Running MLwiN from within Stata: the `runmlwin` command

Modern Modeling Methods (M3) Conference, University of Connecticut
26th May 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 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. Constraints allowing models such as the social relations models and behavioural genetics models to be formulated as multilevel models

3. Fast estimation via classical and Bayesian methods

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

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

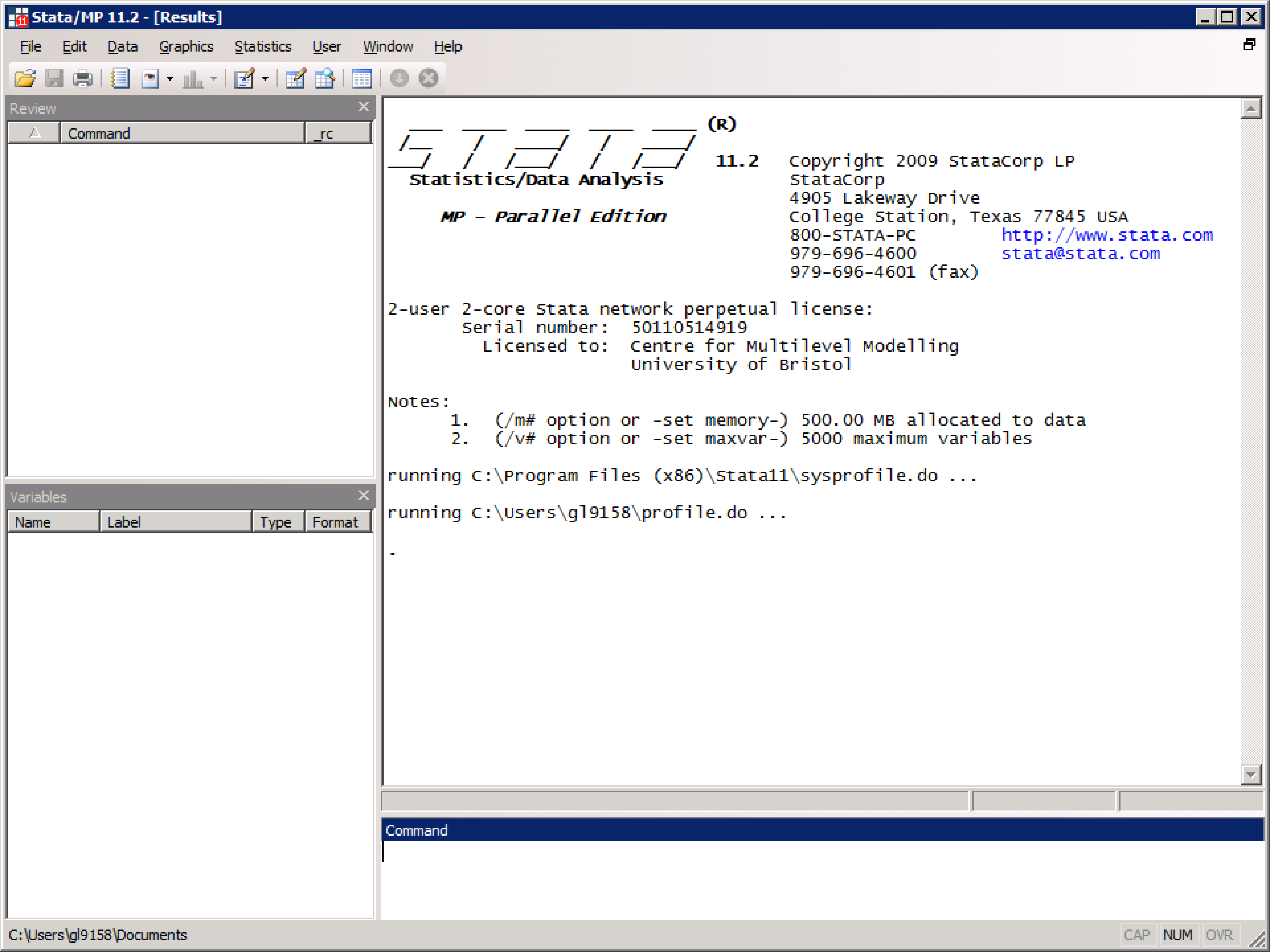
Examples

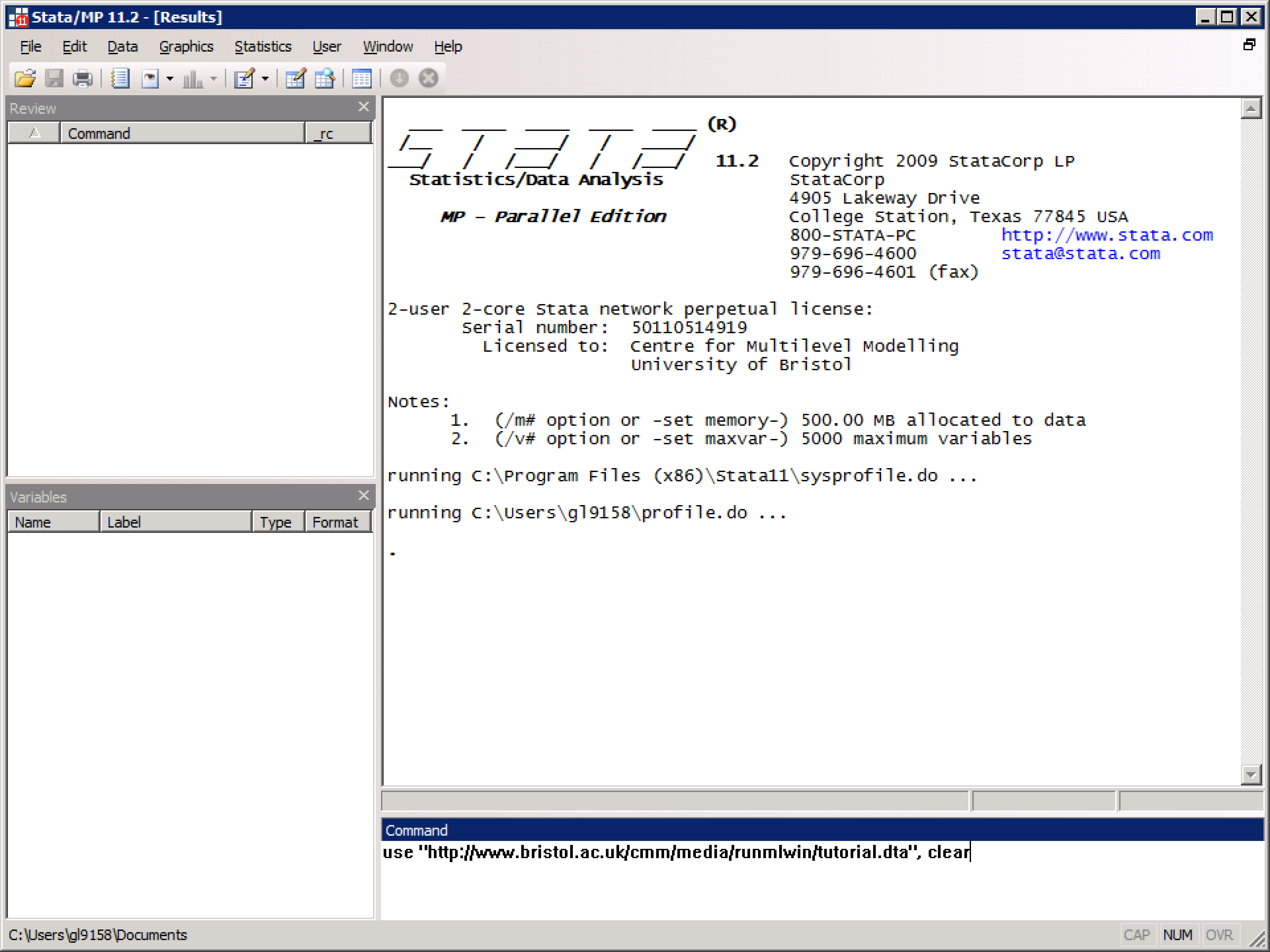
1. Two-level multilevel model
2. Growth curve models
3. Multilevel models for binary responses
4. Simulation studies are easy
5. MCMC estimation
6. Cross-classified models
7. Spatial multilevel models
8. Export models to WinBUGS
9. Work efficiently
10. Resources to help you learn `runmlwin`

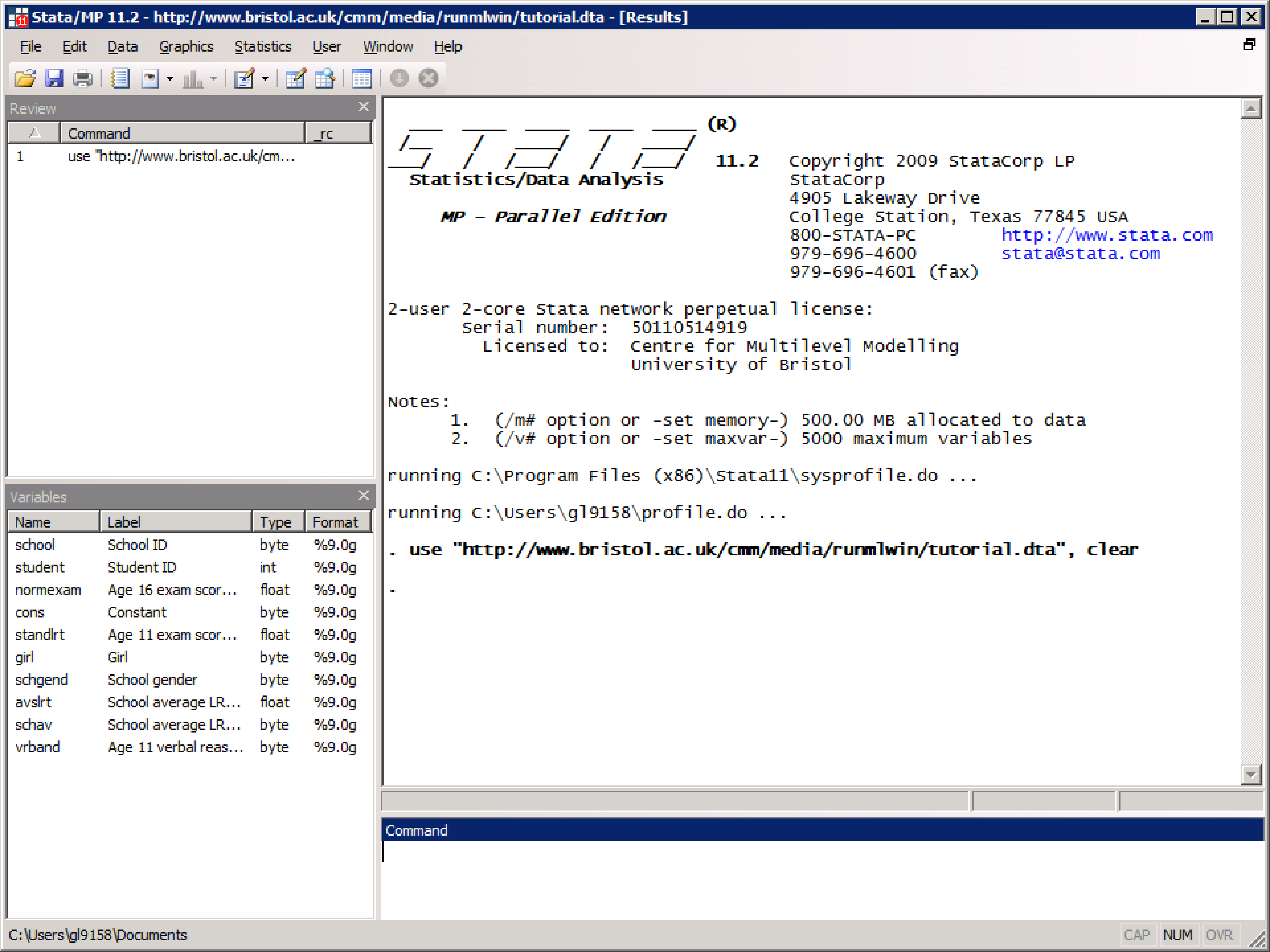
1. TWO-LEVEL MULTILEVEL MODELS

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 (2003), Rabe-Hesketh and Skrondal (2008), Rasbash et al. (2009) and others
- 4059 students nested within 65 schools


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
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 ...
.
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
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
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 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.bristol.ac.uk/cmm/media/runmlwin/tutorial.dta", clear
.
Command
C:\Users\gl9158\Documents
CAP NUM OVR

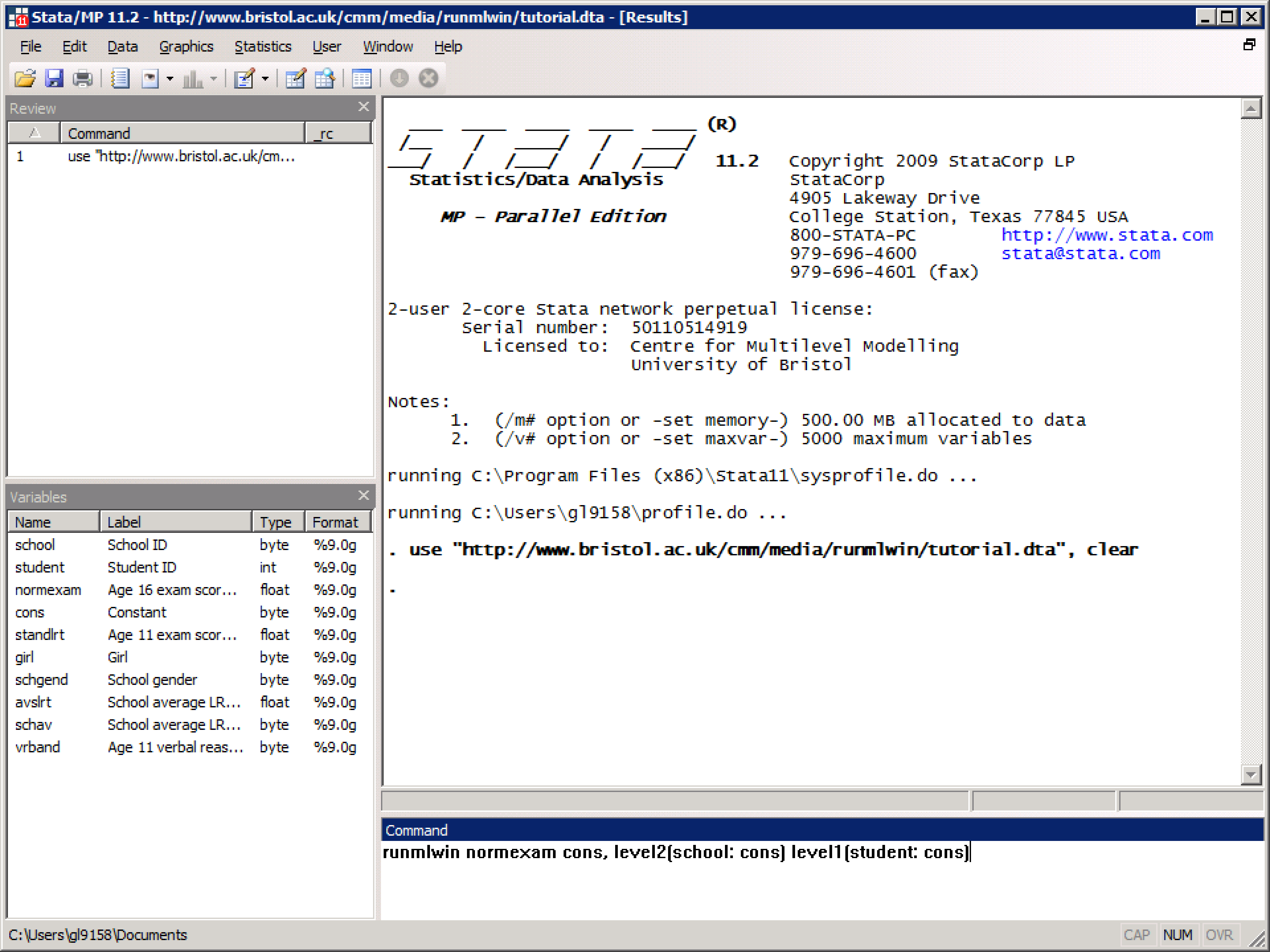
The `runmlwin` command syntax

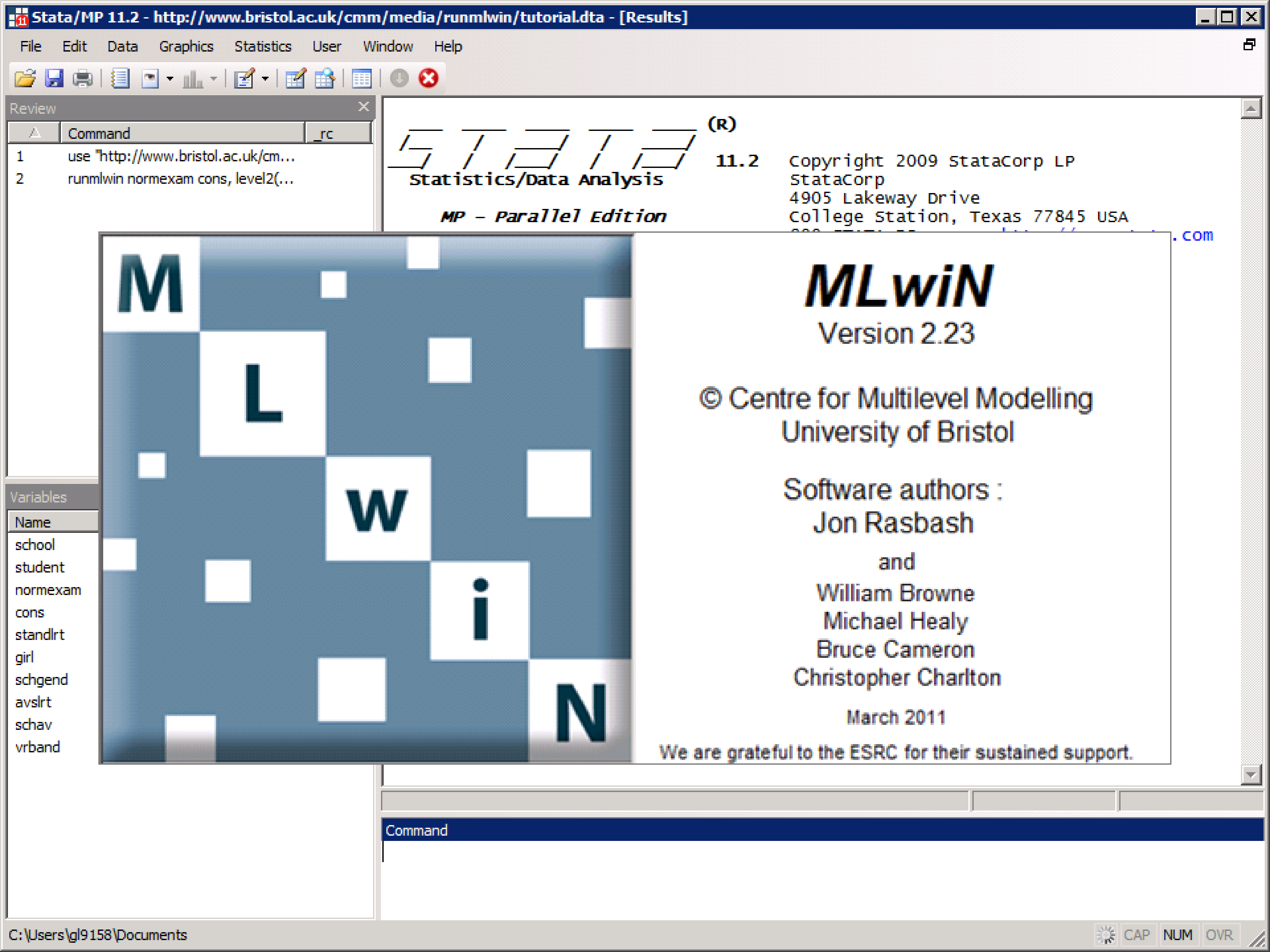
$$normexam_{ij} = \beta_0 + u_j + e_{ij}$$

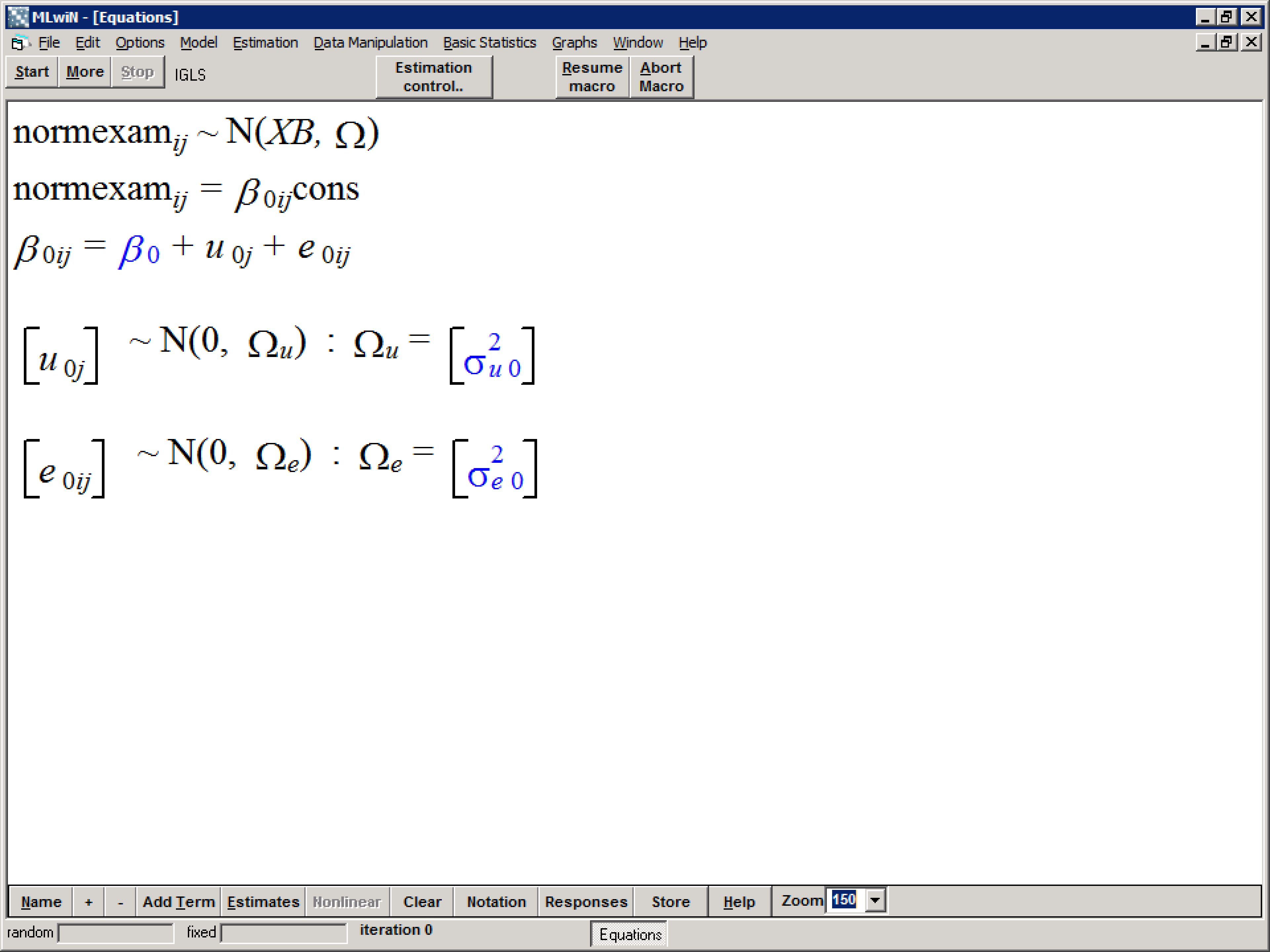
$$u_j \sim \mathrm{N}(0, \sigma_u^2)$$

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

```
. 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...
(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.bristol.ac.uk/cmm/media/runmlwin/tutorial.dta", clear
.
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
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(...
(R)
11.2
Statistics/Data Analysis
MP – Parallel Edition
Copyright 2009 StataCorp LP
StataCorp
4905 Lakeway Drive
College Station, Texas 77845 USA
Variables
Name
school
student
normexam
cons
standlrt
girl
schgend
avslrt
schav
vrband
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
$\text{normexam}_{ij} \sim \text{N}(XB, \Omega)$
$\text{normexam}_{ij} = \beta_{0ij}\text{cons}$
$\beta_{0ij} = \beta_0 + u_{0j} + e_{0ij}$
$\left[u_{0j}\right] \sim \text{N}(0, \Omega_u) : \Omega_u = \left[\sigma^2_{u0}\right]$
$\left[e_{0ij}\right] \sim \text{N}(0, \Omega_e) : \Omega_e = \left[\sigma^2_{e0}\right]$
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

$$\text{normexam}_{ij} \sim \mathrm{N}(XB,\ \Omega)$$

$$\text{normexam}_{ij} = \beta_{0ij}\text{cons}$$

$$\beta_{0ij} = -0.013(0.054) + u_{0j} + e_{0ij}$$

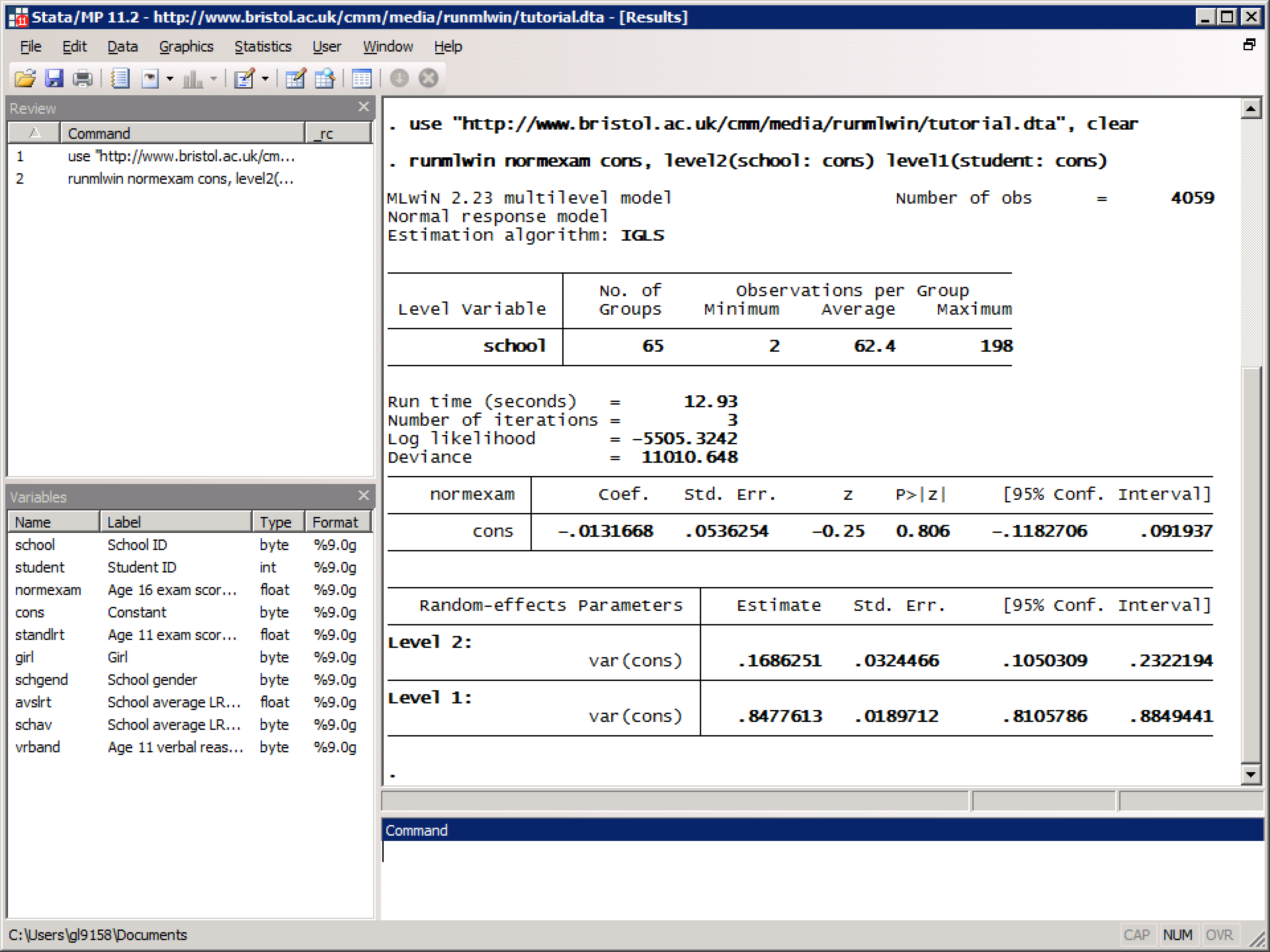
$$\left[u_{0j}\right] \sim \mathrm{N}(0,\ \Omega_u)\ :\ \Omega_u = \left[0.169(0.032)\right]$$

$$\left[e_{0ij}\right] \sim \mathrm{N}(0,\ \Omega_e)\ :\ \Omega_e = \left[0.848(0.019)\right]$$

*-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

-----------------------------------------------------------
                |   No. of       Observations per Group
 Level Variable |   Groups    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

Retrieve the level 2 residuals

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

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

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

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

Do not pause in MLwiN

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

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

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

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

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

File Edit Data Graphics Statistics User Window Help

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

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

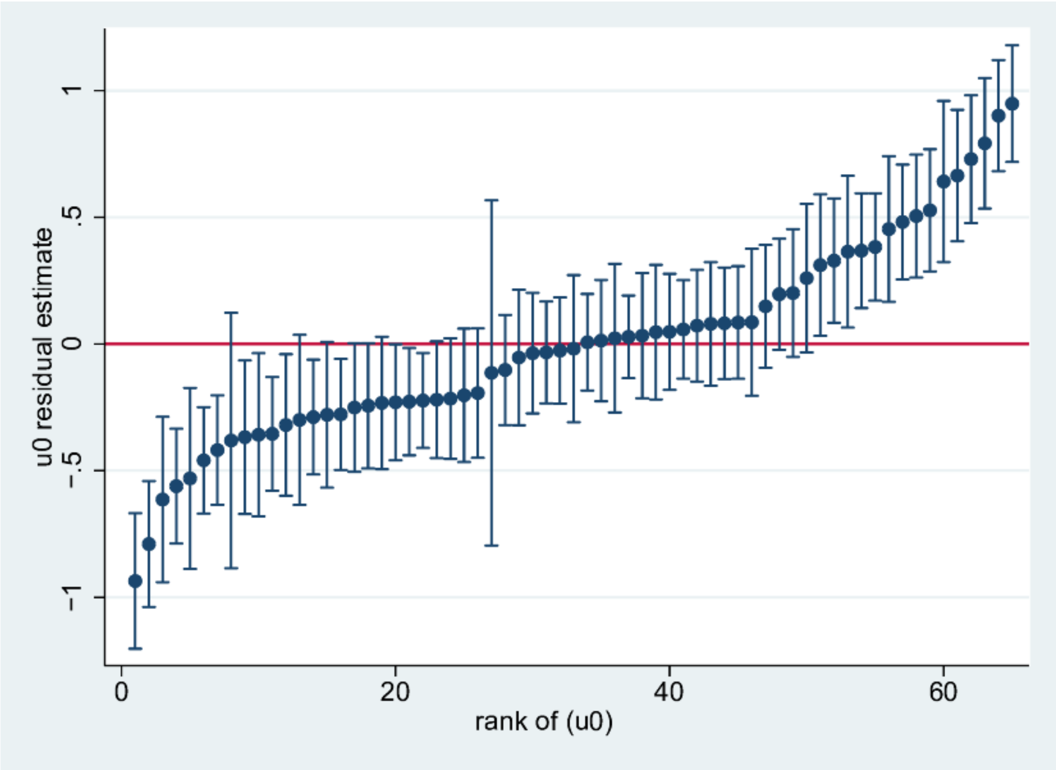
-----------------------------------------------------------
                |   No. of       Observations per Group
 Level Variable |   Groups    Minimum    Average    Maximum
----------------+------------------------------------------
         school |       65          2       62.4        198
-----------------------------------------------------------

Run time (seconds)   =       1.47
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
------------------------------------------------------------------------------

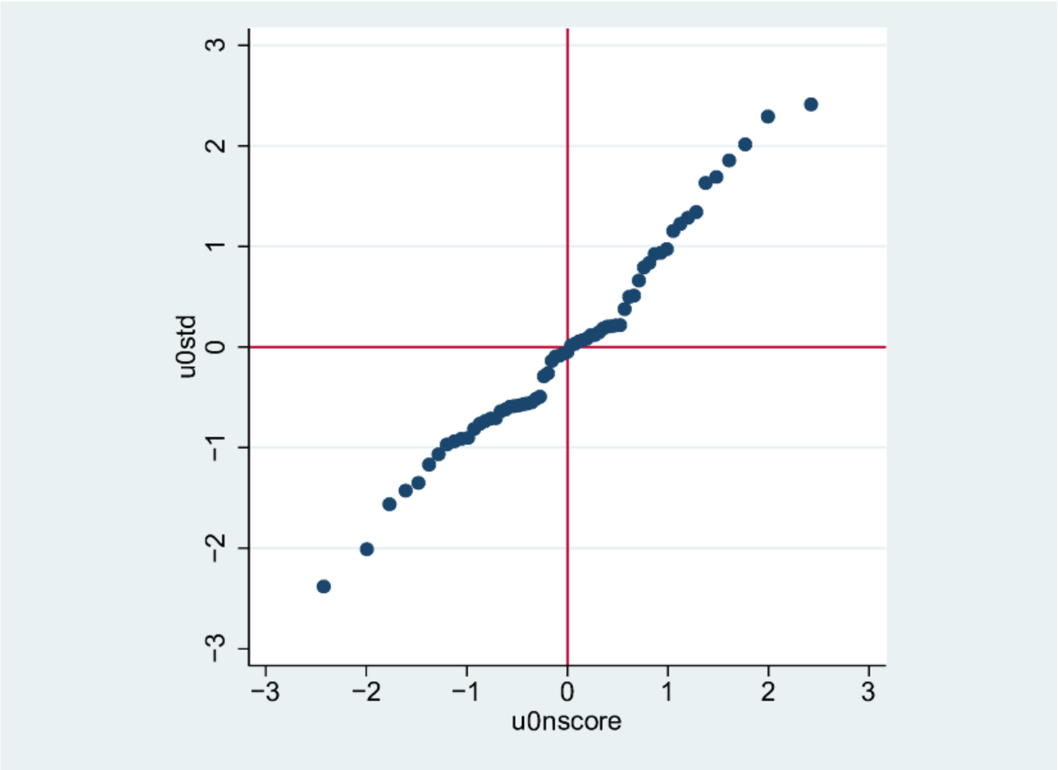
.
```

C:\Users\gl9158\Documents   CAP NUM OVR




```
. egen u0rank = rank(u0)

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




```
. summarize u0

. generate u0std = (u0 - r(mean))/r(sd)

. generate u0uniform = (u0rank - 0.5)/_N

. generate u0nscore = invnorm(u0uniform)

. scatter u0std u0nscore, yline(0) xline(0) ///
      ylabel(-3(1)3) xlabel(-3(1)3) aspectratio(1)
```

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) nopause
```

Include a random slope

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

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

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

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

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) 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                     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)   =       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

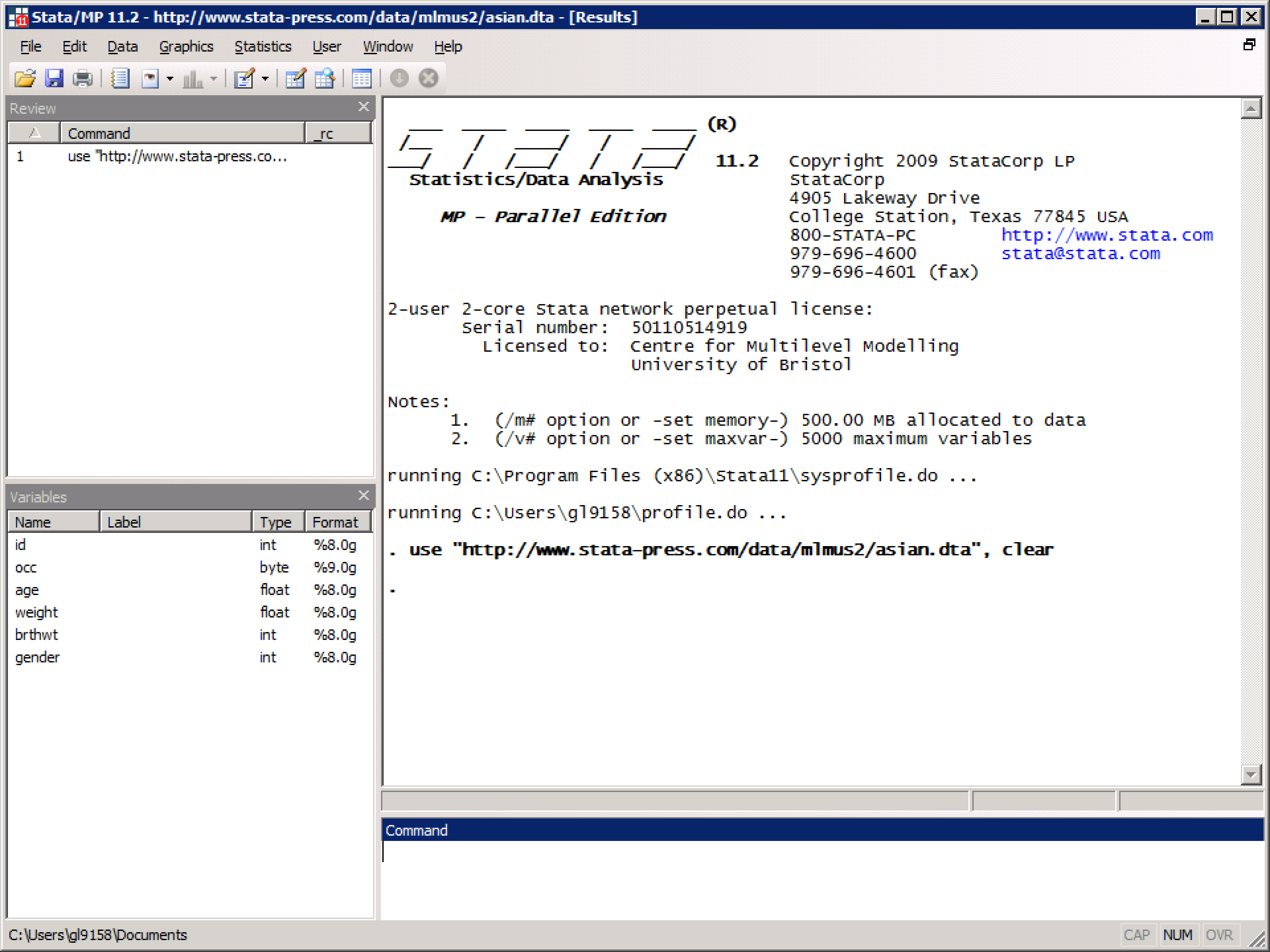
.
.
.
.
.
.
.
.
.
.
.
.
.
.
.
.
.
.
.
.
.
.
.
.
.
.
.
```

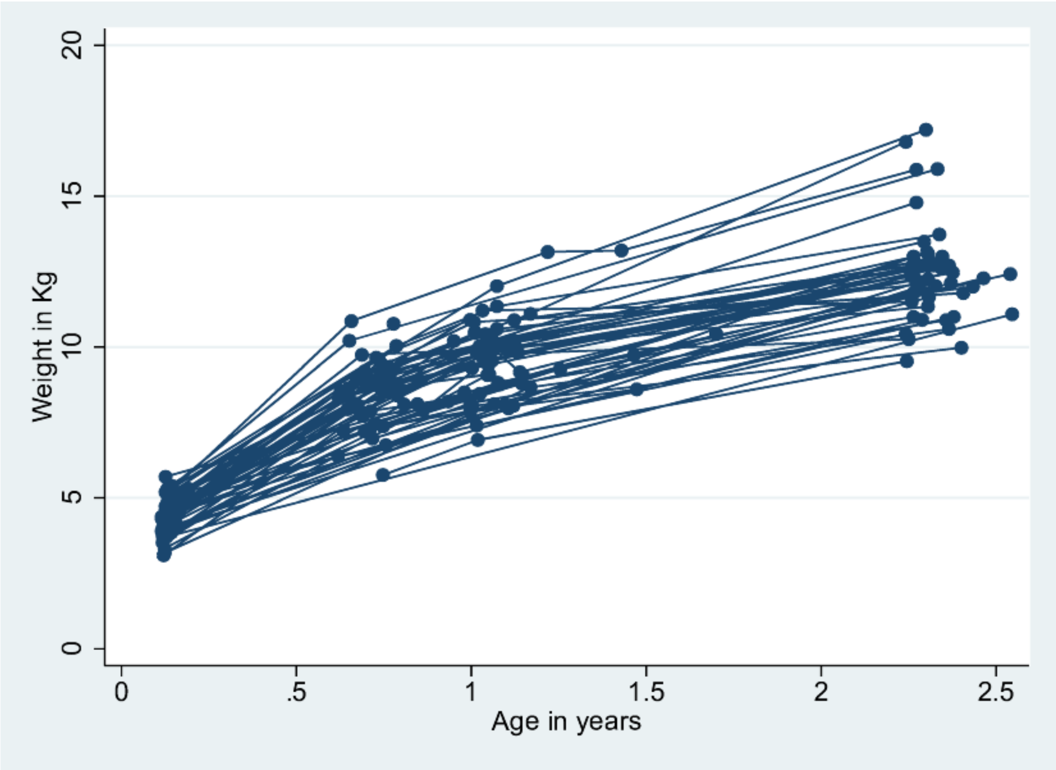
C:\Users\gl9158\Documents CAP NUM OVR

2. GROWTH CURVE MODELS

Child weight data

- Weight gain of Asian children in a British community
- 68 children, one to five measurements per child
- First analysed by Goldstein (1986)
- Re-analysed by Rabe-Hesketh and Skrondal (2008) and others


Stata/MP 11.2 - http://www.stata-press.com/data/mlmus2/asian.dta - [Results]
File Edit Data Graphics Statistics User Window Help
Review
Command
_rc
1 use "http://www.stata-press.co...
Variables
Name Label Type Format
id int %8.0g
occ byte %9.0g
age float %8.0g
weight float %8.0g
brthwt int %8.0g
gender int %8.0g
(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/asian.dta", clear
.
Command
C:\Users\gl9158\Documents
CAP NUM OVR




- ```
  graph twoway ///
       (connect weight age, connect(ascending)), ///
       ytitle("Weight in Kg") xtitle("Age in years")
  ```
  ```

Growth curve model

$$weight_{ij} = \beta_0 + \beta_1 age_{ij} + \beta_2 age_{ij}^2 + u_{0j} + u_{1j} age_{ij} + e_{ij}$$

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

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

```
. runmlwin weight cons age age2, ///
      level2(id: cons age, residuals(u)) ///
      level1(occ: cons) nopause
```

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

File Edit Data Graphics Statistics User Window Help

```
MLwiN 2.23 multilevel model                    Number of obs      =        198
Normal response model
Estimation algorithm: IGLS
```

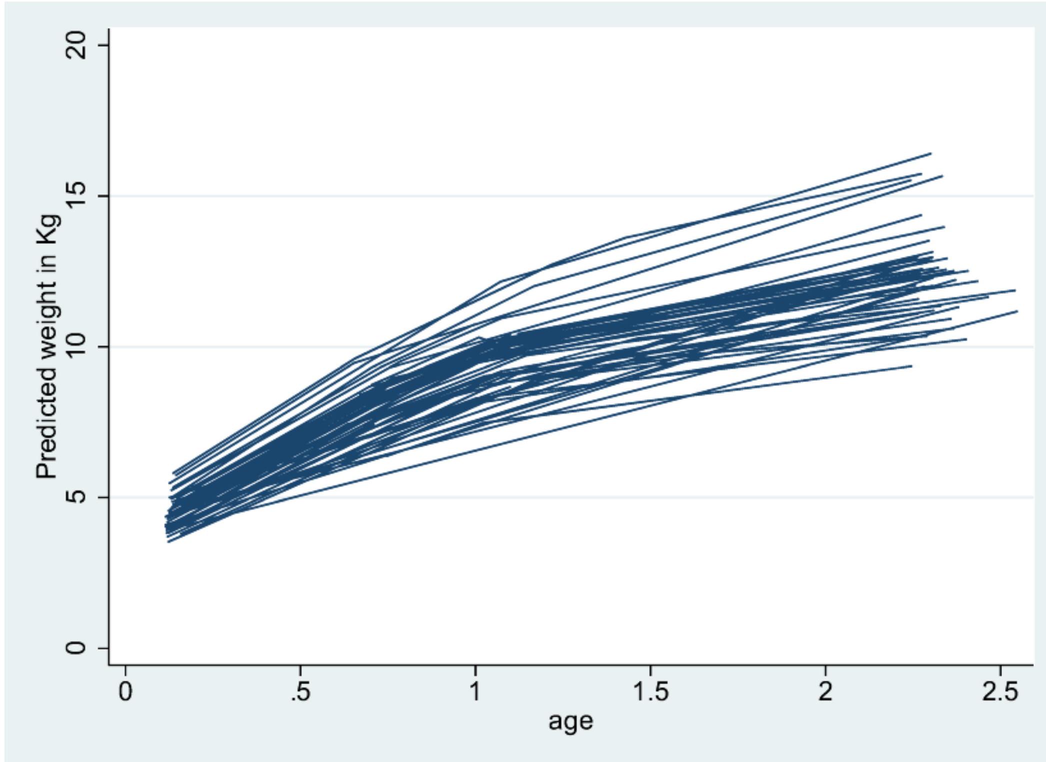
Level Variable	No. of Groups	Observations per Group Minimum	Average	Maximum
id	68	1	2.9	5

```
Run time (seconds)    =       1.62
Number of iterations =          7
Log likelihood        = -258.07785
Deviance              =   516.1557
```

weight	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
cons	3.494518	.1372489	25.46	0.000	3.225515	3.76352
age	7.704002	.2394275	32.18	0.000	7.234733	8.173271
age2	-1.660475	.0885319	-18.76	0.000	-1.833994	-1.486955

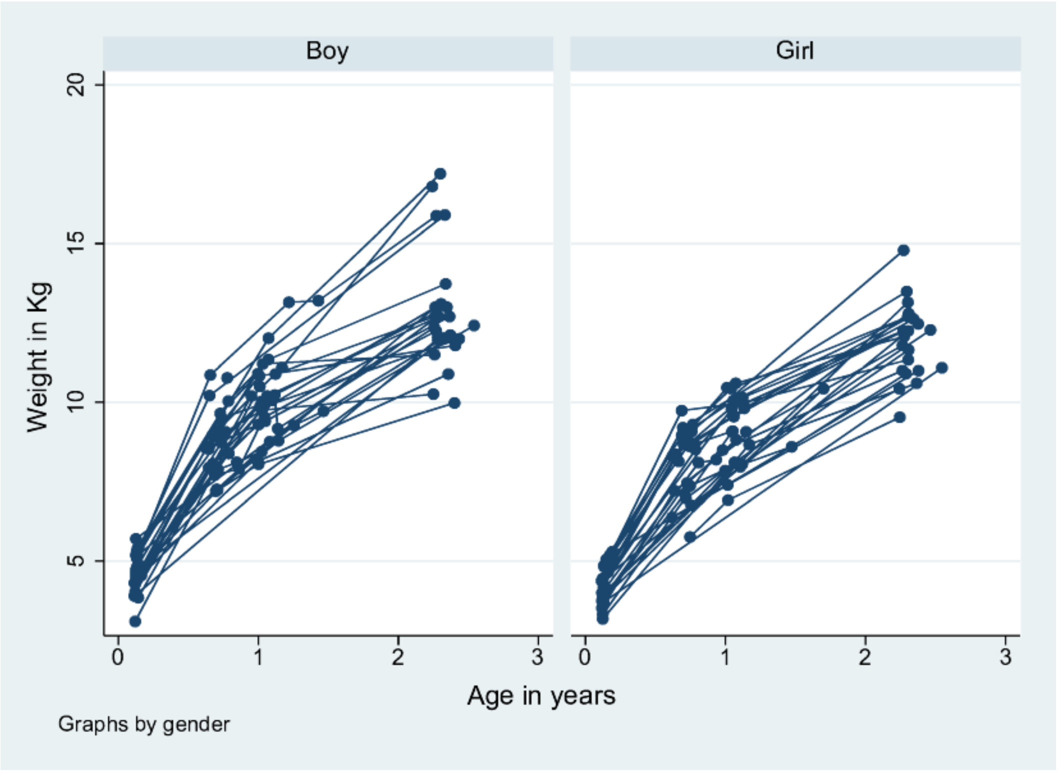
Random-effects Parameters	Estimate	Std. Err.	[95% Conf.	Interval]
Level 2:				
var(cons)	.4040045	.1412488	.1271619	.6808471
cov(cons,age)	.088273	.0812774	-.0710279	.2475738
var(age)	.2539857	.0858503	.0857222	.4222493
Level 1:				
var(cons)	.331641	.0532307	.2273107	.4359712

C:\Users\gl9158\Documents CAP NUM OVR




```
• generate prediction = ///
      _b[cons]*cons + _b[age]*age + _b[age2]*age2 ///
      + u0 + u1*age

• line prediction age, connect(a) ///
      ytitle("Predicted weight in Kg")
```



- `label define genderlabel 1 "Boy" 2 "Girl"`
- `label values gender genderlabel`
- graph twoway:

```
graph twoway (line weight age, connect(ascending)), ///
     by(gender) ///
     xtitle("Age in years") ytitle("Weight in Kg")
```

Growth curve model by gender

$$weight_{ij} = \beta_0 boy_j + \beta_1 boy_j \times age_{ij} + \beta_2 girl_j + \beta_3 girl_j \times age_{ij} + \beta_4 age_{ij}^2$$

$$+u_{0j} boy_j + u_{1j} boy_j \times age_{ij} + u_{2j} girl_j + u_{3j} girl_j \times age_{ij} + e_{0ij} boy_j + e_{2ij} girl_j$$

$$\begin{pmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \\ u_{3j} \end{pmatrix} \sim \mathrm{N} \left\{ \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{u0}^2 & & & \\ \sigma_{u01} & \sigma_{u1}^2 & & \\ 0 & 0 & \sigma_{u2}^2 & \\ 0 & 0 & \sigma_{u23} & \sigma_{u3}^2 \end{pmatrix} \right\}$$

$$\begin{pmatrix} e_{0ij} \\ e_{2ij} \end{pmatrix} \sim \mathrm{N} \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{e0}^2 & \\ 0 & \sigma_{e2}^2 \end{pmatrix} \right\}$$

```
. matrix a = (1,1,1,0,0,1,0,0,1,1)

. runmlwin weight boy boyXage girl girlXage age2, ///

      level2(id: boy boyXage girl girlXage, elements(a)) ///

      level1(occ: boy girl, diagonal) nopause
```

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

File Edit Data Graphics Statistics User Window Help

Level Variable	No. of Groups	Observations per Group Minimum	Average	Maximum
id	68	1	2.9	5

```
Run time (seconds)    =        1.58
Number of iterations  =           7
Log likelihood        = -247.49434
Deviance              =  494.98868
```

weight	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
boy	3.78267	.1563113	24.20	0.000	3.476305	4.089034
boyXage	7.728288	.2567359	30.10	0.000	7.225095	8.231481
girl	3.266411	.1796806	18.18	0.000	2.914244	3.618579
girlXage	7.502467	.2341932	32.04	0.000	7.043457	7.961477
age2	-1.624745	.0849193	-19.13	0.000	-1.791184	-1.458306

Random-effects Parameters	Estimate	Std. Err.	[95% Conf.	Interval]
Level 2:				
var(boy)	.1553577	.1659469	-.1698922	.4806076
cov(boy,boyXage)	.102065	.1232655	-.1395309	.3436609
var(boyXage)	.3869624	.1692804	.055179	.7187458
var(girl)	.5685636	.2111509	.1547155	.9824117
cov(girl,girlXage)	.0161196	.0864426	-.1533048	.185544
var(girlXage)	.0799457	.0608557	-.0393292	.1992206
Level 1:				
var(boy)	.4182827	.0929099	.2361826	.6003828
var(girl)	.2429176	.0555108	.1341183	.3517168

C:\Users\gl9158\Documents

CAP NUM OVR

3. MULTILEVEL MODELS FOR BINARY RESPONSES

Guatemalan immunization campaign

- Child immunization data
- 2159 children within 1595 mothers within 161 communities
- First analysed by Pebley, Goldman and Rodriguez (1996) and Rodriguez and Goldman (2001)
- Reanalysed by Rabe-Hesketh and Skrondal (2008) and others

Three-level binary response model

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

$$logit(\pi_{ijk}) = \beta_0 + \beta_1 kid2p_{ijk} + \beta_2 rural_k + \beta_3 pcInd81_k + v_k + u_{jk}$$

$$v_k \sim \text{N}(0, \sigma_u^2)$$

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

```
. runmlwin immun cons kid2p rural pcInd81, ///
      level3(cluster: cons) ///
      level2(mom: cons) ///
      level1(kid:) ///
      discrete(dist(binomial) link(logit) denom(cons)) ///
      nopause
```

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

File Edit Data Graphics Statistics User Window Help

```
MLwiN 2.23 multilevel model                    Number of obs      =      2159
Binomial logit response model
Estimation algorithm: IGLS, MQL1
```

Level Variable	No. of Groups	Observations per Group Minimum	Average	Maximum
cluster	161	1	13.4	55
mom	1595	1	1.4	3

```
Run time (seconds)   =       2.89
Number of iterations =          5
```

immun	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
cons	-.1433676	.1721252	-0.83	0.405	-.4807268	.1939915
kid2p	.9173057	.1179051	7.78	0.000	.6862159	1.148395
rural	-.5668908	.1480174	-3.83	0.000	-.8569995	-.276782
pcInd81	-.8460267	.1797028	-4.71	0.000	-1.198238	-.4938157

Random-effects Parameters	Estimate	Std. Err.	[95% Conf.	Interval]
Level 3:				
var(cons)	.2960818	.0772351	.1447038	.4474597
Level 2:				
var(cons)	.3674519	.1286113	.1153784	.6195254

C:\Users\gl9158\Documents   CAP NUM OVR

Refit the model using PQL2

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

$$logit(\pi_{ijk}) = \beta_0 + \beta_1 kid2p_{ijk} + \beta_2 rural_k + \beta_3 pcInd81_k + v_k + u_{jk}$$

$$v_k \sim \text{N}(0, \sigma_u^2)$$

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

```
. runmlwin immun cons kid2p rural pcInd81, ///
      level3(cluster: cons) ///
      level2(mom: cons) ///
      level1(kid:) ///
      discrete(d(binomial) l(logit) de(cons) pql2) ///
      initsprevious maxiterations(40) nopause
```

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

File Edit Data Graphics Statistics User Window Help

```
Model fitted using initial values specified as parameter estimates from previous model

MLwiN 2.23 multilevel model                      Number of obs      =       2159
Binomial logit response model
Estimation algorithm: IGLS, PQL2
```

Level Variable	No. of Groups	Observations per Group Minimum	Average	Maximum
cluster	161	1	13.4	55
mom	1595	1	1.4	3

```
Run time (seconds)   =       8.71
Number of iterations =         29
```

immun	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
cons	-.1897211	.2437707	-0.78	0.436	-.6675029	.2880607
kid2p	1.363391	.1542231	8.84	0.000	1.061119	1.665663
rural	-.8506848	.2157692	-3.94	0.000	-1.273585	-.427785
pcInd81	-1.313231	.2638969	-4.98	0.000	-1.83046	-.7960028

Random-effects Parameters	Estimate	Std. Err.	[95% Conf.	Interval]
Level 3:				
var(cons)	.6723186	.1658299	.3472979	.9973393
Level 2:				
var(cons)	2.352684	.2589447	1.845162	2.860206

C:\Users\gl9158\Documents CAP NUM OVR

4. SIMULATION STUDIES ARE EASY

Do-file Editor - rodriguez and goldman (1995).do

File Edit Tools View

rodriguez and goldman (1995).do

```
set seed 12345
postfile MQL1 ix fx cx sigmaf sigmac using "MQL1.dta", replace
set obs 2
generate cx = _n - 1
expand 10
sort cx
generate cid = _n
expand 2
bysort cid: gen fx = _n - 1
expand 10
bysort cid (fx): generate fid = _n
expand 2
bysort cid fid: gen ix = _n - 1
expand 10
bysort cid fid (ix): gen iid = _n
generate cons = 1
forvalues iteration = 1/10 {
    display _n(5) as txt "Iteration " as res "`iteration'" as txt " of " as res "100"
    generate c = rnormal(0,1)
    bysort cid (fid iid): replace c = c[1]
    generate f = rnormal(0,1)
    bysort cid fid (iid): replace f = f[1]
    generate y = rbinomial(1,invlogit(0*cons + 1*ix + 1*fx + 1*cx + f + c))
    runmlwin y cons ix fx cx, level3(cid: cons) level2(fid: cons) level1(iid:) ///
        discrete(distribution(binomial) link(logit) denominator(cons)) ///
        nopause
    post MQL1 ([FP1]ix) ([FP1]fx) ([FP1]cx) (sqrt([RP2]var(cons))) (sqrt([RP3]var(cons)))
    drop c f y
}
postclose MQL1
use "MQL1.dta", clear
tabstat ix fx cx sigmaf sigmac, format(%3.2f)
```

Ready Line: 1, Col: 0 CAP NUM OVR

5. MCMC ESTIMATION

Refit the model using MCMC

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

$$logit(\pi_{ijk}) = \beta_0 + \beta_1 kid2p_{ijk} + \beta_2 rural_k + \beta_3 pcInd81_k + v_k + u_{jk}$$

$$v_k \sim \text{N}(0, \sigma_u^2)$$

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

```
. runmlwin immun cons kid2p rural pcInd81, ///
      level3(cluster: cons) ///
      level2(mom: cons) ///
      level1(kid:) ///
      discrete(d(binomial) l(logit) de(cons) pql2) ///
      mcmc(on) initsprevious nopause
```

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

File Edit Data Graphics Statistics User Window Help

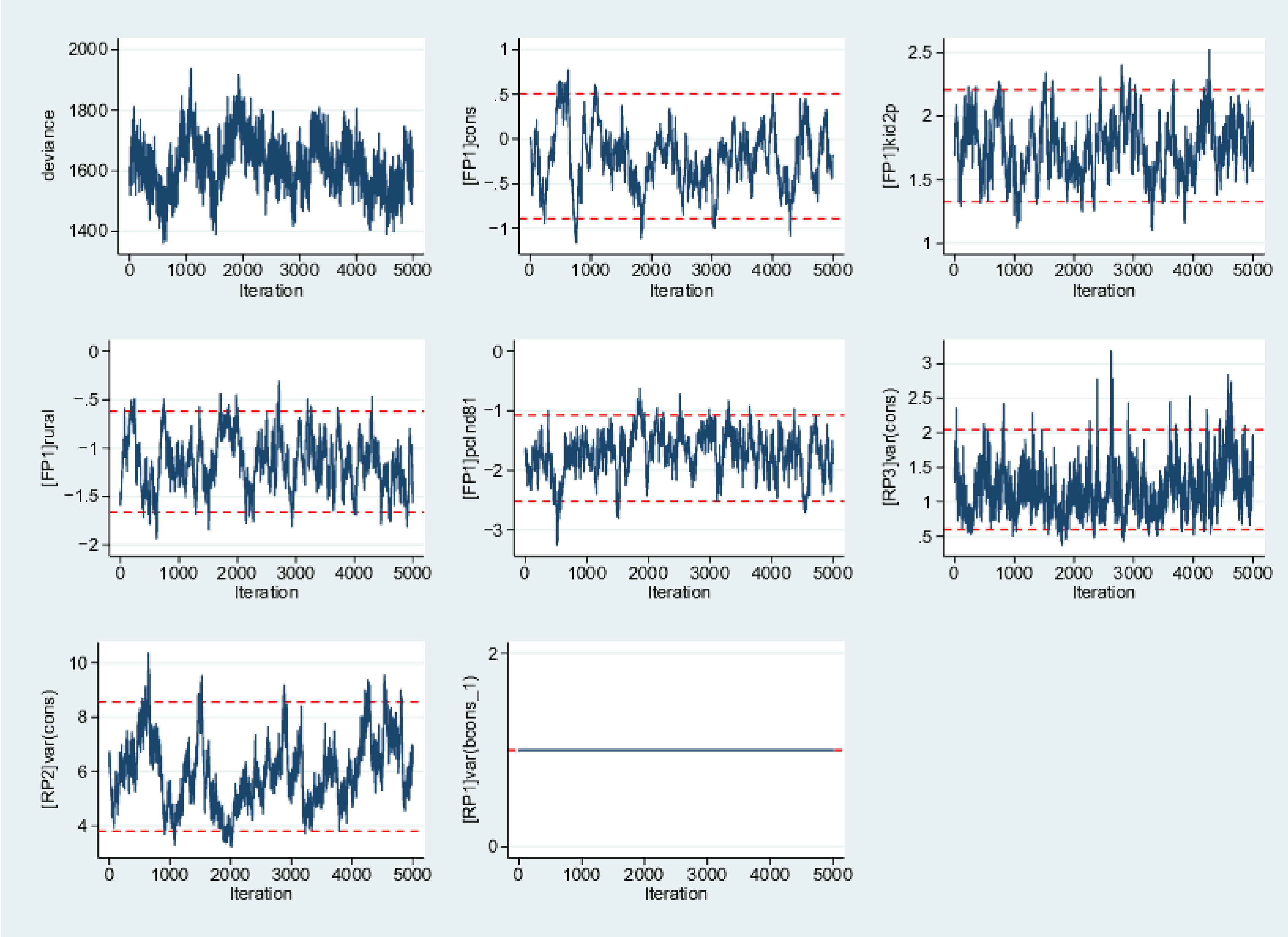
Level Variable	No. of Groups	Observations per Group Minimum	Average	Maximum
cluster	161	1	13.4	55
mom	1595	1	1.4	3

```
Burnin                         =        500
Chain                          =       5000
Run time (seconds)             =         30
Deviance (dbar)                =    1619.22
Deviance (thetabar)            =     866.88
Effective no. of pars (pd)     =     752.34
Bayesian DIC                   =    2371.56
```

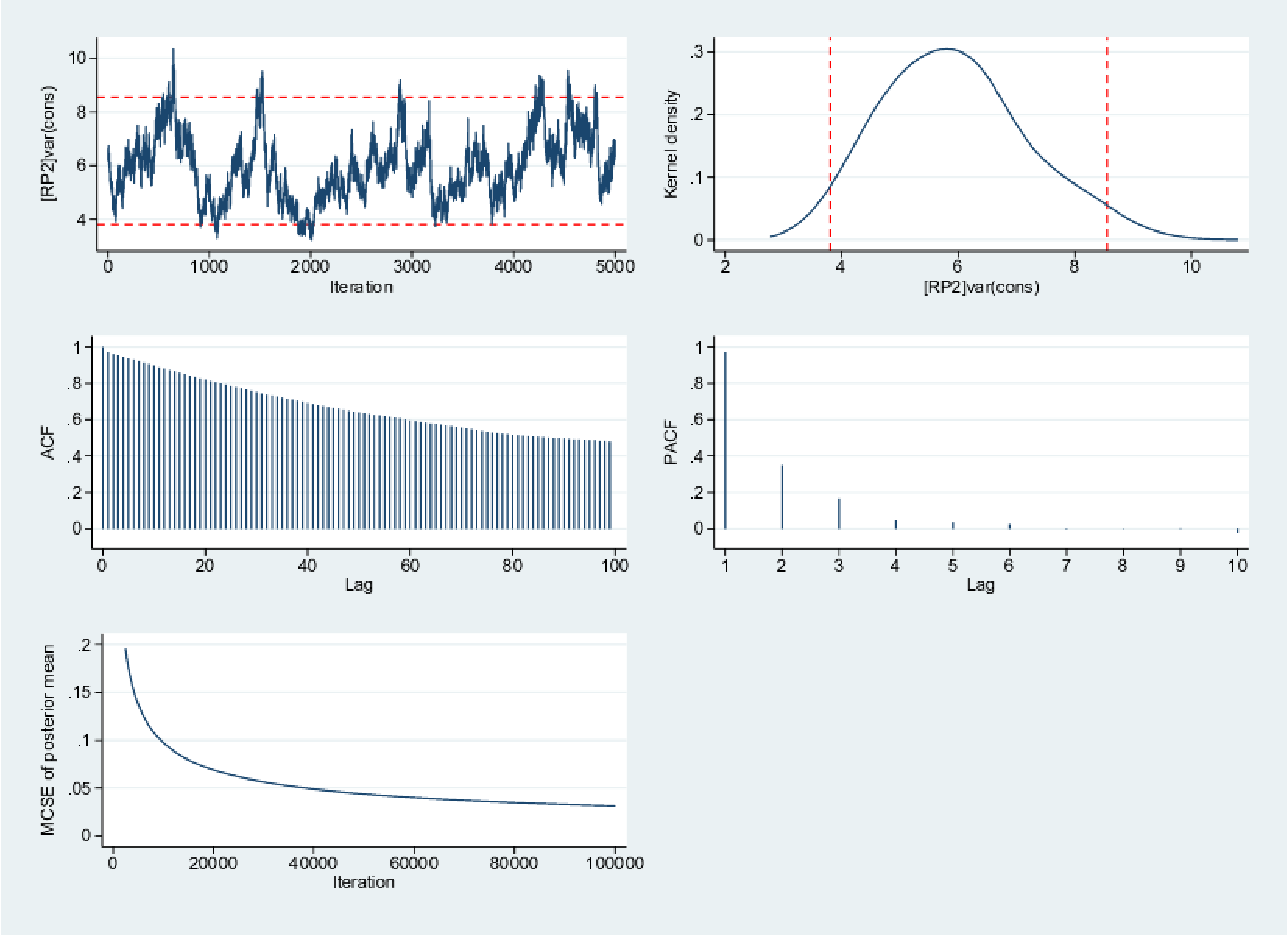
immun	Mean	Std. Dev.	z	ESS	[95% Cred.	Interval]
cons	-.2200957	.3358961	-0.66	42	-.889235	.5050995
kid2p	1.754326	.2300542	7.63	53	1.3257	2.208608
rural	-1.145384	.2759653	-4.15	56	-1.6628	-.6193233
pcInd81	-1.709476	.3672927	-4.65	61	-2.520294	-1.065675

Random-effects Parameters	Mean	Std. Dev.	ESS	[95% Cred.	Int]
Level 3:					
var(cons)	1.161717	.3641234	81	.6004681	2.046271
Level 2:					
var(cons)	5.934905	1.221375	21	3.805253	8.552722

C:\Users\gl9158\Documents CAP NUM OVR




```
. mcmcsum, trajectories
```




```
. mcmcsum [RP2]var(cons), 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              .0862781       0.5%    .0193246     Thinned Chain Length     5000
MCSE of Mean      .0024099       2.5%    .0243268     Effective Sample Size      99
Std. Dev.         .0467082         5%    .0298808     Raftery Lewis (2.5%)    25770
Mode              .0631075        25%    .0520173     Raftery Lewis (97.5%)   23976
                                                      Brooks Draper (mean)   446390
                                  50%    .0765091

                                  75%    .1100566
                                  95%     .179421
                                97.5%    .2023509
                                99.5%    .2549108
--------------------------------------------------------------------------------

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

C:\Users\gl9158\Documents CAP NUM OVR

Run the model for longer

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

$$logit(\pi_{ijk}) = \beta_0 + \beta_1 kid2p_{ijk} + \beta_2 rural_k + \beta_3 pcInd81_k + v_k + u_{jk}$$

$$v_k \sim \text{N}(0, \sigma_u^2)$$

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

```
. runmlwin immunized cons kid2p rural pcInd81, ///
      level3(cluster: cons) ///
      level2(mom: cons) ///
      level1(kid:) ///
      discrete(d(binomial) l(logit) de(cons) pql2) ///
      mcmc(burnin(5000) chain(50000) thinning(10)) ///
      initsprevious nopause
```

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

File Edit Data Graphics Statistics User Window Help

Level Variable	No. of Groups	Observations per Group Minimum	Average	Maximum
cluster	161	1	13.4	55
mom	1595	1	1.4	3

Burnin	=	5000
Chain	=	50000
Run time (seconds)	=	257
Deviance (dbar)	=	1641.35
Deviance (thetabar)	=	895.18
Effective no. of pars (pd)	=	746.17
Bayesian DIC	=	2387.51

immun	Mean	Std. Dev.	z	ESS	[95% Cred.	Interval]
cons	-.2421288	.3079326	-0.79	424	-.8276453	.3486197
kid2p	1.730911	.2184296	7.92	434	1.335295	2.19006
rural	-1.089141	.2954561	-3.69	471	-1.687306	-.5090104
pcInd81	-1.681882	.369106	-4.56	587	-2.450633	-.9707532

Random-effects Parameters	Mean	Std. Dev.	ESS	[95% Cred.	Int]
Level 3:					
var(cons)	1.127521	.3580364	572	.5575184	1.94313
Level 2:					
var(cons)	5.628252	1.25813	186	3.587712	8.415654

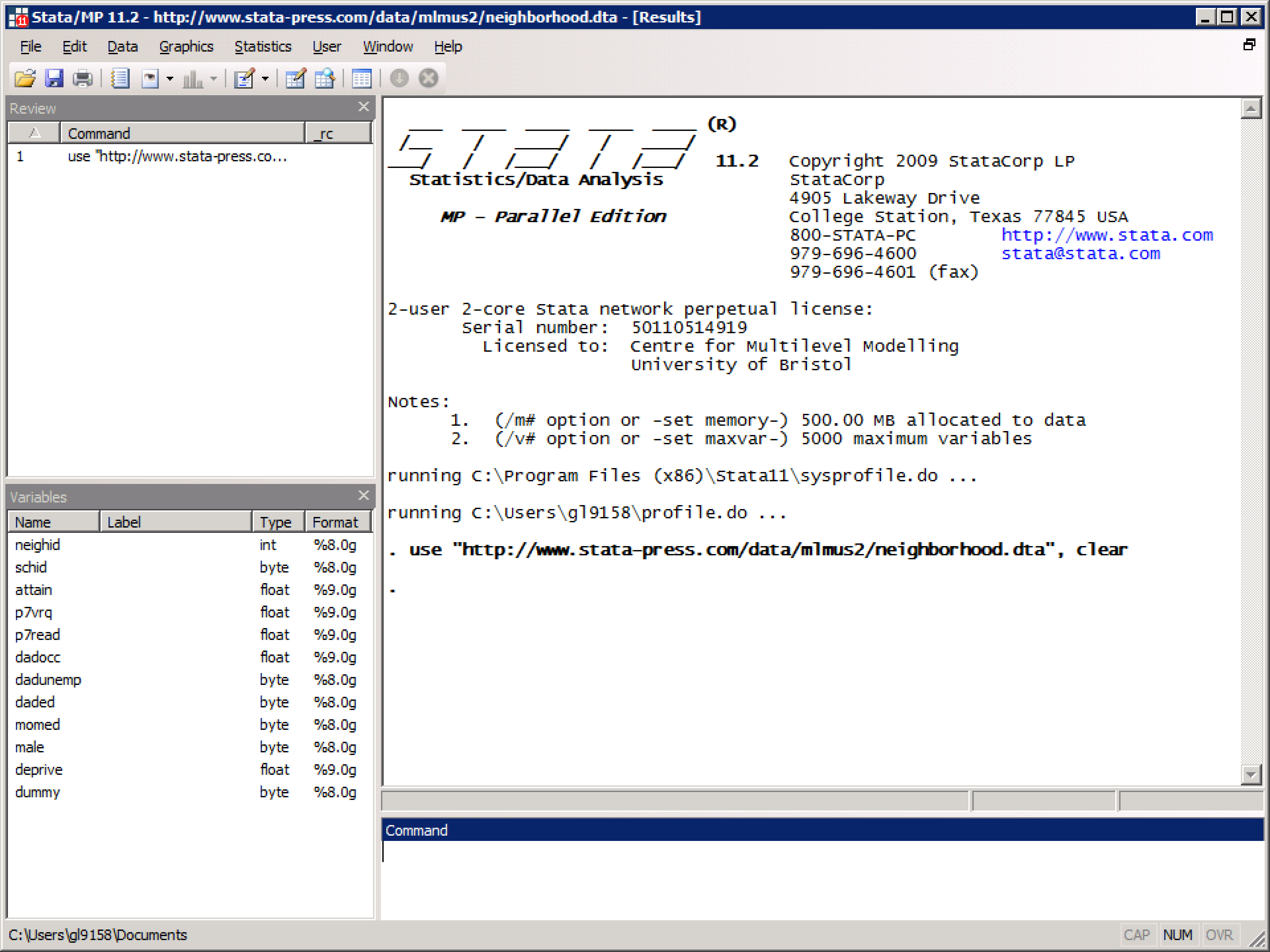
Q:\C-modelling\runmlwin\presentations\2011-05-26 Connecticut

CAP NUM OVR

6. CROSS-CLASSIFIED MODELS

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



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...
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
(R)
STATA
11.2 Copyright 2009 StataCorp LP
StataCorp
Statistics/Data Analysis
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
.
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

```
. 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

Two-way cross-classified model

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

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

```
. matrix b = (0,.33,.33,.33)

. 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,.33,.33,.33)

. runmlwin attain cons, level3(schid: cons) level2(neighid: cons) level1(student
> id: 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)         =       6.88
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) |  .0995634   .0492998    2175    .0387661   .2261923
----------------------------+-------------------------------------------------
Level 2:                    |
                  var(cons) |  .1422148   .0217524     458    .1033897   .1892418
----------------------------+-------------------------------------------------
Level 1:                    |
                  var(cons) |  .8002955   .0260423    2583    .7513126   .8540478
------------------------------------------------------------------------------
```

C:\Users\gl9158\Documents   CAP NUM OVR

Two-way cross-classified model

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

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

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

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

```
. 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 female 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)         =       26.6
Deviance (dbar)            =    4744.77
Deviance (thetabar)        =    4704.11
Effective no. of pars (pd) =      40.67
Bayesian DIC               =    4785.44
------------------------------------------------------------------------------
      attain |      Mean   Std. Dev.       z      ESS     [95% Cred. Interval]
-------------+----------------------------------------------------------------
        cons |  .0351255   .0291963     1.20     1706    -.0220368    .0938575
       p7vrq |  .0275779   .0022758    12.12     4556     .0231709      .03204
      p7read |  .0262253   .0017897    14.65     4689     .0226729    .0297626
      dadocc |  .0080741   .0013761     5.87     4680     .0053839    .0107416
       daded |   .142757   .0411453     3.47     5452     .0615814    .2230719
       momed |  .0605109   .0379741     1.59     4703    -.0132536    .1342922
    dadunemp | -.1224487   .0468065    -2.62     4505    -.2130983    -.028468
      female |  .0558048   .0280615     1.99     4699     .0015048    .1117451
     deprive | -.1562503   .0260965    -5.99     3705     -.207776   -.1055711
------------------------------------------------------------------------------

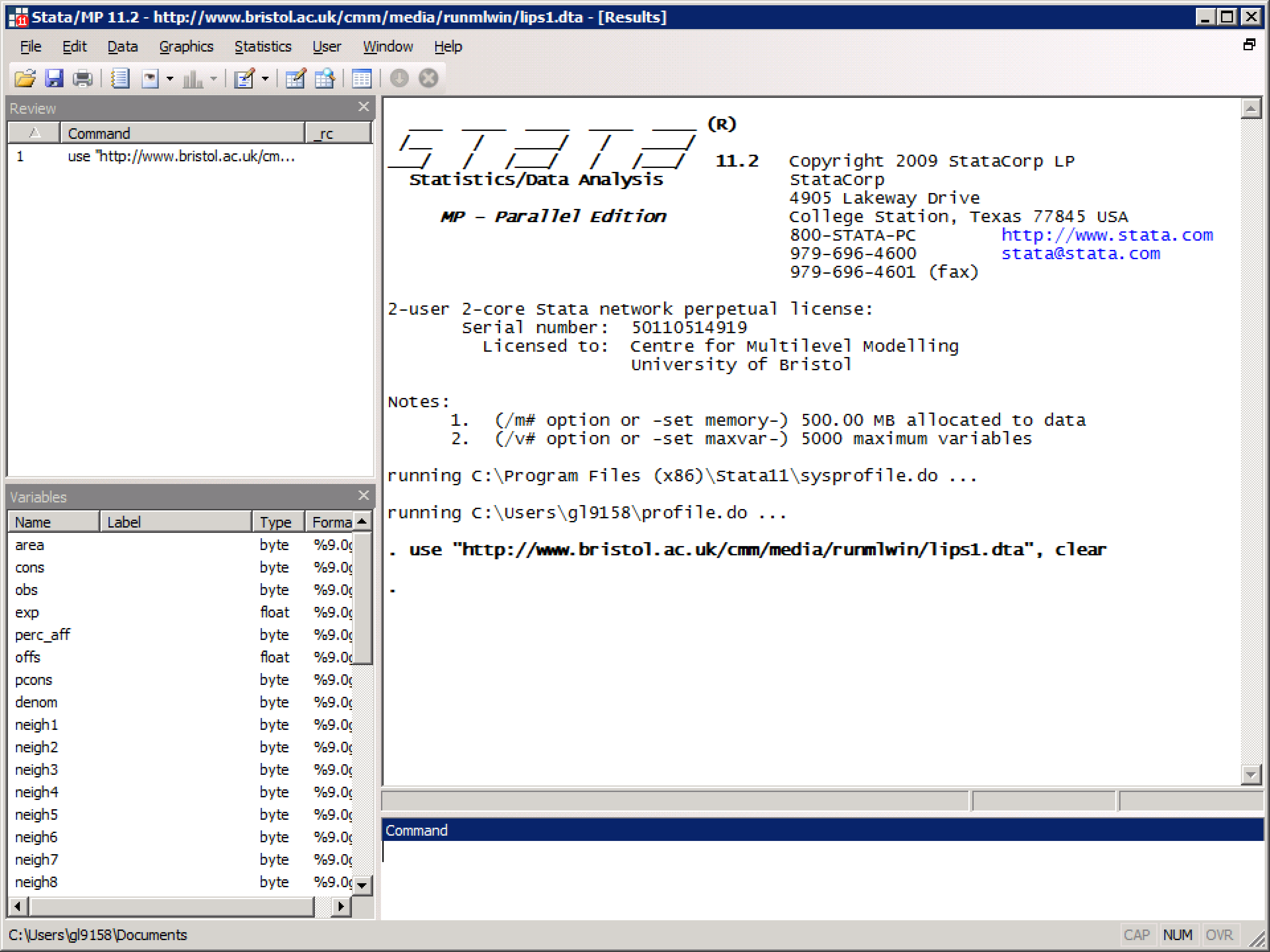
------------------------------------------------------------------------------
```

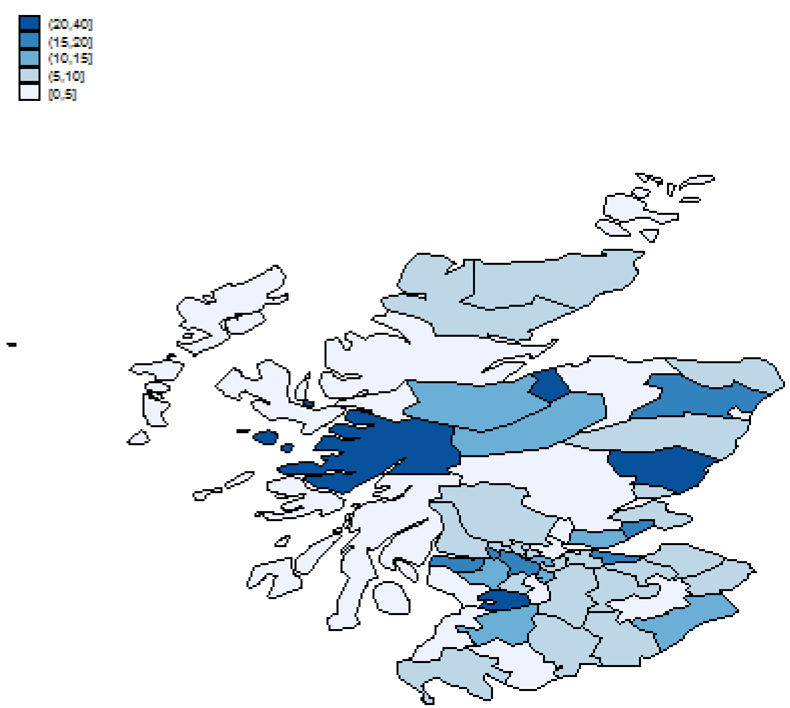
C:\Users\gl9158\Documents   CAP NUM OVR

7. SPATIAL MULTILEVEL MODELS

Scottish lip cancer

- County level lip cancer counts between 1975 and 1980
- 56 Scottish counties
- First analysed by Clayton and Kaldor (1987)
- Re-analysed by Breslow and Clayton (1993), Leyland and Goldstein (2001), Rabe-Hesketh and Skrondal (2008) and others


Stata/MP 11.2 - http://www.bristol.ac.uk/cmm/media/runmlwin/lips1.dta - [Results]
File Edit Data Graphics Statistics User Window Help
Review
Command _rc
1 use "http://www.bristol.ac.uk/cm...
(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.bristol.ac.uk/cmm/media/runmlwin/lips1.dta", clear
.
Variables
Name Label Type Forma
area byte %9.0g
cons byte %9.0g
obs byte %9.0g
exp float %9.0g
perc_aff byte %9.0g
offs float %9.0g
pcons byte %9.0g
denom byte %9.0g
neigh1 byte %9.0g
neigh2 byte %9.0g
neigh3 byte %9.0g
neigh4 byte %9.0g
neigh5 byte %9.0g
neigh6 byte %9.0g
neigh7 byte %9.0g
neigh8 byte %9.0g
Command
C:\Users\gl9158\Documents
CAP NUM OVR



- use "http://www.bristol.ac.uk/cmm/media/runmlwin/lips1.dta", clear

- merge 1:1 area using "scotdb.dta"

- spmap obs using "scotcoord.dta", id(area) ///
      fcolor(Blues) legend(position(10)) ///
      clmethod(custom) clbreaks(0 5 10 15 20 40)

Over-dispersed Poisson model

$$obs_i \sim \text{Poisson}(\pi_i)$$

$$log(\pi_i) = offs_i + \beta_0 + \beta_1 perc\_aff_i + u_i$$

$$u_i \sim \text{N}(0, \sigma_u^2)$$

```
runmlwin obs cons perc_aff, ///
      level2(area: cons) ///
      level1(area:) ///
      discrete(dist(poisson) link(log) offset(offs)) ///
      mcmc(chain(50000)) ///
      initsprevious nopause
```

Stata/MP 11.2 - http://www.bristol.ac.uk/cmm/media/runmlwin/lips1.dta - [Results]

File Edit Data Graphics Statistics User Window Help

```
. runmlwin obs cons perc_aff, level2(area:cons) level1(area:) discrete(distribut
> ion(poisson) link(log) offset(offs)) mcmc(chain(50000) refresh(500)) initsprev
> ious nopause

MLwiN 2.23 multilevel model                    Number of obs      =         56
Poisson response model
Estimation algorithm: MCMC, MQL1
```

Level Variable	No. of Groups	Observations per Group Minimum	Average	Maximum
area	56	1	1.0	1

```
Burnin                     =         500
Chain                      =       50000
Run time (seconds)         =        10.3
Deviance (dbar)            =      270.38
Deviance (thetabar)        =      230.65
Effective no. of pars (pd) =       39.73
Bayesian DIC               =      310.11
```

obs	Mean	Std. Dev.	z	ESS	[95% Cred.	Interval]
cons	-.4835851	.1628447	-2.97	494	-.8143693	-.1711187
perc_aff	.0675449	.0143747	4.70	475	.0388835	.0956604

Random-effects Parameters	Mean	Std. Dev.	ESS	[95% Cred. Int]	
Level 2:					
var(cons)	.3852672	.1125743	6787	.2107703	.645173

C:\Users\gl9158\Documents CAP NUM OVR

CAR model

$$obs_i \sim \text{Poisson}(\pi_i)$$

$$\log(\pi_i) = offs_i + \beta_0 + \beta_1 perc\_aff_i + u_i$$

$$u_i \sim \text{N}\left(\bar{u}_i, \frac{\sigma_u^2}{r_i}\right), \qquad \bar{u}_i = \sum_{j \in neigh(i)} \frac{w_{i,j} u_i}{r_i}$$

```
. runmlwin obs perc_aff, ///
      level2(area: cons, carids(neigh1-neigh11) ///
      carweights(wcar1-wcar11)) ///
      level1(cons:) ///
      discrete(dist(poisson) link(log) offset(offs)) ///
```

Stata/MP 11.2 - http://www.bristol.ac.uk/cmm/media/runmlwin/lips1.dta - [Results]

File Edit Data Graphics Statistics User Window Help

```
. runmlwin obs perc_aff, level2(area: cons, carids(neigh1-neigh11) carweights(wcar1-wcar11))
>  level1(cons:) discrete(distribution(poisson) link(log) offset(offs)) mcmc(chain(50000) re
> fresh(500)) initsprevious nopause

MLwiN 2.23 multilevel model                     Number of obs      =        56
Poisson response model
Estimation algorithm: MCMC, MQL1



-----------------------------------------------------------
                |   No. of       Observations per Group
 Level Variable |   Groups    Minimum    Average    Maximum
----------------+------------------------------------------
           area |       56          1        1.0          1
-----------------------------------------------------------


Burnin                     =        500
Chain                      =      50000
Run time (seconds)         =       9.67
Deviance (dbar)            =     268.77
Deviance (thetabar)        =     240.42
Effective no. of pars (pd) =      28.35
Bayesian DIC               =     297.13
------------------------------------------------------------------------------
         obs |      Mean    Std. Dev.       z     ESS     [95% Cred. Interval]
-------------+----------------------------------------------------------------
    perc_aff |   .035667    .0128288     2.78     354     .0090298    .0591634
------------------------------------------------------------------------------


------------------------------------------------------------------------------
   Random-effects Parameters |    Mean    Std. Dev.    ESS     [95% Cred. Int]
-----------------------------+------------------------------------------------
Level 2:                     |
                   var(cons) |  .5337886   .1900985    3602    .2512767   .9866003
```

C:\Users\gl9158\Documents   CAP NUM OVR

Convolution model

$$obs_i \sim \text{Poisson}(\pi_i)$$

$$\log(\pi_i) = offs_i + \beta_0 + \beta_1 perc_{aff_i} + v_i + u_i$$

$$v_i \sim \text{N}\left(\bar{v}_i, \frac{\sigma_v^2}{r_i}\right), \qquad \bar{v}_i = \sum_{j \in neigh(i)} \frac{w_{i,j} v_i}{r_i}$$

$$u_i \sim \text{N}(0, \sigma_u^2)$$

```
. runmlwin observed cons perc_aff, ///
      level3(area: cons, carids(neigh1-neigh11) ///
      carweights(wcar1-wcar11)) ///
      level2(area: cons) level1(county:) ///
      discrete(d(binomial) l(log) offset(offs)) ////
```

Stata/MP 11.2 - http://www.bristol.ac.uk/cmm/media/runmlwin/lips1.dta - [Results]

File Edit Data Graphics Statistics User Window Help

```
MLwiN 2.23 multilevel model                      Number of obs      =        56
Poisson response model
Estimation algorithm: MCMC, MQL1
```

Level Variable	No. of Groups	Observations per Group Minimum	Average	Maximum
area	56	1	1.0	1

```
Burnin                          =        500
Chain                           =      50000
Run time (seconds)              =       14.6
Deviance (dbar)                 =     267.86
Deviance (thetabar)             =     238.12
Effective no. of pars (pd)      =      29.74
Bayesian DIC                    =     297.60
```

obs	Mean	Std. Dev.	z	ESS	[95% Cred.	Interval]
cons	.8812095	1.762997	0.50	26	-2.068313	3.722563
perc_aff	.0365398	.0136445	2.68	472	.0084325	.0622706

Random-effects Parameters	Mean	Std. Dev.	ESS	[95% Cred.	Int]
Level 3:					
var(cons)	.4903258	.1894587	2166	.2036356	.9411708
Level 2:					
var(cons)	.0236907	.0324559	208	.0006044	.1154397

C:\Users\gl9158\Documents CAP NUM OVR

8. EXPORT MODELS TO WINBUGS

The `runmlwin` command syntax

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

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

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

```
. runmlwin normexam cons standlrt, ///
      level2(school: cons) ///
      level1(student: cons) ///
      mcmc(savewinbugs(model(m.txt) inits(i.txt) ///
      data(d.txt) nofit)) ///
      initsprevious nopause
```

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) {
normexam[i] ~ dnorm(mu[i],tau)
mu[i]<- beta[1] * cons[i]
+ beta[2] * standlrt[i]
+ u2[school[i]] * cons[i]
}
# Higher level definitions
for (j in 1:n2) {
u2[j] ~ dnorm(0,tau.u2)
}
# Priors for fixed effects
for (k in 1:2) { beta[k] ~ dflat() }
# Priors for random terms
tau ~ dgamma(0.001000,0.001000)
sigma2 <- 1/tau
tau.u2 ~ dgamma(0.001000,0.001000)
sigma2.u2 <- 1/tau.u2
}
```

t-distributed level 2 residuals

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

$$u_j \sim t(0, \sigma_u^2, \nu)$$

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

Viewer (#1) [view m.txt]

view m.txt

Advice | Contents | What's New | News

```
# WINBUGS 1.4 code generated from MLwiN program

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

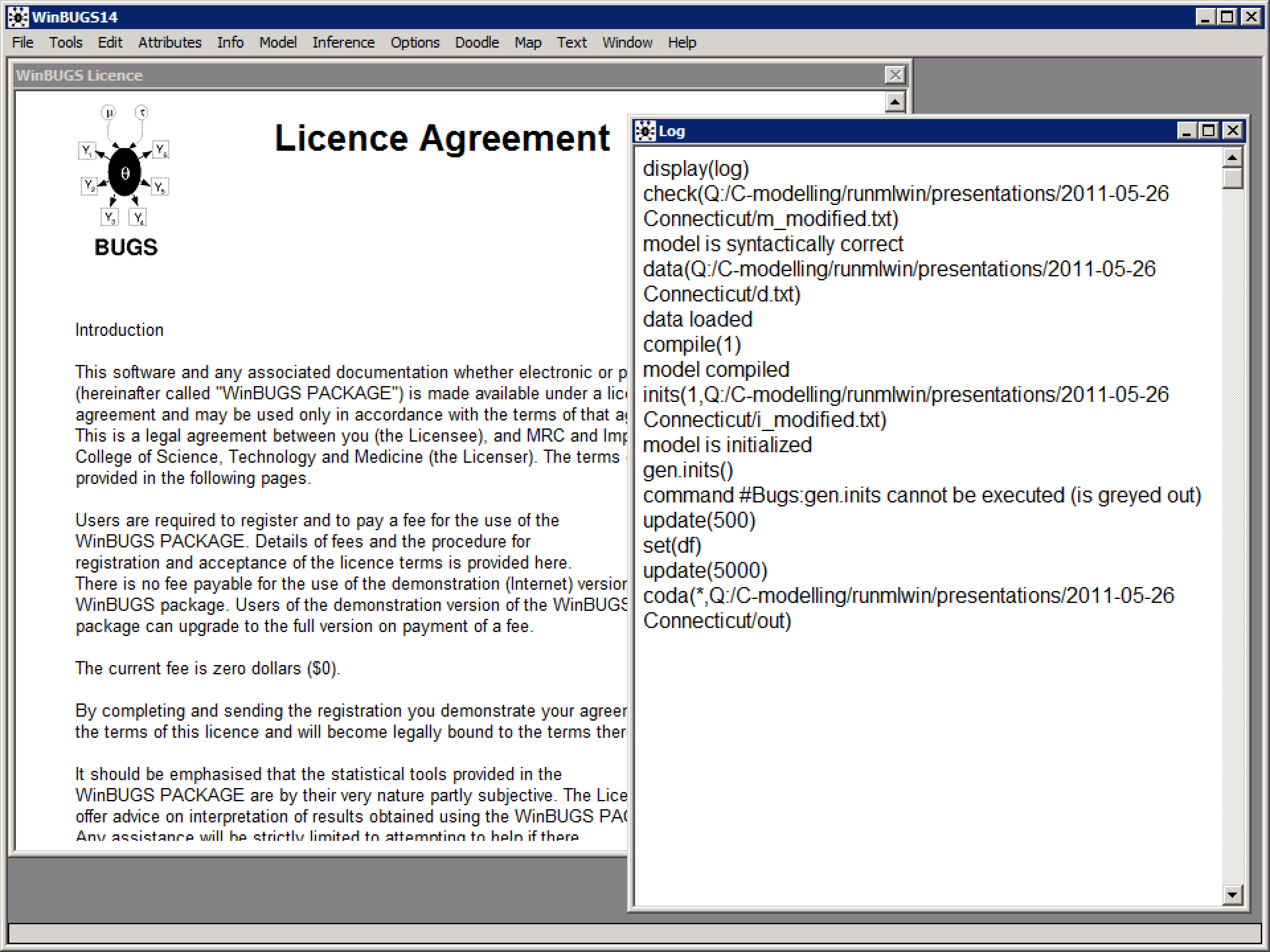
model
{
# Level 1 definition
for(i in 1:N) {
normexam[i] ~ dnorm(mu[i],tau)
mu[i]<- beta[1] * cons[i]
+ beta[2] * standlrt[i]
+ u2[school[i]] * cons[i]
}
# Higher level definitions
for (j in 1:n2) {
u2[j] ~ dt(0,tau.u2,df)
}
# Priors for fixed effects
for (k in 1:2) { beta[k] ~ dflat() }
# Priors for random terms
tau ~ dgamma(0.001000,0.001000)
sigma2 <- 1/tau
tau.u2 ~ dgamma(0.001000,0.001000)
sigma2.u2 <- 1/tau.u2
df ~ dunif(2,200)
}
```

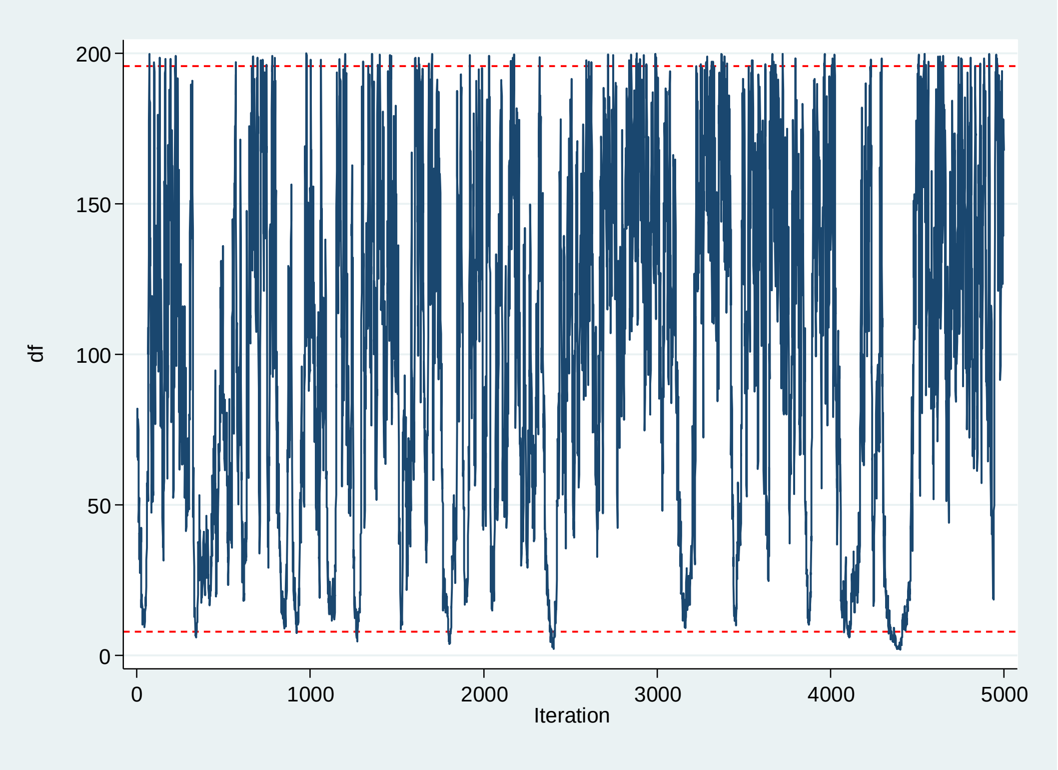
The `winbugs` suite of commands

```
. wbscript , ///
      model("`c(pwd)'\m.txt") inits("`c(pwd)'\i.txt") ///
      data("`c(pwd)'\d.txt") coda("`c(pwd)'\out") ///
      set(df) burn(500) update(5000) ///
      saving("`c(pwd)'\script.txt", replace) quit

. wbrun, script("`c(pwd)'\script.txt") ///
      winbugs("C:\Users\gl9158\WinBUGS14\winbugs14.exe")

. wbcoda, root("`c(pwd)'\out") clear
```


WinBUGS14
File Tools Edit Attributes Info Model Inference Options Doodle Map Text Window Help
WinBUGS Licence
BUGS
Licence Agreement
Introduction
This software and any associated documentation whether electronic or p
(hereinafter called "WinBUGS PACKAGE") is made available under a lic
agreement and may be used only in accordance with the terms of that a
This is a legal agreement between you (the Licensee), and MRC and Imp
College of Science, Technology and Medicine (the Licenser). The terms
provided in the following pages.
Users are required to register and to pay a fee for the use of the
WinBUGS PACKAGE. Details of fees and the procedure for
registration and acceptance of the licence terms is provided here.
There is no fee payable for the use of the demonstration (Internet) versior
WinBUGS package. Users of the demonstration version of the WinBUGS
package can upgrade to the full version on payment of a fee.
The current fee is zero dollars ($0).
By completing and sending the registration you demonstrate your agreer
the terms of this licence and will become legally bound to the terms ther
It should be emphasised that the statistical tools provided in the
WinBUGS PACKAGE are by their very nature partly subjective. The Lice
offer advice on interpretation of results obtained using the WinBUGS PA
Log
display(log)
check(Q:/C-modelling/runmlwin/presentations/2011-05-26
Connecticut/m_modified.txt)
model is syntactically correct
data(Q:/C-modelling/runmlwin/presentations/2011-05-26
Connecticut/d.txt)
data loaded
compile(1)
model compiled
inits(1,Q:/C-modelling/runmlwin/presentations/2011-05-26
Connecticut/i_modified.txt)
model is initialized
gen.inits()
command #Bugs:gen.inits cannot be executed (is greyed out)
update(500)
set(df)
update(5000)
coda(*,Q:/C-modelling/runmlwin/presentations/2011-05-26
Connecticut/out)



. mcmcsum df, trajectories variables

Stata/MP 11.2 - [Results]

File Edit Data Graphics Statistics User Window Help

```
.
. mcmcsum df, variables

                                         df
-------------------------------------------------------------------------------------
                                    Percentiles
Mean              100.3878      0.5%   3.102855     Thinned Chain Length        5000
MCSE of Mean      4.172605      2.5%   7.934025     Effective Sample Size        107
Std. Dev.         57.77448        5%     11.977     Raftery Lewis (2.5%)       57207
Mode              21.98396       25%    49.3425     Raftery Lewis (97.5%)       5464
                                                    Brooks Draper (mean)       13383
                                 50%     102.25

                                 75%      149.4
                                 95%      191.2
                               97.5%      195.8
                               99.5%      199.3
-------------------------------------------------------------------------------------

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

Q:\C-modelling\runmlwin\presentations\2011-05-26 Connecticut   CAP NUM OVR

9. WORK EFFICIENTLY

Do-file Editor - Amsterdam2011.do

File Edit Tools View

Amsterdam2011.do

```
* Open the tutorial data set
use "http://www.bristol.ac.uk/cmm/media/runmlwin/tutorial.dta", clear


* Fit a two-level (students within schools) variance components model to
* a continuous educational response variable, normexam. Note, you will need
* to click the "Resume Macro" button twice in MLwiN to return the model
* results to the Stata output window.
runmlwin normexam cons, ///
    level2(school: cons) ///
    level1(student: cons)

* Generate a boy dummy variable
generate boy = 1 - girl

* Extend the previous model to include fixed part covariates, a random school
* level slope and separate level 1 residuals for boys and girls. The runmlwin
* command also requests that runmlwin extracts the predicted values for the
* school level residuals from MLwiN and returns them to Stata. The nopause
* option prevents MLwiN from pausing before and after model estimation and so
* returns the model results automatically to Stata.
runmlwin normexam cons standlrt girl, ///
    level2(school: cons standlrt, residuals(u)) ///
    level1(student: girl boy, diagonal) nopause

* Perform a Wald test to compare the boy and girl residual variances
test [RP1]var(girl) = [RP1]var(boy)

* Preserve the data as we will shortly be collapsing the data to the school
* level, but afterwards we will want to return to the original data
preserve

    * Tag one child in each school
    egen pickone = tag(school)

    * Collapse the data to one row per school
```

Ready   Line: 4, Col: 50   CAP   NUM   OVR

10. RESOURCES TO HELP YOU LEARN RUNMLWIN

Viewer (#1) [help runmlwin]

help runmlwin

Advice | Contents | What's New | News

---

help for **runmlwin**

Title

runmlwin - Running the MLwiN multilevel modelling package from within Stata

Syntax

runmlwin *responses_and_fixed_part*, *random_part* [discrete(*discrete_options*)] [*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 multinomial ordered and unordered 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** the mlwin.exe file address. See **Remarks on installation instructions** above.

(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.
(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
(See page 59 of the MLwiN User Manual)

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

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

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

Refit the model surpressing the two pauses in MLwiN
(See page 59 of the MLwiN User Manual)

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

Two-level random-intercept and random-slope (coefficient) model with a complex level 1 variance function
(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))
```

Multivariate response models

Setup

```
. use http://www.bristol.ac.uk/cmm/media/runmlwin/gcsemv1, clear
```

Random-intercept bivariate response model
(See page 214 of the MLwiN User Manual)

```
. runmlwin (written cons female, eq(1)) (csework cons female, eq(2)), level2(school: (cons, eq(1))
    (cons, eq(2))) level1(student:   (cons, eq(1)) (cons, eq(2)))
```

Cross-classified models

runmlwin: Running MLwiN from within Stata

runmlwin is a user written Stata command to fit multilevel models in MLwiN from within Stata. Models can be fit to both hierarchical and non-hierarchical (cross-classified and multiple membership) data structures and to both univariate and multivariate responses. Models can be fit to continuous, categorical (binary, proportion, nominal, ordinal) and count data.

The multilevel models fitted by **runmlwin** are analogous to those fitted by the Stata's **xtmixed, xtmelogit** and **xtmepoisson** commands and by the user written **gllamm** command.

running the **runmlwin** command in Stata carries out the following steps:

1. Writes an MLwiN macro for the specified multilevel model.
2. Opens MLwiN and runs the MLwiN macro.
3. Pauses MLwiN once the model is specified. This allows the user to check that the model is specified as expected.
4. Fits the model in MLwiN.
5. Pauses MLwiN once the model has been fitted (i.e. converged). This allows the user to examine the model results.
6. Stores and displays the model results in Stata

MLwiN and Stata are both required to use **runmlwin.**

Download

runmlwin is now available as a beta release. Users should be able to fit all models in the MLwiN User Manual. We are currently developing **runmlwin** so that users will soon be able to additionally fit all models in the MCMC MLwiN Manual.

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

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

Presentations using runmlwin

- University of Bristol, Mplus/MlwiN User Group (MUGS) meeting (14th June 2011)
  - Slides (PDF, 1.0mb)
  - 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, 3.0mb)
  - Stata do-file (do, 0.3mb) to replicate all analyses presented in the slides.
- 2011 American Sociological Association Spring Methodology Conference, Tilburg University (20th May 2011)
  - Slides (PDF, 1.0mb)
  - 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)
  - Slides (PDF, 2.0mb)
  - Stata do-file (do, 0.1mb) to replicate all analyses presented in the slides.

Examples using runmlwin

MLwiN User Manual

These do-files and log files replicate the analyses reported in the MLwiN User Manual.

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)
- 15 - Diagnostics for Multilevel Models (do | log)

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

File Edit View History Bookmarks Tools Help

www.cmm.bristol.ac.uk/forum • View forum ...

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

Wiktionary (en)

runmlwin user forum

Forum rules

NEWTOPIC Search this forum... Search

12 topics • Page 1 of 1

ANNOUNCEMENTS	REPLIES	VIEWS	LAST POST
Replicate the entire MLwiN User Manual using runmlwin by GeorgeLeckie » Mon Apr 18, 2011 5:30 pm	0	68	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	87	by GeorgeLeckie Fri Apr 01, 2011 4:06 pm

TOPICS	REPLIES	VIEWS	LAST POST
MCMC estimation by janna » Fri Apr 08, 2011 9:29 am	5	116	by ChrisCharlton Thu Jun 09, 2011 10:09 am
error: too few quotes by laura » Wed May 18, 2011 5:54 pm	4	73	by laura Thu May 19, 2011 1:01 pm
Fixed parameters as Odds ratio or Relative risk by Beatriz » Fri Apr 08, 2011 5:29 pm	4	145	by GeorgeLeckie Thu May 12, 2011 3:51 pm
Can you use Stata's predict command after runmlwin? by jbherman » Wed May 11, 2011 7:53 pm	1	59	by GeorgeLeckie Thu May 12, 2011 11:12 am
Error message: too many macros by Corrie » Wed Apr 27, 2011 3:58 pm	2	107	by ChrisCharlton Wed Apr 27, 2011 11:06 pm
Error message: No version information foundr(198) by Aknigge » Wed Apr 20, 2011 2:29 pm	3	86	by GeorgeLeckie Fri Apr 22, 2011 10:56 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