept and random-slope (coefficient) model (fitted using IGLS)
(See page 59 of the MLwiN User Manual)

```
. runmlwin normexam cons standlrt, level2 (school: cons standlrt) level1 (student: cons)
```

Refit the model suppressing the two pauses in MLwiN (fitted using IGLS)
(See page 59 of the MLwiN User Manual)

```
. runmlwin normexam cons standlrt, level2 (school: cons standlrt) level1 (student: cons) nopause
```

Refit the model, where this time we additionally calculate the level 2 residuals (fitted using IGLS)
(See page 59 of the MLwiN User Manual)

```
. runmlwin normexam cons standlrt, level2 (school: cons standlrt, residuals(u)) level1 (student: cons)
```

Two-level random-intercept and random-slope (coefficient) model with a complex level 1 variance function (fitted using IGLS)
(See page 99 of the MLwiN User Manual)

```
. matrix A = (1,1,0,0,0,1)
. runmlwin normexam cons standlrt girl, level2(school: cons standlrt) level1(student: cons standlrt
    girl, elements(A))
```

Two-level random-intercept and random-slope (coefficient) model using MCMC (where we first fit the model using IGLS to obtain initial values for the MCMC chains)
(See page 71 of the MLwiN MCMC Manual)

```
. runmlwin normexam cons standlrt, level2 (school: cons standlrt) level1 (student: cons)
. runmlwin normexam cons standlrt, level2 (school: cons standlrt) level1 (student: cons) mcmc(on)
    initsprevious
```

Multivariate response models

SOFTWARE

- MLwiN
- Realcom
- MLPowSim
- runmlwin
  - Presentations
  - Examples
  - Citations
  - User Forum
- CMM software support

University home > Centre for Multilevel Modelling... > Software > runmlwin

runmlwin: Running MLwiN from within Stata

runmlwin is a Stata command which allows Stata users to run the powerful MLwiN multilevel modelling software from within Stata.

The multilevel models fitted by **runmlwin** are often considerably faster than those fitted by the Stata's **xtmixed, xtmelogit** and **xtmepoisson** commands. The range of models which can be fitted by **runmlwin** is also much wider than those commands. **runmlwin** also allows fast estimation on large data sets for many of the more complex multilevel models available through the user written **gllamm** command.

MLwiN has the following features:

1. Estimation of multilevel models for continuous, binary, count, ordered categorical and unordered categorical data
2. Fast estimation via classical and Bayesian methods
3. Estimation of multilevel models for cross-classified and multiple membership nonhierarchical data structures
4. Estimation of multilevel multivariate response models, multilevel spatial models, multilevel measurement error models and multilevel multiple imputation models

These details with a screen shot are available on our **runmlwin** leaflet (pdf, 0.1mb)

Presentations

We have provided a range of presentations showcasing **runmlwin**. These presentations provide a quick overview of how the command works and the range of models which can be fitted. More > >

Download

SOFTWARE
MLwiN
Realcom
MLPowSim
runmlwin
Presentations
Examples
Citations
User Forum
CMM software support

University home > Centre for Multilevel Modelling... > Software > runmlwin > Presentations

Presentations using runmlwin

- UK Stata Users' Group, 17th Meeting (16th September 2011)
  - Slides (PDF, 2.0mb)
  - Stata do-file (do, 0.1mb) to replicate all analyses presented in the slides.
- University of Bristol, Mplus/MlwiN User Group (MUGS) meeting (14th June 2011)
  - Slides (PDF, 2.3mb)
  - Stata do-file (do, 0.1mb) to replicate all analyses presented in the slides.
- Modern Modeling Methods (M3) Conference, University of Connecticut (26th May 2011)
  - Slides (PDF, 3.2mb)
  - Stata do-file (do, 0.1mb) to replicate all analyses presented in the slides.
- 2011 American Sociological Association Spring Methodology Conference, Tilburg University (20th May 2011)
  - Slides (PDF, 2.0mb)
  - Stata do-file (do, 0.1mb) to replicate all analyses presented in the slides.
- University of Bristol, e-Stat meeting (7th April 2011)
  - Slides (PDF, 1.7mb)
  - Stata do-file (do, 0.1mb) to replicate all analyses presented in the slides.
- 8th International Amsterdam Multilevel Conference (17th March 2011)

SOFTWARE
MLwiN
Realcom
MLPowSim
runmlwin
Presentations
Examples
Citations
User Forum
CMM software support

University home > Centre for Multilevel Modelling... > Software > runmlwin > Examples

Examples using runmlwin

MLwiN User Manual

These do-files and log files replicate the analyses reported in the MLwiN User Manual (PDF, 4.6 mb) Rasbash, J., Steele, F., Browne, W.J. and Goldstein, H. (2009) Centre for Multilevel Modelling, University of Bristol.

Note that we have not created do-files for Chapters 1, 8 or 19 of the manual as no models are fitted in those chapters. We have also not yet attempted to replicate the analysis in Chapter 17.

- 1 - Introducing Multilevel Models
- 2 - Introduction to Multilevel Modelling (do | log)
- 3 - Residuals (do | log)
- 4 - Random Intercept and Random Slope Models (do | log)
- 5 - Graphical Procedures for Exploring the Model (do | log)
- 6 - Contextual Effects (do | log)
- 7 - Modelling the Variance as a Function of Explanatory Variables (do | log)
- 8 - Getting Started with your Data
- 9 - Logistic Models for Binary and Binomial Responses (do | log)
- 10 - Multinomial Logistic Models for Unordered Categorical Responses (do | log)
- 11 - Fitting an Ordered Category Response Model (do | log)
- 12 - Modelling Count Data (do | log)
- 13 - Fitting Models to Repeated Measures Data (do | log)
- 14 - Multivariate Response Models (do | log)

- Rasbash, J., Charlton, C., Browne, W.J., Healy, M. and Cameron, B. 2009. MLwiN Version 2.1. Centre for Multilevel Modelling, University of Bristol.

For models fitted using MCMC estimation, we ask that you additionally cite:

- Browne, W.J. 2009. MCMC Estimation in MLwiN, v2.13. Centre for Multilevel Modelling, University of Bristol.

Papers using runmlwin

Please let George Leckie (g.leckie@bristol.ac.uk) know of any further publications using **runmlwin** including forthcoming papers, books, PhD theses, etc.

- Cheung, C., Goodman, D., Leckie, G. and Jenkins, J. (2011) Understanding Contextual Effects on Externalizing Behaviors in Children in Out-of-home Care: Influence of Workers and Foster Families. *Children and Youth Services Review*, 33, 2050-2060.

- Chung, H. and Beretvas, S.N. (2011) The Impact of ignoring multiple membership data structures in multilevel models. *British Journal of Mathematical and Statistical Psychology. Forthcoming.*

- Leckie, G. and Baird, J.-A. (2011) Rater effects on essay scoring: A multilevel analysis of severity drift, central tendency and rater experience. *Journal of Educational Measurement. Forthcoming.*

- Leckie, G., Pillinger, R., Jones, K. and Goldstein, H. (2011) Multilevel modelling of social segregation. *Journal of Educational and Behavioral Statistics. Forthcoming.*

- Paternoster, L., Howe, L. D., Tilling, K., Weedon, M. N., Freathy, R. M., Frayling, T. M., Kemp, J. P., Davey Smith, G., Timpson, N. J. Ring, S. M., Evans, D. M. and Lawlor, D. A. (2011) Adult height variants affect birth length and growth rate in children. *Human Molecular Genetics. Forthcoming*

Books discussing runmlwin

- Snijders, T. and Bosker, R. (2011) Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling, Second Edition. *Sage. Forthcoming.*

www.cmm.bristol.ac.uk/forum • View forum - runmlwin user forum - Mozilla Firefox

File Edit View History Bookmarks Tile Tools Help

http://www.cmm.bristol.ac.uk/forum/viewforum.php?f=3

runmlwin user forum

Forum rules

NEWTOPIC Search this forum... Search 20 topics • Page 1 of 1

ANNOUNCEMENTS	REPLIES	VIEWS	LAST POST
Do-files to replicate entire MLwiN User & MCMC Manuals by GeorgeLeckie » Mon Apr 18, 2011 5:30 pm	0	123	by GeorgeLeckie Mon Apr 18, 2011 5:30 pm
Welcome to the runmlwin discussion forum by GeorgeLeckie » Fri Apr 01, 2011 4:06 pm	0	130	by GeorgeLeckie Fri Apr 01, 2011 4:06 pm

TOPICS	REPLIES	VIEWS	LAST POST
MVs & error message 'line too long' by julia1633 » Mon Aug 15, 2011 3:17 pm 1 2	17	125	by julia1633 Sun Aug 28, 2011 12:05 am
runmlwin in Batch mode - gui causing error? by ash » Sat Aug 27, 2011 6:43 am	2	20	by ash Sat Aug 27, 2011 9:07 pm
Bug in residuals(u, savechains("u.dta", replace")) ? by ash » Mon Aug 01, 2011 7:06 pm	4	76	by GeorgeLeckie Wed Aug 03, 2011 6:25 pm
Predictions via the runmlwin interface: a clarification by ewancarr » Tue Jul 26, 2011 6:49 pm	6	84	by GeorgeLeckie Wed Jul 27, 2011 7:04 pm
highly correlated multivariate depenents -> numerical error by ash » Sat Jul 23, 2011 10:48 am	1	67	by GeorgeLeckie Mon Jul 25, 2011 3:49 pm
Input dataset contains double precision data... by ewancarr » Wed Jul 13, 2011 3:55 pm	2	81	by ewancarr Wed Jul 13, 2011 9:16 pm
Modelling Count Data (example do-file) - mismatch error by leap » Tue Jul 12, 2011 10:18 am	1	48	by ChrisCharlton Wed Jul 13, 2011 3:32 pm
Error code: r(-1073740777); by pd65 » Mon Jul 04, 2011 11:01 am	7	207	by GeorgeLeckie Thu Jul 07, 2011 3:15 pm
MCMC estimation by [illegible]	5	293	by ChrisCharlton [illegible]

Citing `runmlwin`

- If you use `runmlwin` in your work, please cite `runmlwin`
- Leckie, G. and Charlton, C. (2011) *runmlwin: Stata module for fitting multilevel models in the MLwiN software package*. Centre for Multilevel Modelling, University of Bristol.
- We can then add you to the list of papers using `runmlwin` on our website
- http://www.bristol.ac.uk/cmm/software/runmlwin/citations