Running MLwiN from within Stata: the `runmlwin` command

17th Stata Users' Group Meeting
Cass Business School, London
16th September 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

TWO-LEVEL MULTILEVEL MODELS

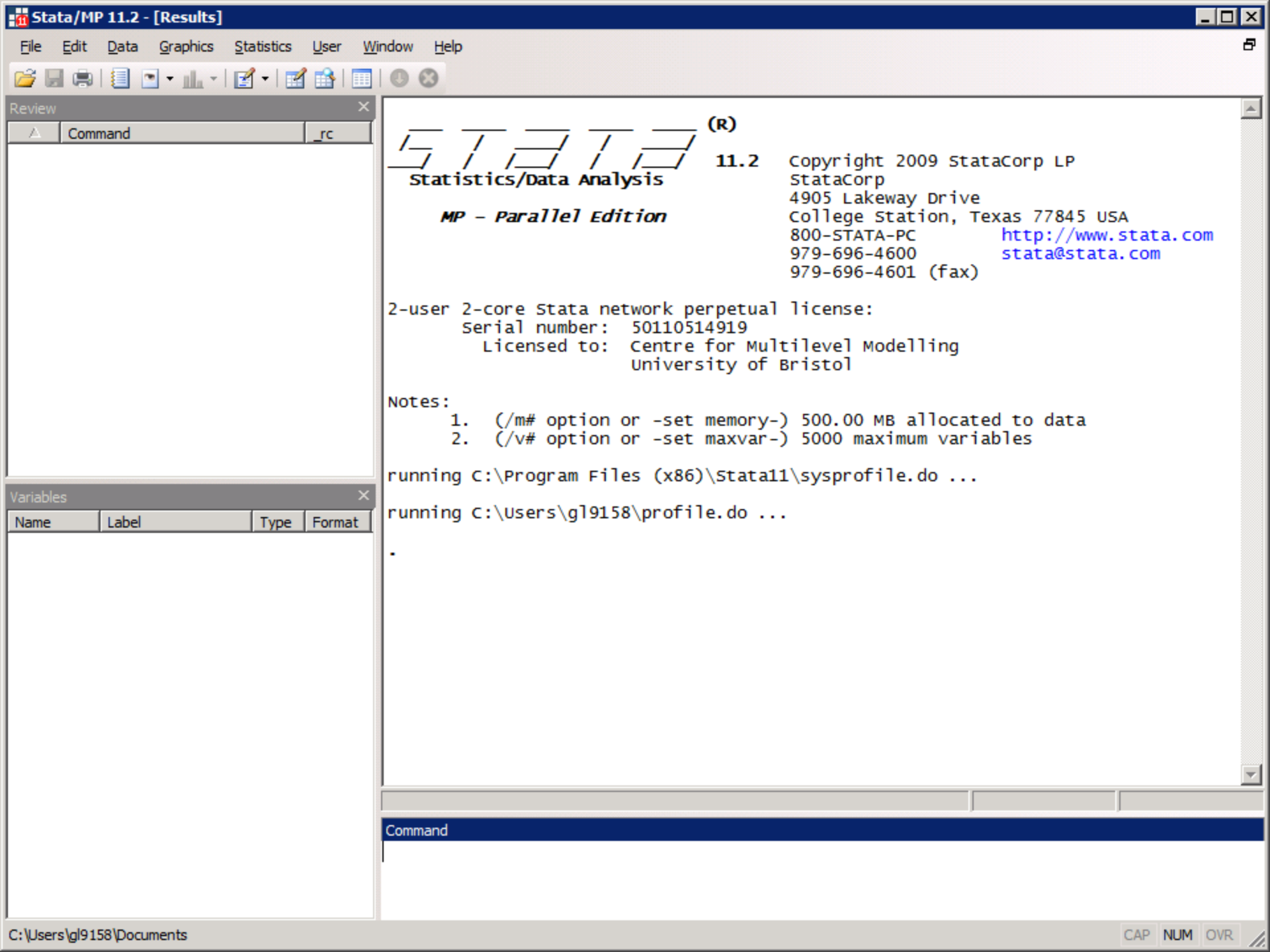
Two-level variance components model

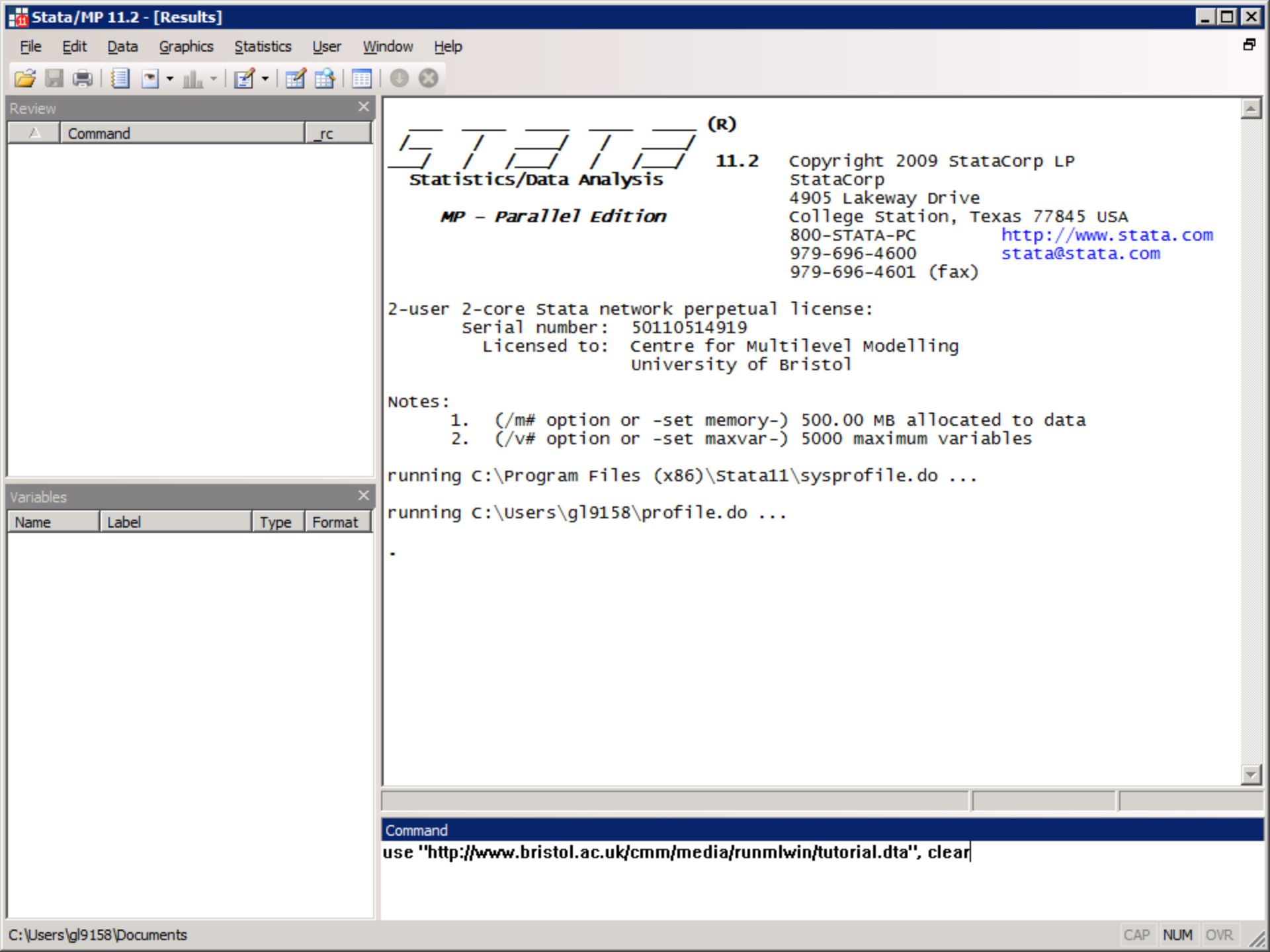
- Inner-London schools exam scores data set
- Classic MLwiN User Manual example
- First analysed by Goldstein et al. (1993)
- Reanalysed by Goldstein (2010), Rabe-Hesketh and Skrondal (2008), Rasbash et al. (2009) and others
- 4059 students nested within 65 schools

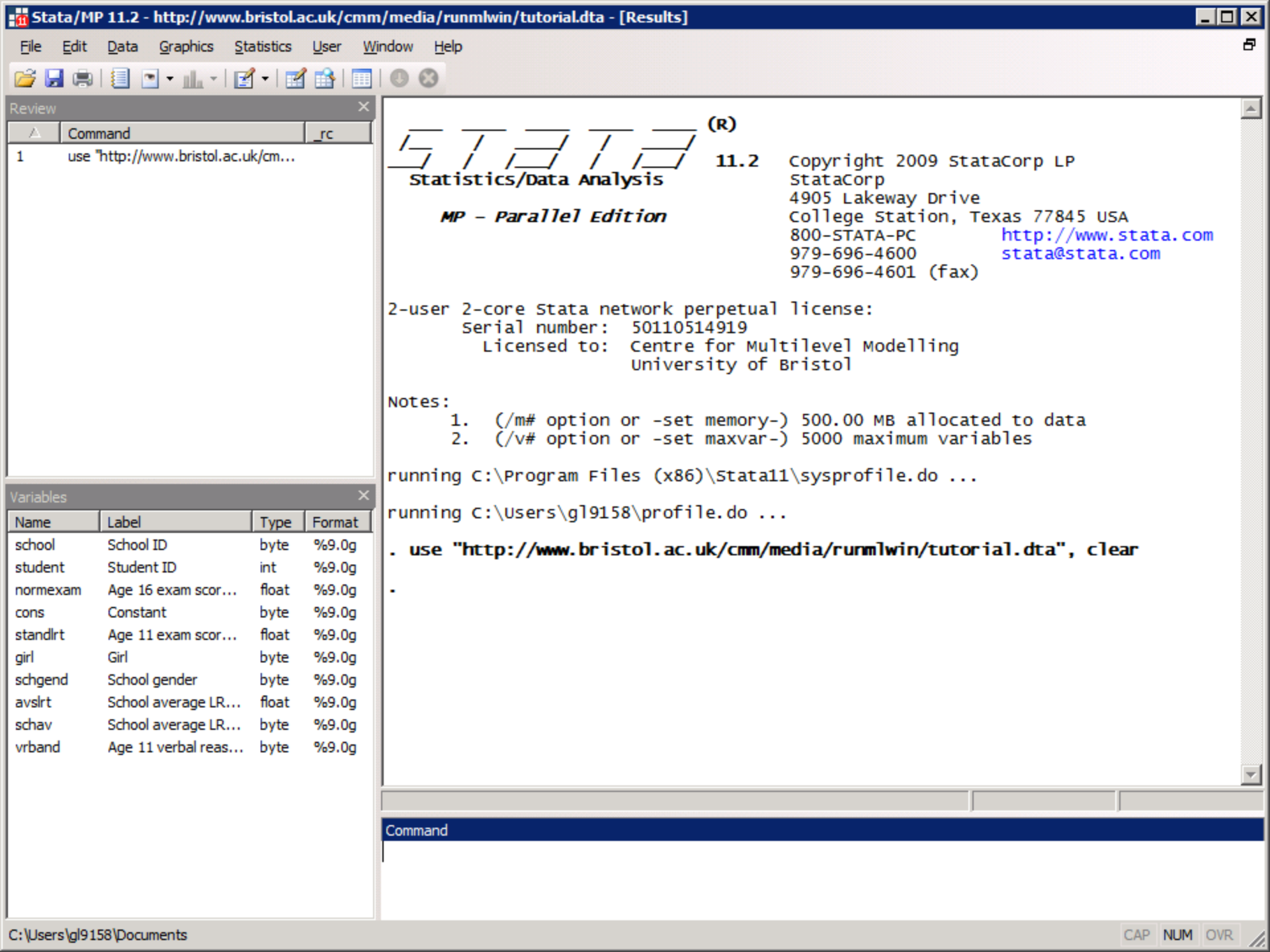
$$normexam_{ij} = \beta_0 + u_j + e_{ij}$$

$$u_j \sim \mathrm{N}(0, \sigma_u^2)$$

$$e_{ij} \sim \mathrm{N}(0, \sigma_e^2)$$


Stata/MP 11.2 - [Results]
File Edit Data Graphics Statistics User Window Help
Review
Command _rc
Variables
Name Label Type Format
(R)
11.2
Statistics/Data Analysis
MP - Parallel Edition
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)
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
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)
Statistics/Data Analysis
11.2
MP - Parallel Edition
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)
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.bristol.ac.uk/cmm/media/runmlwin/tutorial.dta", clear
C:\Users\gl9158\Documents
CAP NUM OVR


Stata/MP 11.2 - http://www.bristol.ac.uk/cmm/media/runmlwin/tutorial.dta - [Results]
File Edit Data Graphics Statistics User Window Help
Review
Command _rc
1 use "http://www.bristol.ac.uk/cm...
Variables
Name Label Type Format
school School ID byte %9.0g
student Student ID int %9.0g
normexam Age 16 exam scor... float %9.0g
cons Constant byte %9.0g
standlrt Age 11 exam scor... float %9.0g
girl Girl byte %9.0g
schgend School gender byte %9.0g
avslrt School average LR... float %9.0g
schav School average LR... byte %9.0g
vrband Age 11 verbal reas... byte %9.0g
(R)
11.2
Statistics/Data Analysis
MP - Parallel Edition
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)
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.bristol.ac.uk/cmm/media/runmlwin/tutorial.dta", clear
.
Command
C:\Users\gl9158\Documents
CAP NUM OVR

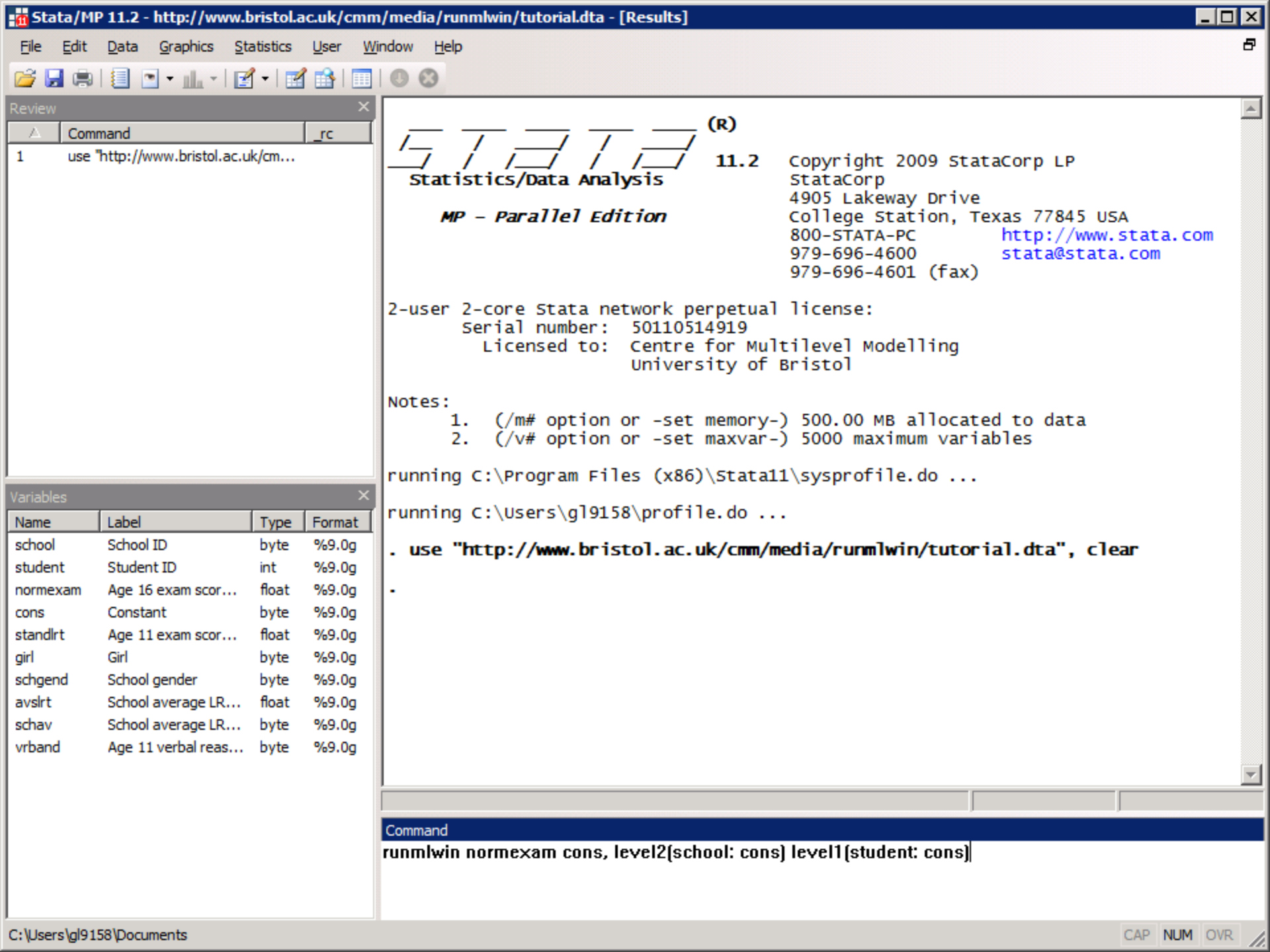
The `runmlwin` command syntax

$$normexam_{ij} = \beta_0 + u_j + e_{ij}$$

$$u_j \sim \mathrm{N}(0, \sigma_u^2)$$

$$e_{ij} \sim \mathrm{N}(0, \sigma_e^2)$$

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


Stata/MP 11.2 - http://www.bristol.ac.uk/cmm/media/runmlwin/tutorial.dta - [Results]
File Edit Data Graphics Statistics User Window Help
Review
Command _rc
1 use "http://www.bristol.ac.uk/cm...
Variables
Name Label Type Format
school School ID byte %9.0g
student Student ID int %9.0g
normexam Age 16 exam scor... float %9.0g
cons Constant byte %9.0g
standlrt Age 11 exam scor... float %9.0g
girl Girl byte %9.0g
schgend School gender byte %9.0g
avslrt School average LR... float %9.0g
schav School average LR... byte %9.0g
vrband Age 11 verbal reas... byte %9.0g
(R)
Statistics/Data Analysis
11.2
MP - Parallel Edition
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)
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.bristol.ac.uk/cmm/media/runmlwin/tutorial.dta", clear
.
Command
runmlwin normexam cons, level2(school: cons) level1(student: cons)
C:\Users\gl9158\Documents
CAP NUM OVR

Stata/MP 11.2 - http://www.bristol.ac.uk/cmm/media/runmlwin/tutorial.dta - [Results]

File Edit Data Graphics Statistics User Window Help

Review

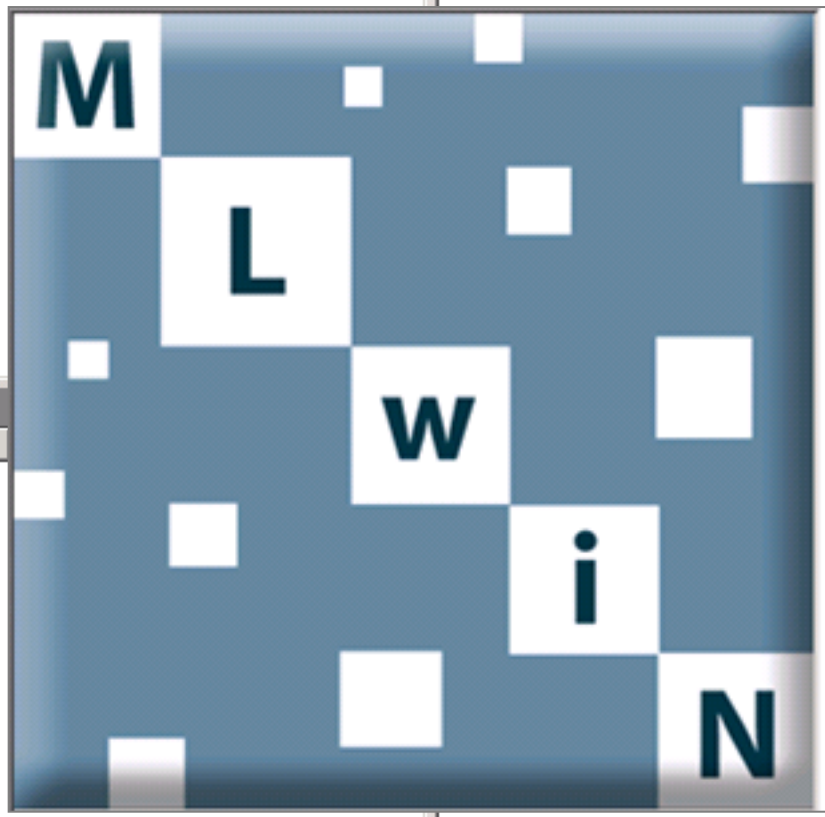
	Command	_rc
1	use "http://www.bristol.ac.uk/cm...	
2	runmlwin normexam cons, level2(...	

Variables

Name
school
student
normexam
cons
standlrt
girl
schgend
avslrt
schav
vrband

(R)
Statistics/Data Analysis
11.2
MP - Parallel Edition

Copyright 2009 StataCorp LP
StataCorp
4905 Lakeway Drive
College Station, Texas 77845 USA
.com



MLwiN
Version 2.23

© Centre for Multilevel Modelling
University of Bristol

Software authors :
Jon Rasbash
and
William Browne
Michael Healy
Bruce Cameron
Christopher Charlton

March 2011

We are grateful to the ESRC for their sustained support.

Command

C:\Users\gl9158\Documents

CAP NUM OVR

MLwiN - [Equations]
File Edit Options Model Estimation Data Manipulation Basic Statistics Graphs Window Help
Start More Stop IGLS
Estimation control..
Resume macro
Abort Macro
$\text{normexam}_{ij} \sim \text{N}(XB, \Omega)$
$\text{normexam}_{ij} = \beta_{0ij}\text{cons}$
$\beta_{0ij} = \beta_0 + u_{0j} + e_{0ij}$
$\begin{bmatrix} u_{0j} \end{bmatrix} \sim \text{N}(0, \Omega_u) : \Omega_u = \begin{bmatrix} \sigma^2_{u0} \end{bmatrix}$
$\begin{bmatrix} e_{0ij} \end{bmatrix} \sim \text{N}(0, \Omega_e) : \Omega_e = \begin{bmatrix} \sigma^2_{e0} \end{bmatrix}$
Name + - Add Term Estimates Nonlinear Clear Notation Responses Store Help Zoom 150
random fixed iteration 0
Equations

MLwiN - [Equations]
File Edit Options Model Estimation Data Manipulation Basic Statistics Graphs Window Help
Start More Stop IGLS
Estimation control..
Resume macro
Abort Macro
$\text{normexam}_{ij} \sim \text{N}(XB, \Omega)$
$\text{normexam}_{ij} = \beta_{0ij}\text{cons}$
$\beta_{0ij} = -0.013(0.054) + u_{0j} + e_{0ij}$
$\left[u_{0j}\right] \sim \text{N}(0, \Omega_u) : \Omega_u = \left[0.169(0.032)\right]$
$\left[e_{0ij}\right] \sim \text{N}(0, \Omega_e) : \Omega_e = \left[0.848(0.019)\right]$
-2*loglikelihood(IGLS Deviance) = 11010.648(4059 of 4059 cases in use)
Name + - Add Term Estimates Nonlinear Clear Notation Responses Store Help Zoom 150
random fixed iteration 3
Equations

Stata/MP 11.2 - http://www.bristol.ac.uk/cmm/media/runmlwin/tutorial.dta - [Results]

File Edit Data Graphics Statistics User Window Help

Review

	Command	_rc
1	use "http://www.bristol.ac.uk/cm...	
2	runmlwin normexam cons, level2(...	

Variables

Name	Label	Type	Format
school	School ID	byte	%9.0g
student	Student ID	int	%9.0g
normexam	Age 16 exam scor...	float	%9.0g
cons	Constant	byte	%9.0g
standlrt	Age 11 exam scor...	float	%9.0g
girl	Girl	byte	%9.0g
schgend	School gender	byte	%9.0g
avslrt	School average LR...	float	%9.0g
schav	School average LR...	byte	%9.0g
vrband	Age 11 verbal reas...	byte	%9.0g

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

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

MLwiN 2.23 multilevel model                     Number of obs      =      4059
Normal response model
Estimation algorithm: IGLS
```

Level Variable	No. of Groups	Observations per Group Minimum	Average	Maximum
school	**65**	**2**	**62.4**	**198**

```
Run time (seconds)   =      12.93
Number of iterations =          3
Log likelihood       = -5505.3242
Deviance             =  11010.648
```

normexam	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
cons	-.0131668	.0536254	-0.25	0.806	-.1182706	.091937

Random-effects Parameters		Estimate	Std. Err.	[95% Conf.	Interval]
Level 2:	var(cons)	**.1686251**	**.0324466**	**.1050309**	**.2322194**
Level 1:	var(cons)	**.8477613**	**.0189712**	**.8105786**	**.8849441**

```
.
```

Command

C:\Users\gl9158\Documents

CAP NUM OVR

Stata/MP 11.2 - http://www.bristol.ac.uk/cmm/media/runmlwin/tutorial.dta - [Results]

File Edit Data Graphics Statistics User Window Help

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

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

MLwiN 2.23 multilevel model                     Number of obs      =      4059
Normal response model
Estimation algorithm: IGLS
```

Level Variable	No. of Groups	Observations per Group Minimum	Average	Maximum
school	**65**	**2**	**62.4**	**198**

```
Run time (seconds)   =      12.93
Number of iterations =          3
Log likelihood       = -5505.3242
Deviance             =  11010.648
```

normexam	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
cons	-.0131668	.0536254	-0.25	0.806	-.1182706	.091937

Random-effects Parameters		Estimate	Std. Err.	[95% Conf.	Interval]
Level 2:	var(cons)	.1686251	.0324466	.1050309	.2322194
Level 1:	var(cons)	.8477613	.0189712	.8105786	.8849441

C:\Users\gl9158\Documents

CAP NUM OVR

Add covariates

$$normexam_{ij} = \beta_0 + \beta_1 standlrt_{ij} + \beta_2 girl_{ij} + u_j + e_{ij}$$

$$u_j \sim \mathrm{N}(0, \sigma_u^2)$$

$$e_{ij} \sim \mathrm{N}(0, \sigma_e^2)$$

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

Include a random slope

$$normexam_{ij} = \beta_0 + \beta_1 standlrt_{ij} + \beta_2 girl_{ij} + u_{0j} + u_{1j} standlrt_{ij} + e_{ij}$$

$$\begin{pmatrix} u_{0j} \\ u_{1j} \end{pmatrix} \sim \mathrm{N}\left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2_{u0} & \\ \sigma_{u01} & \sigma^2_{u1} \end{pmatrix} \right\}$$

$$e_{ij} \sim \mathrm{N}(0, \sigma^2_e)$$

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

Allow for level 1 heteroskedasticity

$$normexam_{ij} = \beta_0 + \beta_1 standlrt_{ij} + \beta_2 girl_{ij} + u_{0j} + u_{1j} standlrt_{ij}$$

$$+ e_{2ij} girl_{ij} + e_{3ij} boy_{ij}$$

$$\begin{pmatrix} u_{0j} \\ u_{1j} \end{pmatrix} \sim \mathrm{N}\left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2_{u0} & \\ \sigma_{u01} & \sigma^2_{u1} \end{pmatrix} \right\}$$

$$\begin{pmatrix} e_{2ij} \\ e_{3ij} \end{pmatrix} \sim \mathrm{N}\left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2_{e2} & \\ 0 & \sigma^2_{e3} \end{pmatrix} \right\}$$

```
. runmlwin normexam cons standlrt girl, ///
      level2(school: cons standlrt) ///
      level1(student: girl boy, diagonal)
```

Retrieve the level 2 residuals

$$normexam_{ij} = \beta_0 + \beta_1 standlrt_{ij} + \beta_2 girl_{ij} + u_{0j} + u_{1j} standlrt_{ij}$$

$$+ e_{2ij} girl_{ij} + e_{3ij} boy_{ij}$$

$$\begin{pmatrix} u_{0j} \\ u_{1j} \end{pmatrix} \sim \mathrm{N}\left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2_{u0} & \\ \sigma_{u01} & \sigma^2_{u1} \end{pmatrix} \right\}$$

$$\begin{pmatrix} e_{2ij} \\ e_{3ij} \end{pmatrix} \sim \mathrm{N}\left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2_{e2} & \\ 0 & \sigma^2_{e3} \end{pmatrix} \right\}$$

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

Do not pause in MLwiN

$$normexam_{ij} = \beta_0 + \beta_1 standlrt_{ij} + \beta_2 girl_{ij} + u_{0j} + u_{1j} standlrt_{ij}$$

$$+ e_{2ij} girl_{ij} + e_{3ij} boy_{ij}$$

$$\begin{pmatrix} u_{0j} \\ u_{1j} \end{pmatrix} \sim \mathrm{N}\left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{u0}^2 & \\ \sigma_{u01} & \sigma_{u1}^2 \end{pmatrix} \right\}$$

$$\begin{pmatrix} e_{2ij} \\ e_{3ij} \end{pmatrix} \sim \mathrm{N}\left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{e2}^2 & \\ 0 & \sigma_{e3}^2 \end{pmatrix} \right\}$$

```
. runmlwin normexam cons standlrt girl, ///
      level2(school: cons standlrt, residuals(u)) ///
      level1(student: girl boy, diagonal) nopause
```

Stata/MP 11.2 - http://www.bristol.ac.uk/cmm/media/runmlwin/tutorial.dta - [Results]

File Edit Data Graphics Statistics User Window Help

MLwiN 2.23 multilevel model
Normal response model
Estimation algorithm: **IGLS**

Number of obs = **4059**

Level Variable	No. of Groups	Observations per Group Minimum	Average	Maximum
school	**65**	**2**	**62.4**	**198**

Run time (seconds) = **1.61**
Number of iterations = **4**
Log likelihood = **-4640.71**
Deviance = **9281.4199**

normexam	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
cons	-.111534	.0433072	-2.58	0.010	-.1964145	-.0266536
standlrt	.5529361	.0200758	27.54	0.000	.5135882	.5922841
girl	.1752785	.0324156	5.41	0.000	.1117451	.238812

Random-effects Parameters	Estimate	Std. Err.	[95% Conf.	Interval]
Level 2:				
var(cons)	.0862511	.017175	.0525887	.1199135
cov(cons,standlrt)	.0190537	.0066789	.0059632	.0321441
var(standlrt)	.0148919	.0044702	.0061304	.0236534
Level 1:				
var(girl)	.5251641	.0152836	.4952088	.5551194
var(boy)	.5874345	.0209983	.5462786	.6285904

C:\Users\gl9158\Documents

CAP NUM OVR

Stata/MP 11.2 - http://www.bristol.ac.uk/cmm/media/runmlwin/tutorial.dta - [Results]

File Edit Data Graphics Statistics User Window Help

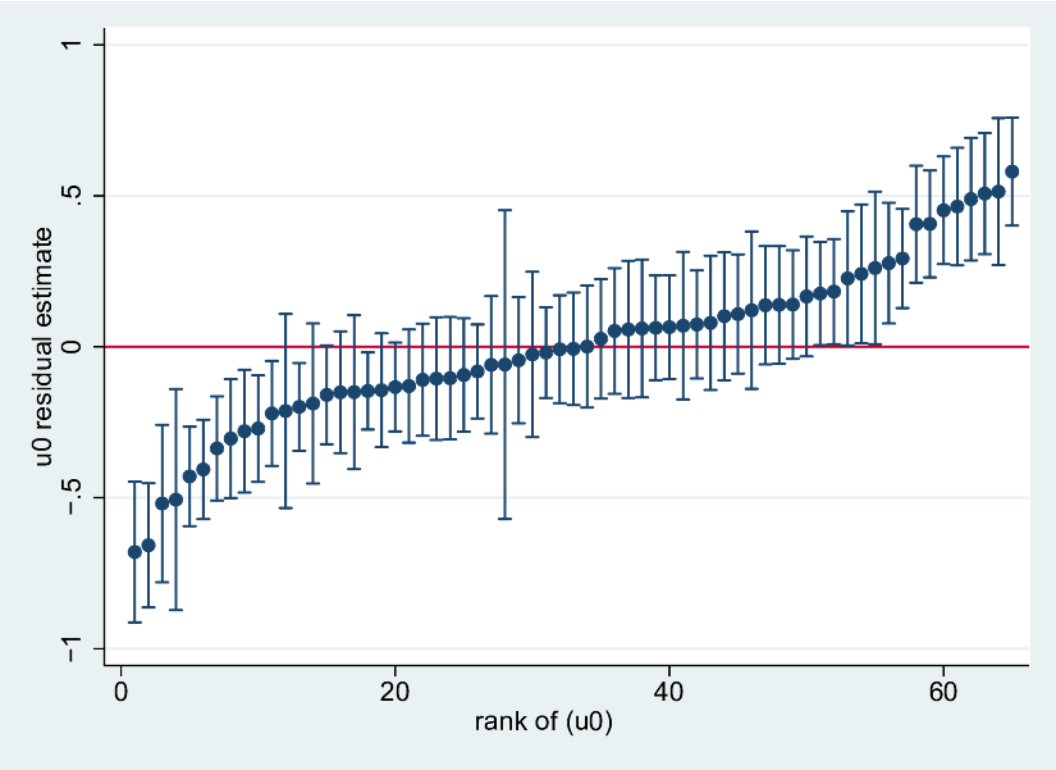
```
. test [RP1]var(girl) = [RP1]var(boy)

 ( 1)  [RP1]var(girl) - [RP1]var(boy) = 0

           chi2(  1) =    5.74
         Prob > chi2 =    0.0166
```

C:\Users\gl9158\Documents

CAP NUM OVR




```
. egen u0rank = rank(u0)

. serrbar u0 u0se u0rank, scale(1.96) yline(0)
```

runmlwin vs. xtmixed

- 100 simulated data sets
- 500,000 students in 20,000 classrooms in 2,500 schools in 100 districts

$$y_{ijkl} = \beta_0 + \beta_1 x_{ijkl} + \beta_2 x_{jkl} + \beta_3 x_{kl} + \beta_4 x_l + f_l + v_{kl} + u_{jkl} + e_{ijkl}$$

$$f_l \sim \mathrm{N}(0, \sigma_f^2), \qquad v_{kl} \sim \mathrm{N}(0, \sigma_v^2), \qquad u_{jkl} \sim \mathrm{N}(0, \sigma_u^2), \qquad e_{ijkl} \sim \mathrm{N}(0, \sigma_e^2)$$

Level	Seconds	β_0	β_1	β_2	β_3	β_4	σ_f^2	σ_v^2	σ_u^2	σ_e^2
True values		0	1	1	1	1	1	1	1	1
`xtmixed`	132	0.99	1.00	1.00	0.99	0.98	0.99	1.00	1.00	1.00
`xtmixed, emonly`	35	0.99	1.00	1.00	0.99	0.98	0.99	1.00	1.00	1.00
`runmlwin`	7	0.99	1.00	1.00	0.99	0.98	0.99	1.00	1.00	1.00

MULTILEVEL MODELS FOR BINARY RESPONSES

Guatemalan immunization campaign

- Child immunization data
- First analysed by Pebley, Goldman and Rodriguez (1996)
- Motivated a simulation study by Rodriguez and Goldman (1995) to analyse alternative estimation methods for binary response multilevel models
- 100 simulated data sets
- 2159 children within 1595 mothers within 161 communities
- Three-level random intercepts binary response multilevel model
- Simulation study data reanalysed by Browne and Draper (2006), Goldstein and Rasbash (1996), Rabe-Hesketh and Skrondal (2008), Rodriguez and Goldman (2001, 2008) and others

Three-level binary response model

$$y_{ijk} \sim \text{Binomial}(1, \pi_{ijk})$$

$$\text{logit}(\pi_{ijk}) = \beta_0 + \beta_1 x_{1ijk} + \beta_2 x_{2jk} + \beta_3 x_{3k} + v_k + u_{jk}$$

$$v_k \sim \text{N}(0, \sigma_v^2), \qquad u_{jk} \sim \text{N}(0, \sigma_u^2)$$

```
. runmlwin y cons x1 x2 x3, ///
      level3(community: cons) ///
      level2(mother: cons) ///
      level1(child:) ///
      discrete(d(binomial) l(logit) denom(cons) pql2) ///
      nopause
```

Three-level binary response model

$$y_{ijk} \sim \text{Binomial}(1, \pi_{ijk})$$

$$\text{logit}(\pi_{ijk}) = \beta_0 + \beta_1 x_{1ijk} + \beta_2 x_{2jk} + \beta_3 x_{3k} + v_k + u_{jk}$$

$$v_k \sim \text{N}(0, \sigma_v^2), \qquad u_{jk} \sim \text{N}(0, \sigma_u^2)$$

```
. runmlwin y cons x1 x2 x3, ///
      level3(community: cons) ///
      level2(mother: cons) ///
      level1(child:) ///
      discrete(d(binomial) l(logit) denom(cons)) ///
      mcmc(on) initsprevious nopause
```

`runmlwin` vs. `xtmelogit`

Estimation method	Seconds	β_0	β_1	β_2	β_3	σ_v^2	σ_u^2
True values		0.67	1	1	1	1	1
`xtmelogit, intpoints(1)`	53	0.62	0.90	0.91	0.95	0.78	0.35
`xtmelogit, intpoints(2)`	54	0.64	0.93	0.93	0.98	0.83	0.55
`xtmelogit, intpoints(3)`	56	0.67	0.97	0.98	1.02	0.92	0.87
`xtmelogit, intpoints(4)`	67	0.68	0.98	0.99	1.04	0.95	0.98
`xtmelogit, intpoints(7)`	125	0.68	0.98	0.99	1.04	0.96	0.98
`runmlwin, pql2`	13	0.65	0.94	0.95	1.00	0.90	0.58
`runmlwin, b(200) c(1000)`	19	0.67	0.98	0.98	1.03	0.99	0.93
`runmlwin, b(500) c(5000)`	34	0.67	0.98	0.99	1.03	1.00	0.95

- PQL2 performs better than Laplace, but both are severely biased
- MCMC performs as well as adaptive quadrature, but is considerably quicker
- These differences will be much greater in data sets with random slopes
- Similar results will apply for `runmlwin` vs. `xtmepoisson`

MCMC ESTIMATION

The `runmlwin` command syntax

$$binexam_{ij} \sim \text{Binomial}(1, \pi_{ij})$$

$$logit(\pi_{ij}) = \beta_0 + \beta_1 standlrt_{ij} + \beta_2 girl_{ij} + u_{0j} + u_{1j} standlrt_{ij}$$

$$\begin{pmatrix} u_{0j} \\ u_{1j} \end{pmatrix} \sim \text{N}\left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2_{u0} & \\ \sigma_{u01} & \sigma^2_{u1} \end{pmatrix} \right\}$$

```
. matrix a = (1,1,1,1,1,1,1)

. runmlwin binexam cons standlrt girl, ///
      level2(school: cons standlrt) ///
      level1(student:) ///
      discrete(d(binomial) l(logit) de(cons)) ///
      mcmc(burnin(500) chain(5000)) initsb(a) nopause
```

Stata/MP 11.2 - http://www.bristol.ac.uk/cmm/media/runmlwin/tutorial.dta - [Results]

File Edit Data Graphics Statistics User Window Help

MLwiN 2.23 multilevel model
Binomial logit response model
Estimation algorithm: **MCMC**

Number of obs = **4059**

Level Variable	No. of Groups	Observations per Group Minimum	Average	Maximum
school	**65**	**2**	**62.4**	**198**

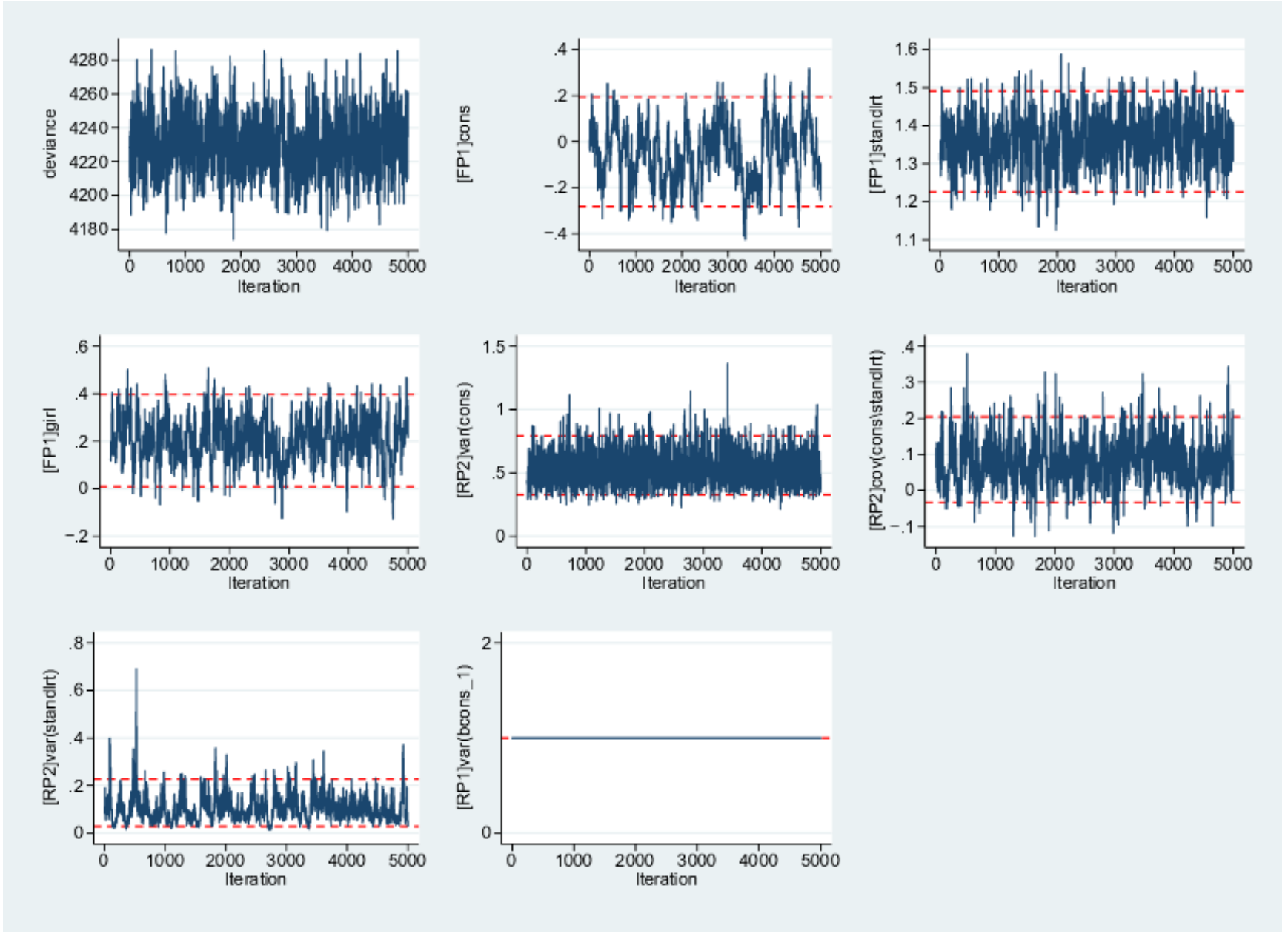
Burnin	=	**500**
Chain	=	**5000**
Thinning	=	**1**
Run time (seconds)	=	**29.1**
Deviance (dbar)	=	**4229.40**
Deviance (thetabar)	=	**4154.23**
Effective no. of pars (pd)	=	**75.17**
Bayesian DIC	=	**4304.57**

binexam	Mean	Std. Dev.	z	ESS	[95% Cred.	Interval]
cons	-.0512708	.1281822	-0.40	65	-.2820003	.1941637
standlrt	1.358827	.067	20.28	425	1.226268	1.491531
girl	.2053954	.0986879	2.08	181	.0094571	.3976289

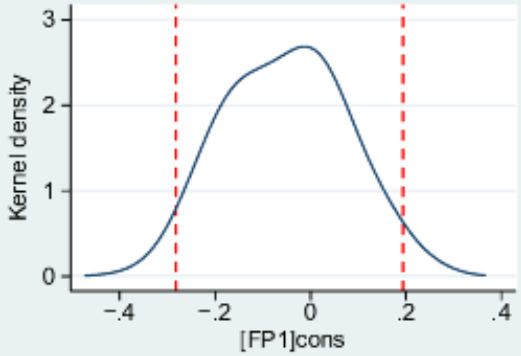
Random-effects Parameters	Mean	Std. Dev.	ESS	[95% Cred.	Int]
Level 2: school					
var(cons)	.5198903	.1202148	1139	.328118	.793948
cov(cons,standlrt)	.076277	.0594963	254	-.0333807	.203841
var(standlrt)	.1020427	.0522664	137	.0279808	.2268086

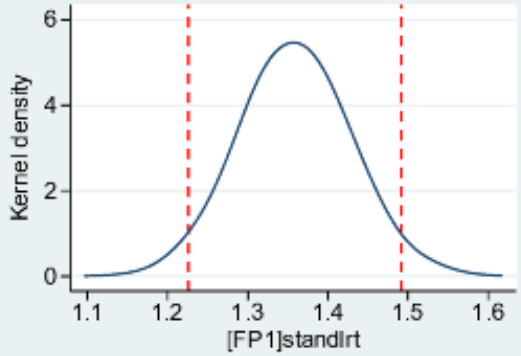
Q:\C-modelling\runmlwin\presentations\2011-09-16 London

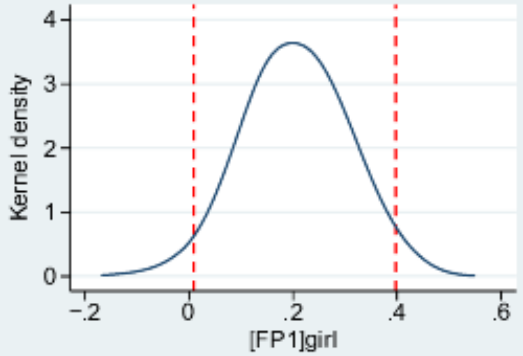
CAP NUM OVR

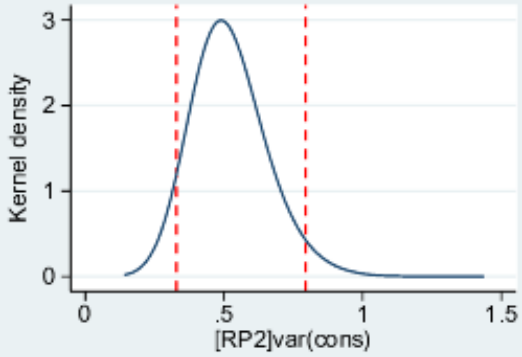


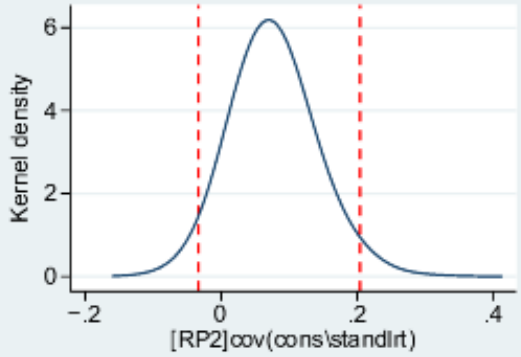

```
. mcmcsum, trajectories
```

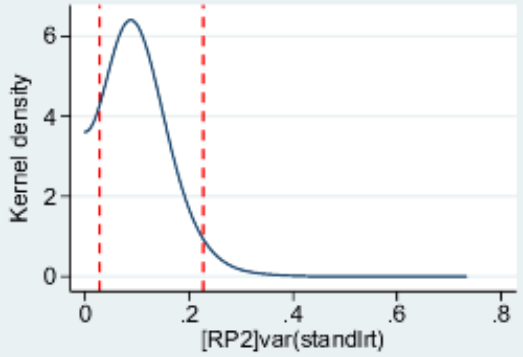




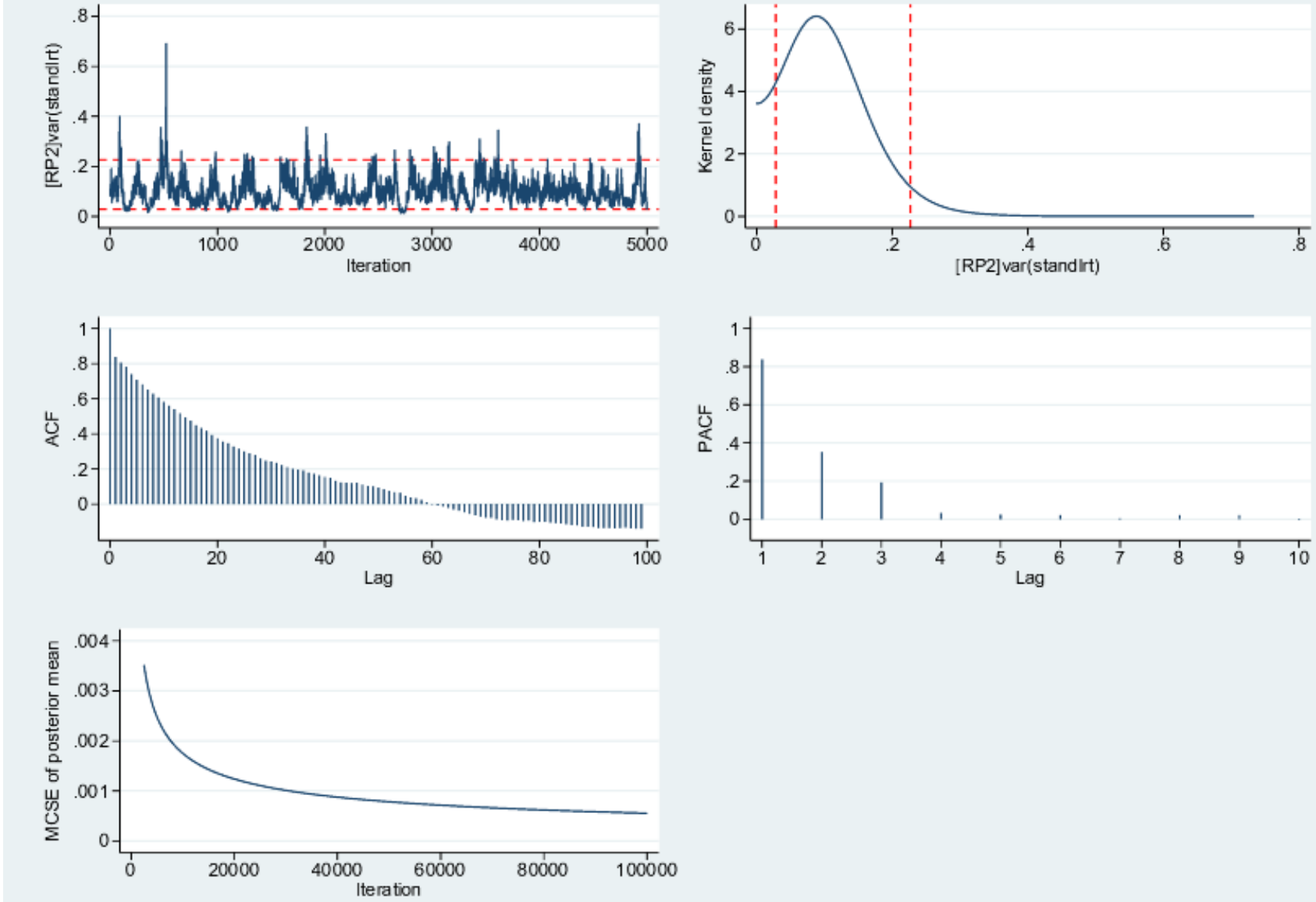









```
. mcmcsum, densities
```




```
. mcmcsum [RP2]var(standlrt), fiveplot
```

Stata/MP 11.2 - http://www.bristol.ac.uk/cmm/media/runmlwin/tutorial.dta - [Results]

File Edit Data Graphics Statistics User Window Help

```
. mcmcsum [RP2]var(standlrt)

                        [RP2]var(standlrt)
------------------------------------------------------------------------------

                                    Percentiles
Mean              .1020427      0.5%    .0189533   Thinned Chain Length     5000
MCSE of Mean      .0024763      2.5%    .0279808   Effective Sample Size     137
Std. Dev.         .0522664        5%    .0351138   Raftery Lewis (2.5%)    66475
Mode                .08801       25%    .0652843   Raftery Lewis (97.5%)   26115
P(mean)                  0                       Brooks Draper (mean)     4714
P(mode)                  0       50%    .0934135
P(median)                0
                                75%     .128915
                                95%    .1999984
                             97.5%    .2268086
                             99.5%    .3022345
------------------------------------------------------------------------------

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

Q:\C-modelling\runmlwin\presentations\2011-09-16 London

CAP NUM OVR

EXPORT MODELS TO WINBUGS

The `runmlwin` command syntax

$$binexam_{ij} \sim \text{Binomial}(1, \pi_{ij})$$

$$logit(\pi_{ij}) = \beta_0 + \beta_1 standlrt_{ij} + \beta_2 girl_{ij} + u_{0j} + u_{1j} standlrt_{ij}$$

$$\begin{pmatrix} u_{0j} \\ u_{1j} \end{pmatrix} \sim \text{N}\left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2_{u0} & \\ \sigma_{u01} & \sigma^2_{u1} \end{pmatrix} \right\}$$

```
. runmlwin binexam cons standlrt girl, ///
      level2(school: cons standlrt) ///
      level1(student:) ///
      discrete(d(binomial) l(logit) de(cons)) ///
      mcmc(b(500) c(5000) savewinbugs(model(m.txt) ///
      inits(i.txt) data(d.txt) nofit)) initsprevious
```

Viewer (#1) [view m.txt]

view m.txt

Advice | Contents | What's New | News

```
# WINBUGS 1.4 code generated from MLwiN program

#----MODEL Definition----------------

model
{
# Level 1 definition
for(i in 1:N) {
binexam[i] ~ dbin(p[i],denom[i])
logit(p[i]) <- beta[1] * cons[i]
+ beta[2] * standlrt[i]
+ beta[3] * girl[i]
+ u2[school[i],1] * cons[i]
+ u2[school[i],2] * standlrt[i]
}
# Higher level definitions
for (j in 1:n2) {
u2[j,1:2] ~ dmnorm(zero2[1:2],tau.u2[1:2,1:2])
}
# Priors for fixed effects
for (k in 1:3) { beta[k] ~ dflat() }
# Priors for random terms
for (i in 1:2) {zero2[i] <- 0}
tau.u2[1:2,1:2] ~ dwish(R2[1:2, 1:2],2)
sigma2.u2[1:2,1:2] <- inverse(tau.u2[,])
}
```

RESOURCES TO HELP YOU LEARN RUNMLWIN

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*])

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)

CMM software support

- 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.*

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 ianna » Fri Apr 08, 2011 9:20 am	5	293	by ChrisCharlton Thu Jun 09, 2011 10:09 am

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