Introduction to Multilevel Modeling using MLwin version 2.02:

Fitting a two-level model[1]

1. Introduction

The data set you will use has been selected from a much larger data set produced by the Junior School Project (JSP) (Mortimer et al, 1989). This was a longitudinal study of an age cohort of 2000 pupils (level 1) who entered junior school age 7 and left in 1984 aged 11. The JSP pupils attended 48 primary schools (level 2) selected randomly from the 636 primary schools maintained by the Inner London Education Authority.

In this data modelling task the pupil outcome measure (called the 'response' variable) is a score for mathematics at age 9 (MATHS5; that is year 5 of the National Curriculum) and the single predictor (explanatory) variable is the mathematics score at age 7 (MATHS3; that is year 3 of the National Curriculum). We are interested in discovering – via data modelling - the size, nature and extent of the school effect on progress in mathematics.

We will consider the following models:

Model 1. A random intercepts 'null' model with Maths5 as the response. No predictor/explanatory variables apart from the Constant (ie representing the intercept) which is allowed to vary randomly across schools and with the levels defined as pupils (level 1) in schools (level 2);

Model 2. A random intercepts model. Model 1 with also the pre-test score as an explanatory variable (Maths3);

Model 3. A random intercepts/slopes model. Model 2 with also the parameter associated with Maths3 being allowed to vary randomly across schools; that is random slopes as well as intercepts.

For any multilevel model there is a basic sequence of procedures which we will follow:

- data input: sorting, creating the constant term (normally called 'cons');
- model specification: response, predictors (explanatory variables), level, terms for the fixed and random part;
- estimation: the fitting of the specified model by a chosen procedure;
- examining the estimates and values such as standard errors
- estimating the residuals at each level for diagnosis of model problems and sometimes to make substantive interpretations;
- graphing the results both to look at estimate residuals and predictions from the estimated model
- model re-specification and the cycle begins over again.

2. Data input and manipulation

Here is a recommended sequence:

Data input

File on Main Menu

ASCII text file input

[1] Updated and adapted from task created by Kelvyn Jones, School of Geographical Sciences, University of Bristol.

Columns: c1-c5
File: jsp.dat
OK

Name columns

Data manipulation on main menu
Names

C1	School	Enter
C2	Pupil	Enter
C3	Math3	Enter
C4	Math5	Enter
C5	Male	Enter

Naming categories

Highlight 'male'
Categories
0: Female
1: Male

Save the worksheet

File on Main Meun
Save worksheet as: jsp.ws

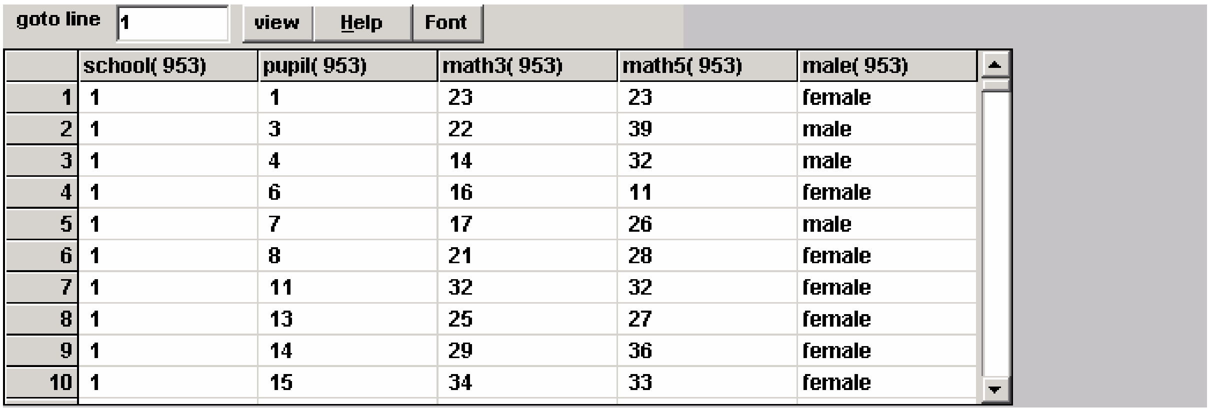
Remember to write down the complete filename you have used.

Sorting the data: pupils within schools

Data manipulation on main menu
Sort
Increase number of keys to 2
Choose 'School' as the highest key [slowest changing]
Choose 'Pupil' as the lowest key [fastest changing]
Highlight 'School' to 'male'
Same as input
Add to Action List
Execute
Close Sort Window

Check data and save sorted worksheet

Data Manipulation on Main Menu
View or Edit data
Select View and Highlight 'School' to 'male' to select columns to view
OK
Resize window to see all 5 columns


goto line 1 | view | Help | Font

	school(953)	pupil(953)	math3(953)	math5(953)	male(953)
1	1	1	23	23	female
2	1	3	22	39	male
3	1	4	14	32	male
4	1	6	16	11	female
5	1	7	17	26	male
6	1	8	21	28	female
7	1	11	32	32	female
8	1	13	25	27	female
9	1	14	29	36	female
10	1	15	34	33	female

Some questions: 1

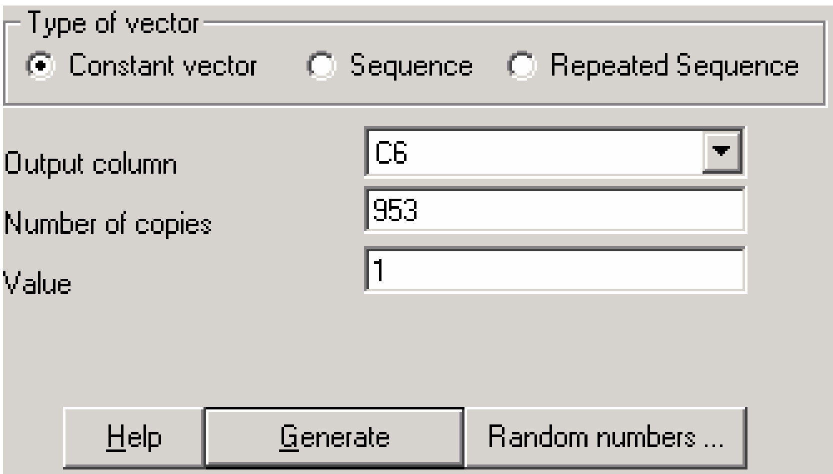
Has pupil 1 in school 1 made progress; what about pupil 4, and 15?

If it looks correct -
File on Main Menu
Save (as jsp.ws)
Yes to overwrite

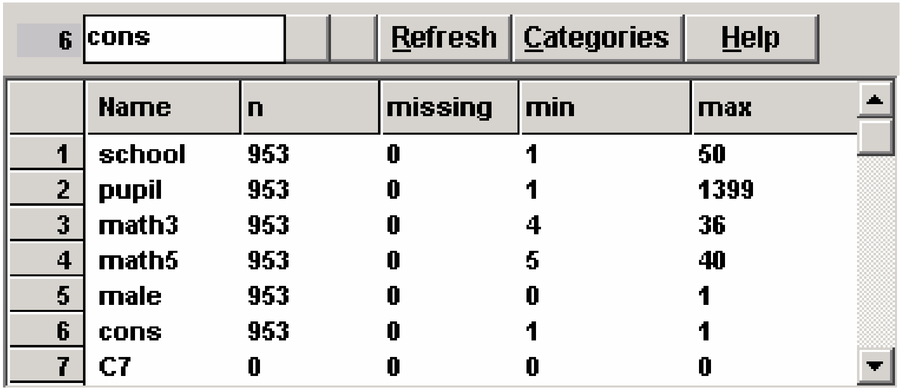
There is a final variable we have to create before we can begin modelling the data – the 'constant' variable. The 'constant' variable takes the value of 1 for every pupil (ie is a vector of 1's) and is used to estimate the intercept term in the regression equation. There are many ways of doing this but you must ensure that there is 1 for each and every pupil. The simplest way to achieve this is:

Data manipulation on the Main Menu
Generate Vector
Constant Vector
Output column: c6
Number of copies: 953
Value: 1
Generate
Close window

The Generate vector just before Generate is clicked should look like:



Name c6 as ‘cons’. The revised top of the worksheet should look like:



6 cons | Refresh | Categories | Help

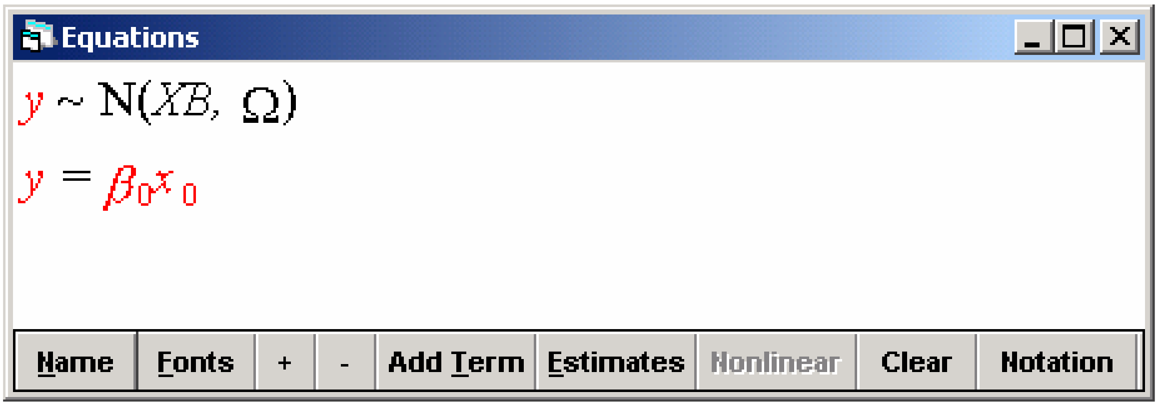
	Name	n	missing	min	max
1	school	953	0	1	50
2	pupil	953	0	1	1399
3	math3	953	0	4	36
4	math5	953	0	5	40
5	male	953	0	0	1
6	cons	953	0	1	1
7	C7	0	0	0	0

After saving the revised data you are ready for modelling. Close the Names and the View data windows, you will not need them for the time being.

3. Model 1: two-level ‘null’ random intercepts

Specifying the model

The most straightforward way to specify the model is through the equations window which you will find under model on the main screen. Clicking on equations (main menu – model) will bring up the following rather uninspiring screen which is the heart of the programme. Here models are specified and estimates displayed. It is also possible to specify models in the command window and to see the equations displayed here.



Ignoring the bottom tool bar for a moment; the equations are as follows:

y	is the response;
N	indicates a normal distribution for a fixed part XB and a random part Ω;
β_0	is the first fixed part estimate to be specified and x_0 is the first predictor variable to be specified.
red	is significant as it indicates that the variable and the parameter associated with it has not yet been specified.

To specify the response, click on either of the y’s and complete the pop up menu as follows:

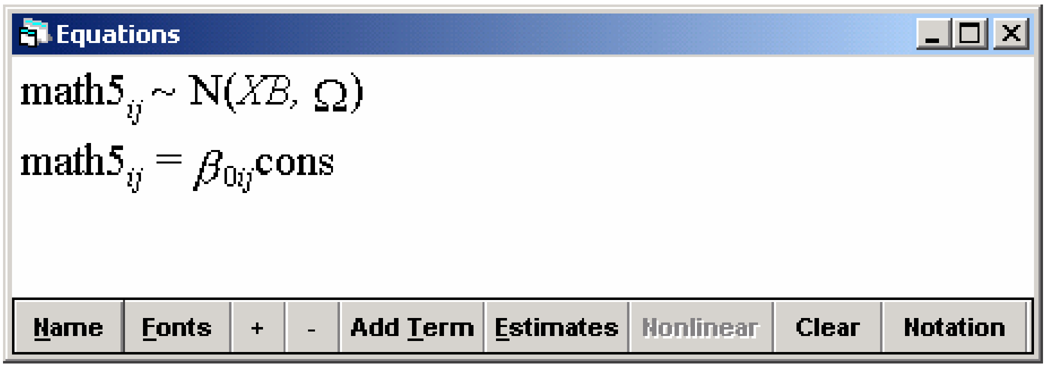
y	math5	[replaces none]
N levels:	2	[that is 2 levels school (j) and pupil (i); replaces none]
Level 2(j):	School	[j is higher level unit]

Level 1 (i)	Pupil	[i is lower level unit]
Done		

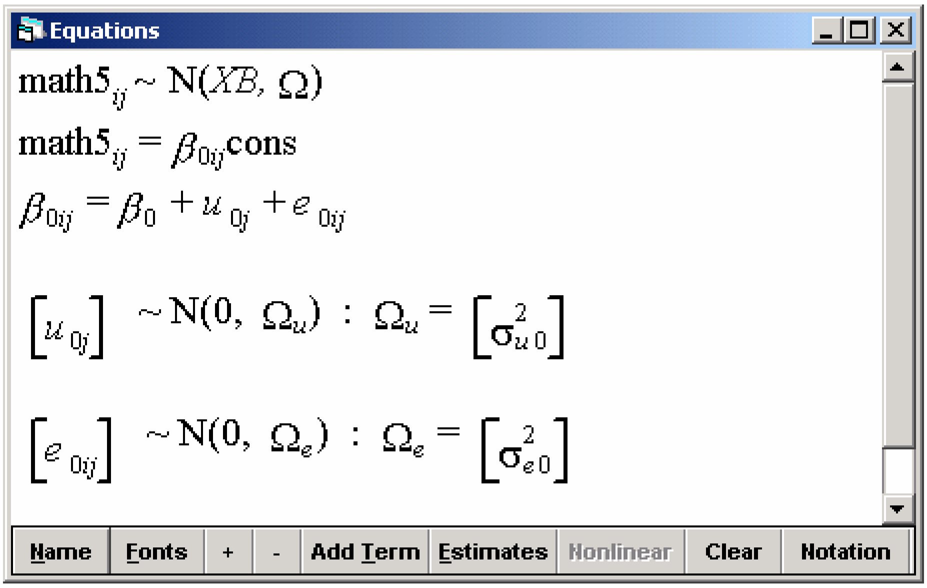
To specify the predictor to be a constant in the random intercepts 'null' model; click on either β_0 or x_0; complete the pop-up menu as follows:

x	cons	[replaces none]
Tick	fixed part	[includes β_0]
Tick	j School	[allows β_0 parameter to vary at level 2]
Tick	i Pupil	[allows β_0 parameter to vary at level 1]

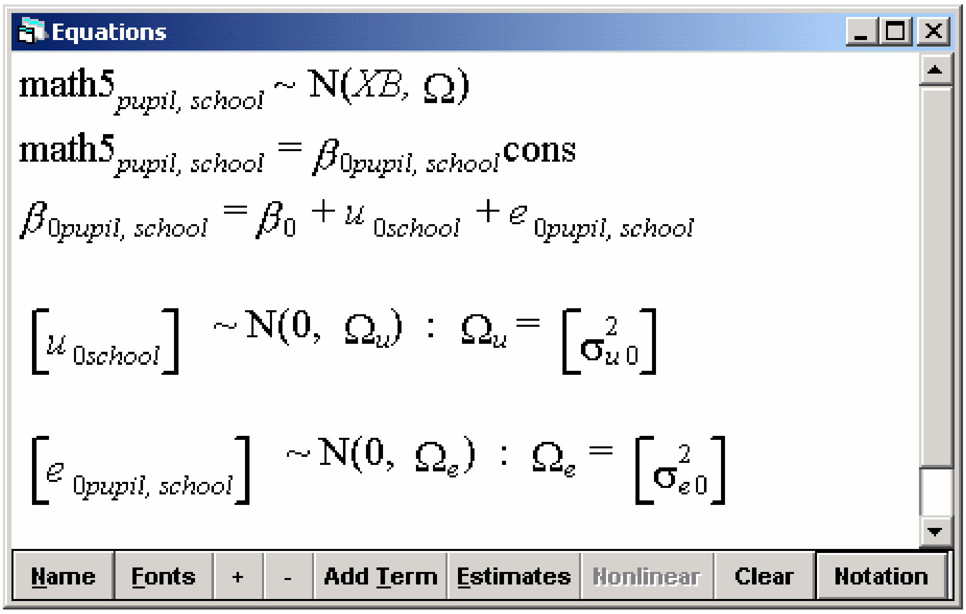
This completes the specification and the revised screen shows the variables and parameters have changed from red to black indicating that specification is complete.



Pressing the + button on the bottom toolbar increases the detail; pressing + again will bring more detail. You should now see the full algebraic specification of the model. Pressing – reduces the detail, clicking Fonts allow the fonts to be changed in terms of size and type.

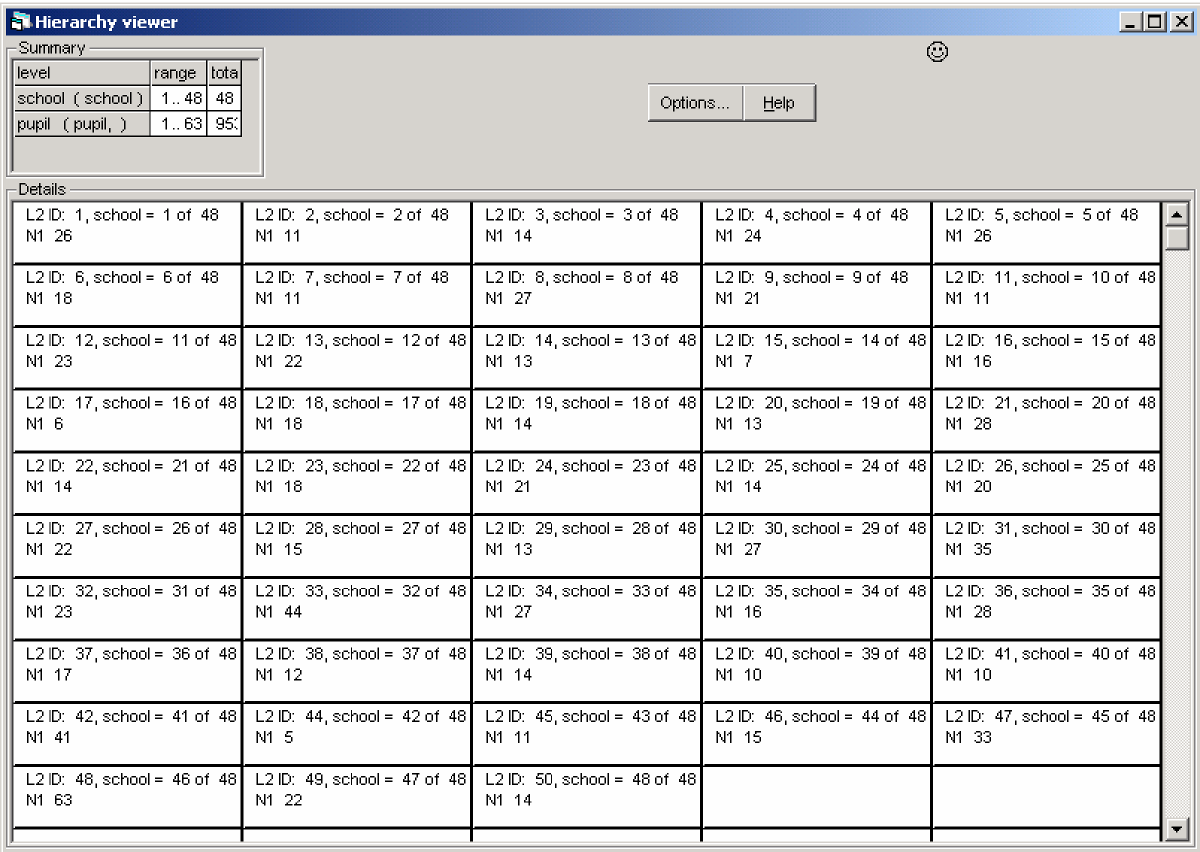


To produce a model that is easier to interpret click on the Name button and then the Notation button (tick subscripts as names) to get the following display



Before proceeding to estimation it is a good idea to check the hierarchy with the following sequence:

Model on main Menu
Hierarchy viewer

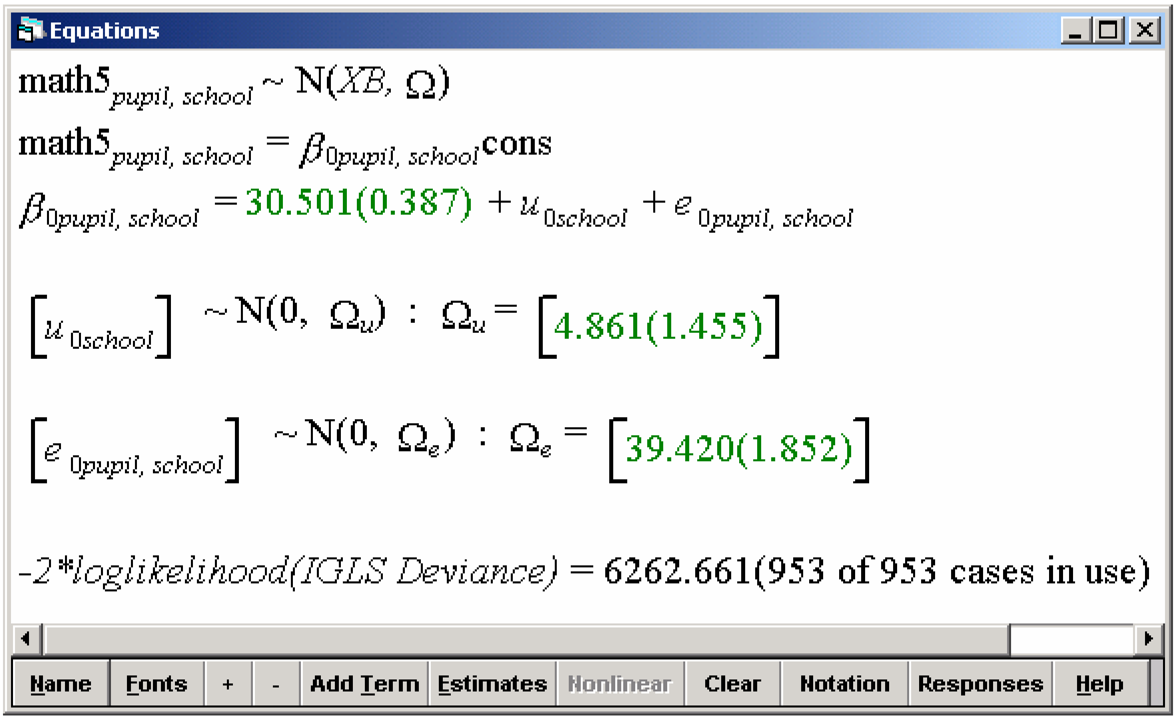


It is possible to see the number of pupils in each and every (higher-level) school. Close the windows when you have examined the structure and it is the same as shown. Any problems are likely to be a result of incorrect sorting. Notice that there are 48 schools, number 10 and 43 are not in our sample.

Estimating the model

Before estimating begins, click on estimates in the lower tool bar twice (on equations window). The blue values are to be ignored, as they are not the converged values. To start estimation click the START button at the top of the screen – watch the screen at the bottom as the fixed and random parameters are estimated school by school and the gauge tanks are filled, and as the iteration counter increases.

As the parameters converge on a stable value, the coefficients in the Equations window will turn green. The letters IGLS next to STOP inform you that the default overestimation procedure is being used: iterative generalised least squares. When they are all green the overall model has converged. For model 1, the following estimates are derived:



The terms in the Equations window represent parameter estimates with their estimated standard errors in brackets; the log-likelihood is a measure of badness of fit, 953 out of 953 cases in use means there are no missing values in our data.

Some questions: 2

What does 30.501 represent?
And 4.861; and is it significantly different from zero? And 39.42?
Does it appear that student achievement varies between schools?

What is overall mean? 30.5
What is total variation around this mean? 4.86 + 39.42
What proportion of the variance is at the school level (ie attributable to the school)?
4.86/ (4.86 + 39.42) = 11%
What are likely bounds of variation on schools (ie confidence interval)? Assuming normality?

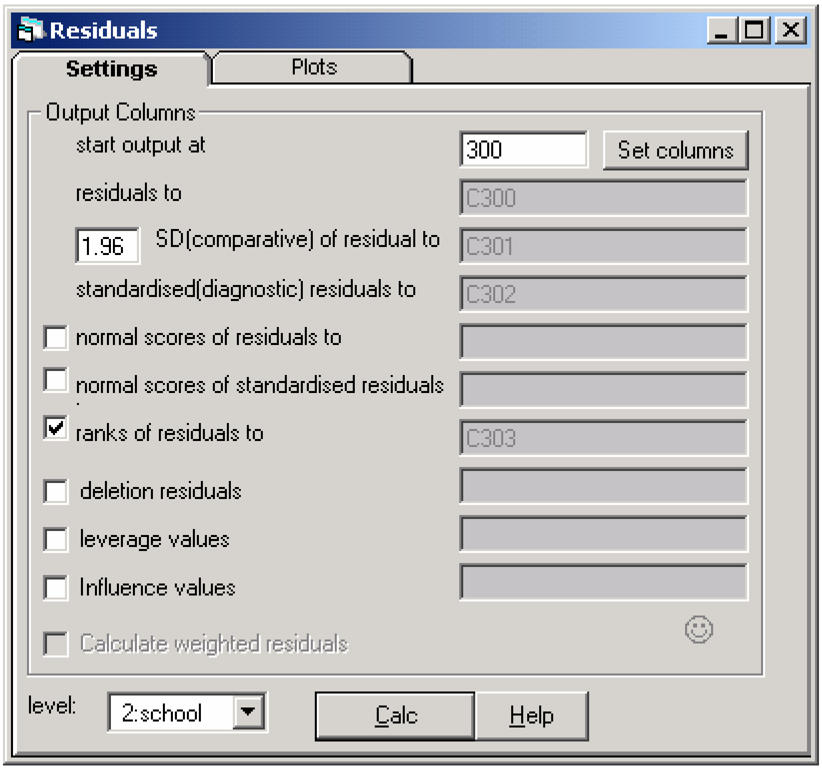
95% of schools lie 30.5 ± 1.96 * sqrt (level 2 variance); that is between 34.8 and 26.2

Estimating Residuals and Graph of Residuals

The next stage is to examine the residuals. One useful procedure is to estimate the level-2 residuals, their ranks and produce a caterpillar plot to see which are significantly different from the overall model. The sequence is:

Model on Main Menu
Residuals
Change 1.0 to 1.96 SD (comparative) of residual [to get standard errors of residuals]
Tick only 'Ranks of residuals to' [untick all other options]
Change level to 2:schools [replace 1:pupil; to get school level residuals]
Click set columns [to get output columns]
Calculate [to estimate]

Don't close down the windows.
The completed screen should look like:



The columns where requested values are to be stored are shown. To view the values you can either use the view data window, or use the command interface to print them out. We chose the latter; the sequence is;

Data Manipulation on Main Menu
Command interface
Type the command in lower left hand box
Print c300 c303 [Level 2 residual and rank]
Press enter

	C300	C303
N =	48	48
1	-3.7826	2.0000
2	-0.49765	20.000
3	1.5372	40.000
4	-1.7136	9.0000
5	0.87883	32.000
6	0.037677	24.000
7	1.2816	37.000
8	-0.10038	22.000
9	0.77229	31.000
10	-0.65464	18.000
11	-0.88464	15.000
12	-0.83096	17.000
13	-1.3033	10.000
14	2.4153	45.000
15	-1.8671	8.0000
16	0.92097	33.000
17	-0.57512	19.000
18	1.2206	35.000
19	0.49685	29.000
20	-2.5208	5.0000
21	-1.2670	11.000
22	-0.99642	13.000
23	2.6617	46.000
24	1.2658	36.000
25	0.67528	30.000
26	-0.067084	23.000
27	-3.8734	1.0000
28	-2.0139	6.0000
29	0.38381	28.000
30	4.0935	48.000
31	-0.85249	16.000
32	-1.0178	12.000
33	2.2921	44.000
34	0.082354	25.000
35	3.6548	47.000
36	0.97511	34.000
37	1.3421	38.000
38	-0.40763	21.000
39	-3.3689	4.0000
40	1.8217	42.000
41	-1.9047	7.0000
42	0.34291	26.000
43	-3.4804	3.0000
44	1.7519	41.000
45	2.2007	43.000
46	-0.93598	14.000
47	0.36467	27.000
48	1.4468	39.000

Some questions: 3

What is the highest achieving school; what does a pupil on average achieve there?
What is the lowest achieving school; what does a pupil on average achieve there?

The extremes ‘add’ and ‘take-away’ 4.1 and -3.9; corresponding to likely 95% confidence interval

Close the command interface and the output window, before proceeding

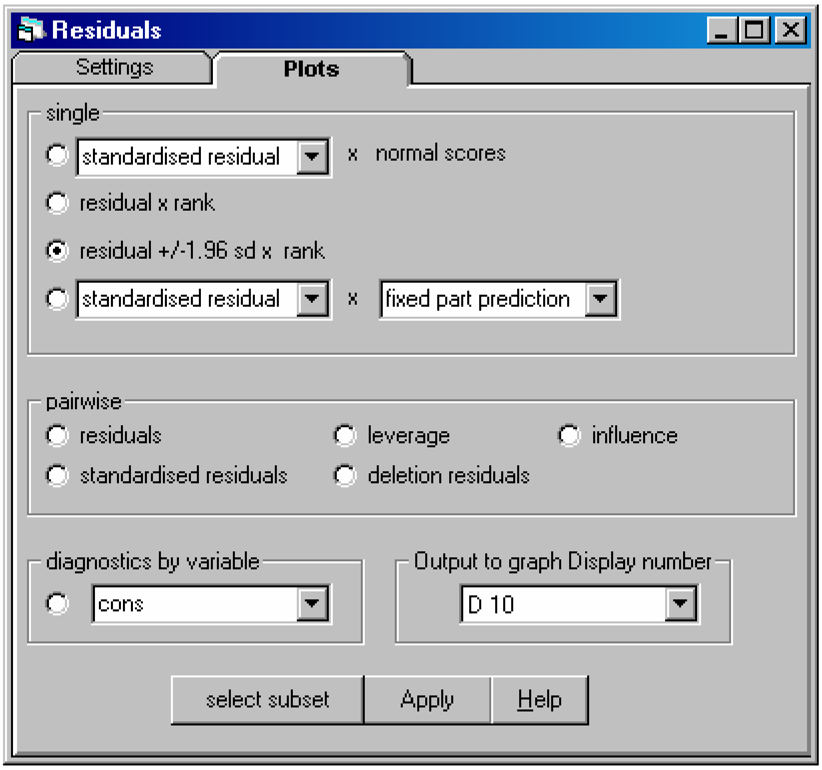
Return to the residuals window and select:

Plot tab

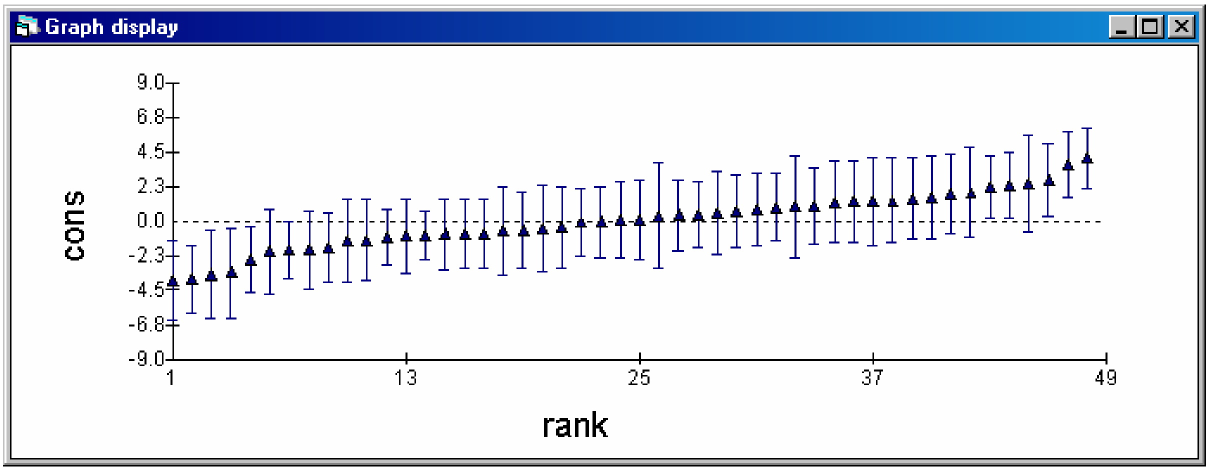
Click residuals +/- 1.96 SD x rank [on single plots section at top of screen]

Apply

Note that D10 is the default graph display for this plot.



This gives a caterpillar plot, which plots each residual with its 95% confidence band against rank.



Some questions: 3

What are the high and low achieving schools?

Click on the graph and use the Identify Points tab

What determines confidence band? Hint: click on one with wide band and look at hierarchy

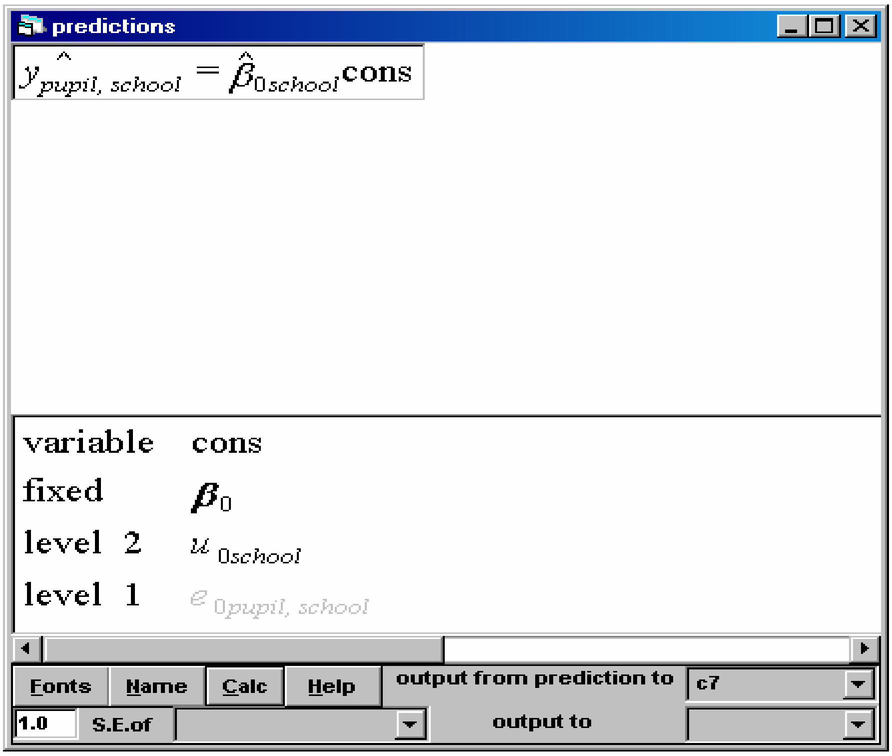
Before proceeding close the Residuals and Graph Display.

Making predications and drawing varying relation plots

The next task is to make predictions of Math5 score in each school and then to plot them on a customised graph.

Model on Main Menu
Predictions

The top screen needs to be completed by choosing items from the middle screen; the bottom buttons control the form of the results and where they are going to be put.

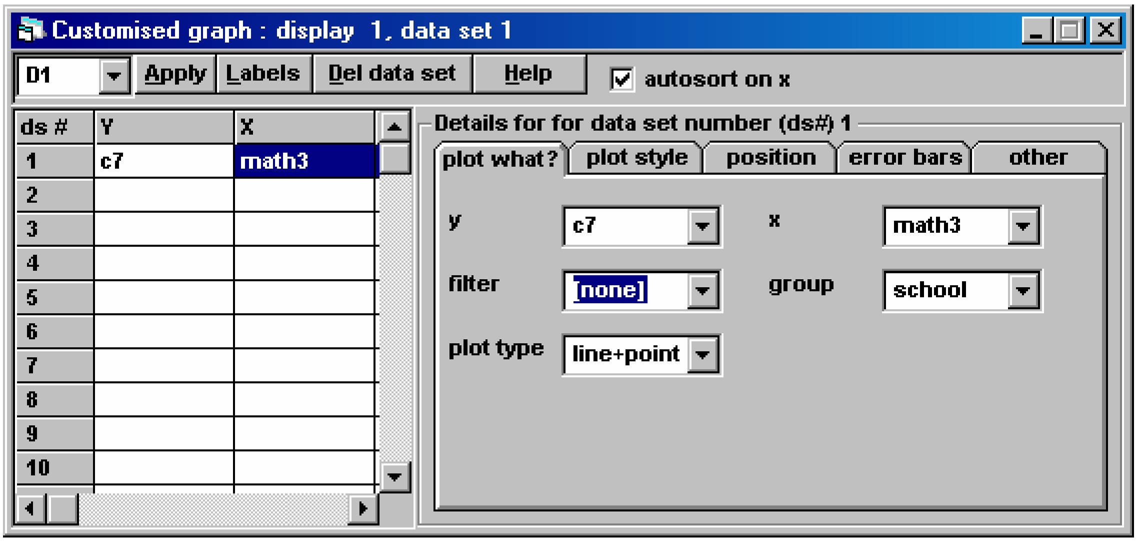


Here is the completed screen to derive the predicted mean Math5 for each school; the level-1 residuals remain greyed out and the results are stored in column 7, which is currently unused. Clicking on an item toggles it in and out of the equation. Calculate needs to be pressed to make the calculations. Nothing appears to happen but if you View the data you will see that a set of predications has been made and put in column 7.

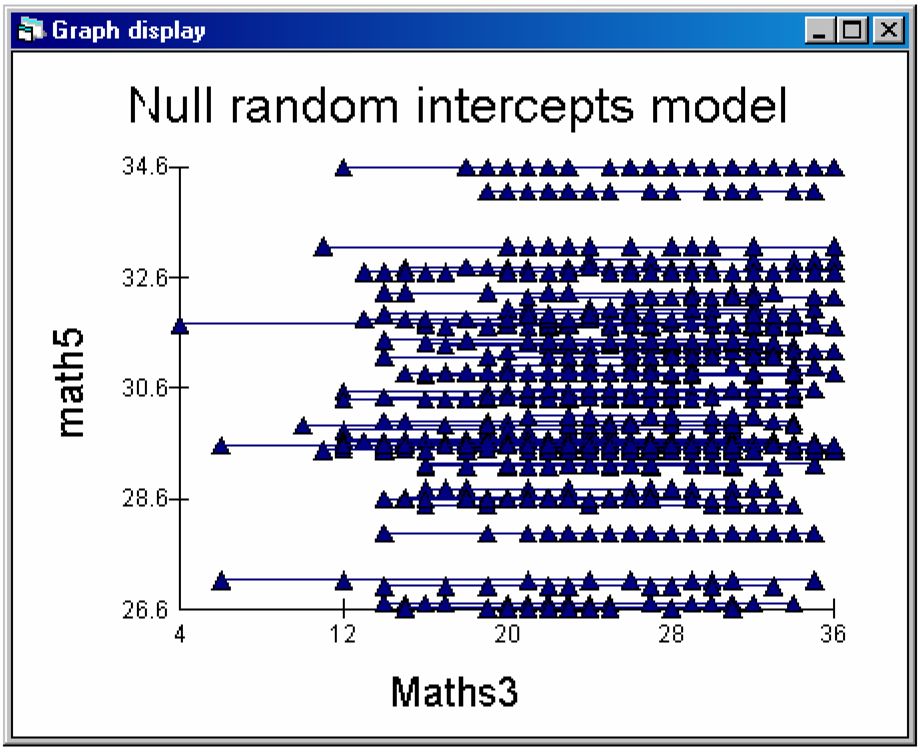
Next bring up the Customised graphics window (Graph on main menu). Currently the D10 graphic display is in operation as this set was used to produce the caterpillar residual plot. Change this to D1.

Choose y is c7 [math5 predictions on y axis]
x is math3 [plot against math3 on x axis; but remember math3 not yet in model]
Group is school [to get a line of predictions for each school]
Plot type is line and point
Apply

Here is the customised setting screen as it should be before pressing Apply:



To add titles to the resultant graph – right click on graph display. Here is the graph after titles have been added:



That completes the first model, save the worksheet as model1.ws, which will include the model equations, graphs and estimates, after giving the name Yhat1 to column 7. Close all windows except the equations window.

4. Model 2: two-level random intercepts model (with a centred predictor: math3)

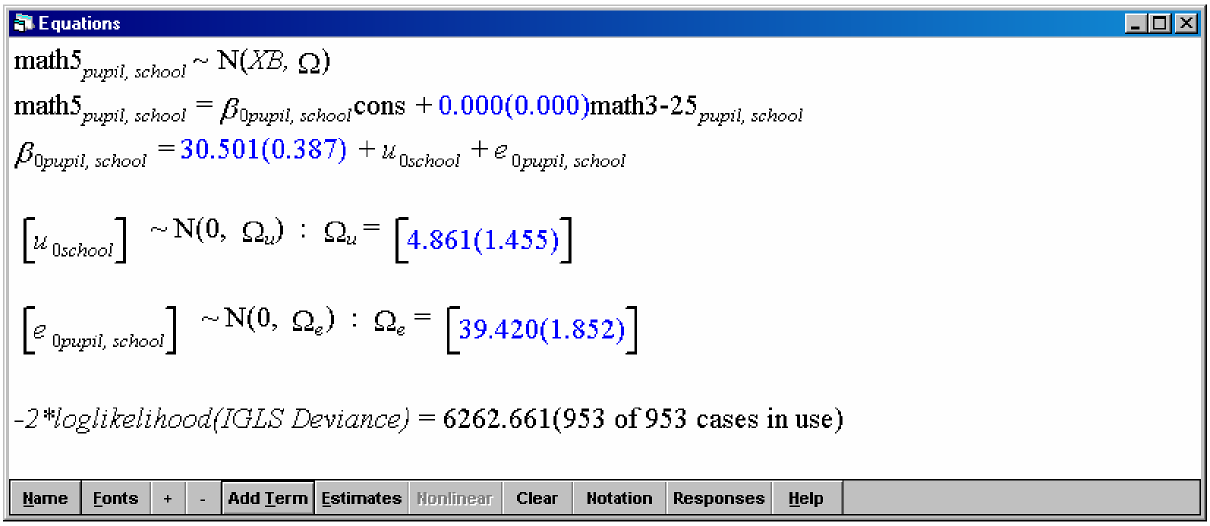
Specifying and estimating the model

We now want to see what happens when we take account of Math3, that is we are modelling *progress* not achievement. It may well be that schools are markedly different in their intake, and this may be the underlying reason for differing school achievement.

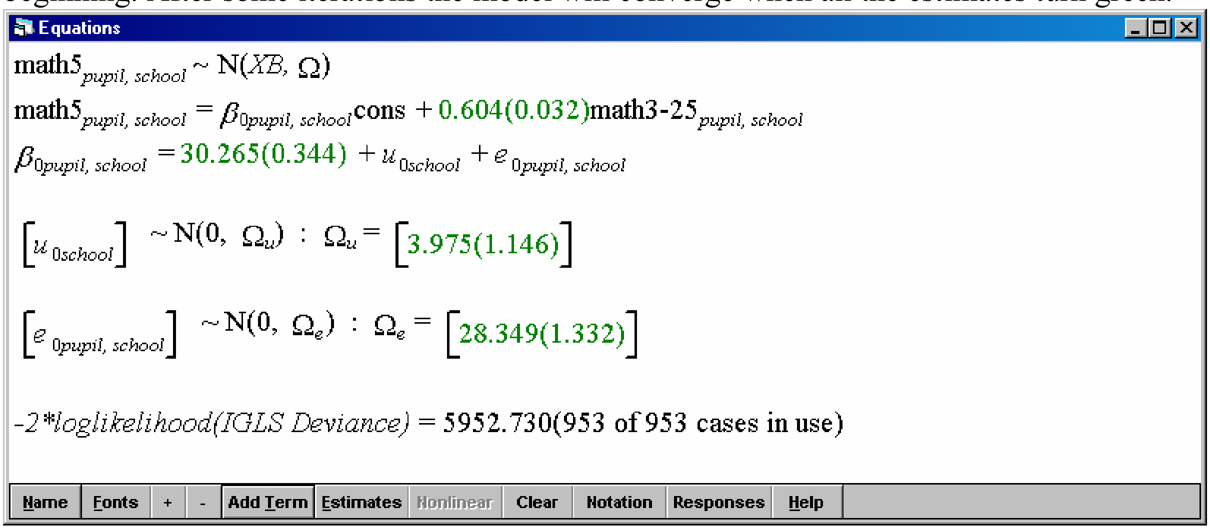
The first thing we have to do is to centre the predictor variable around a convenient value. Use the Basic Statistics window to find that the mean is 25 for Math3. Use the calculate window (or the command window) to create a new variable:

c8 = 'math3' –25

and name c8 with the heading 'Math3-25'. Save the revised worksheet as model2.ws. Return to the equations window. To include the new variable in the fixed part of the model, click on Add Term on the bottom toolbar and specify the variable 'math3-25 by clicking on it from the list. Then click Done.



The initial estimate is zero and the model has to be estimated. By clicking on More in the top toolbar, estimation will progress from the current estimates; START restarts the estimation from the beginning. After some iterations the model will converge when all the estimates turn green.



Some questions: 4

What do the estimates represent?

30.265	(remember Math3 is centred at 25)
0.604	the general rate of progress across all schools
3.975	is there still variation between schools after taking account of math3?
28.349	has pupil level variance changed?

Partitioning the variance

A: Original Variance from null model: 4.861 + 39.24 = 44.28

B: Total residual variance from model 2: 3.975 + 28.349 = 32.224

Proportion of original variance accounted for by 'Math3'

= (B-A)/ A

= (32.224 – 44.28)/ 44.28

= 27% total variance in pupil outcome accounted for

Proportion of remaining variance still unaccounted for at school level

=3.975/ (3.975 + 28.349)

= 12 % (ie 12% of remaining variance is attributable to schools)

What are likely bounds of variation on schools (ie confidence interval)? Assuming normality 95% of schools lie 30.3 ± 1.96 * sqrt (level 2 variance); that is between 34.2 and 26.4

That is typical child starts with 25 score on math3: then progress is typically to 30.3 math5 score - but in top 2.5% of schools the average progress is to 34.2 math5 score and in bottom 2.5% of schools the average progress is to 26.4 math5 score.

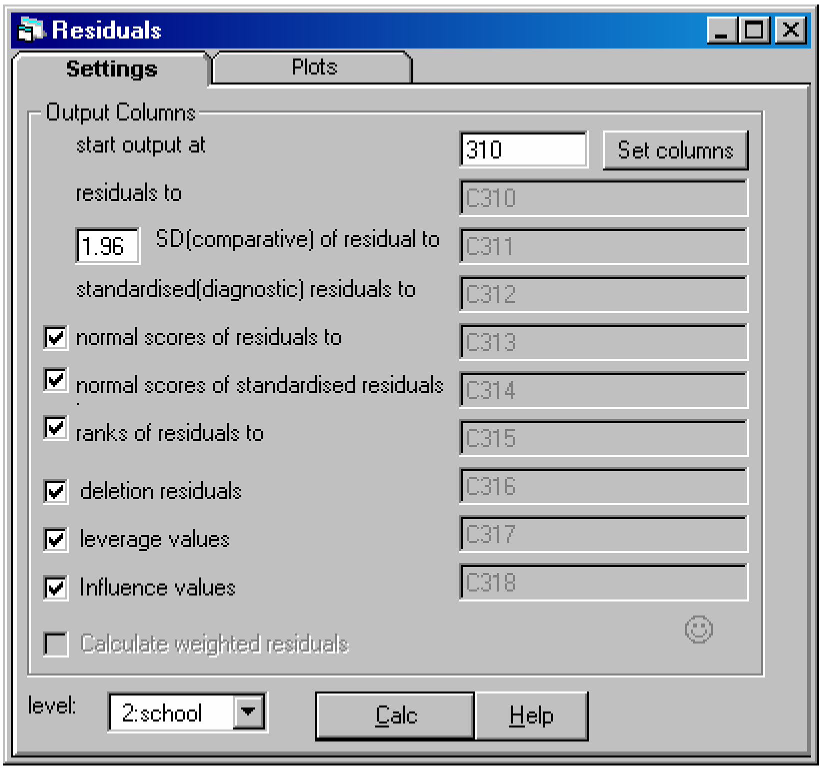
Notice that with 1 extra explanatory variable the loglikelihood/deviance has decreased from to 6263 to 5952 with a single parameter.

Residuals and graph of residuals

Model on Main Menu

Residuals

Start output at c310	[not to overwrite existing columns/residuals]
Change 1.0 to 1.96 standard errors	[to get 95% confidence intervals]
Tick all types of residuals	
Level 2: schools	[replace 1:pupil; to get school level residuals]
Click Set columns	[to get all output columns]
Click Calculate	[to estimate]



Data manipulation on Main Menu

Command interface

Print c300 c303 c310 c315

This gives the following data in the out put window

Model 1: c300: school residuals c303: rank

Model 2: c310: school residuals c315: rank

	C300	C303	C310	C315
N =	48	48	48	48
1	-3.7826	2.0000	-2.6157	5.0000
2	-0.49765	20.000	-0.21464	23.000
3	1.5372	40.000	0.87949	36.000
4	-1.7136	9.0000	-2.0311	7.0000
5	0.87883	32.000	-0.22142	22.000
6	0.037677	24.000	-0.12780	25.000
7	1.2816	37.000	1.5606	40.000
8	-0.10038	22.000	0.26056	27.000
9	0.77229	31.000	-1.8780	8.0000
10	-0.65464	18.000	0.75171	34.000
11	-0.88464	15.000	-0.77298	12.000
12	-0.83096	17.000	-0.14164	24.000
13	-1.3033	10.000	-0.76421	13.000
14	2.4153	45.000	0.69202	33.000
15	-1.8671	8.0000	-0.55611	17.000
16	0.92097	33.000	0.45405	31.000
17	-0.57512	19.000	-0.28410	20.000
18	1.2206	35.000	1.0052	38.000
19	0.49685	29.000	-0.43571	19.000
20	-2.5208	5.0000	-3.0554	2.0000
21	-1.2670	11.000	-1.5689	9.0000
22	-0.99642	13.000	-0.57769	16.000
23	2.6617	46.000	2.9306	44.000
24	1.2658	36.000	1.7952	42.000
25	0.67528	30.000	0.98509	37.000
26	-0.067084	23.000	0.33708	28.000
27	-3.8734	1.0000	-2.8753	3.0000
28	-2.0139	6.0000	-3.3681	1.0000
29	0.38381	28.000	1.3776	39.000
30	4.0935	48.000	3.0525	47.000
31	-0.85249	16.000	0.42203	30.000
32	-1.0178	12.000	-1.3651	10.000
33	2.2921	44.000	3.0222	46.000
34	0.082354	25.000	-0.63784	14.000
35	3.6548	47.000	2.9485	45.000
36	0.97511	34.000	0.83096	35.000
37	1.3421	38.000	3.1995	48.000
38	-0.40763	21.000	-1.3267	11.000
39	-3.3689	4.0000	-2.5776	6.0000
40	1.8217	42.000	2.0989	43.000
41	-1.9047	7.0000	-0.45035	18.000
42	0.34291	26.000	-0.22856	21.000
43	-3.4804	3.0000	-2.8256	4.0000
44	1.7519	41.000	0.40773	29.000
45	2.2007	43.000	1.6816	41.000
46	-0.93598	14.000	-0.62480	15.000
47	0.36467	27.000	0.16175	26.000
48	1.4468	39.000	0.67059	32.000

Some questions: 5

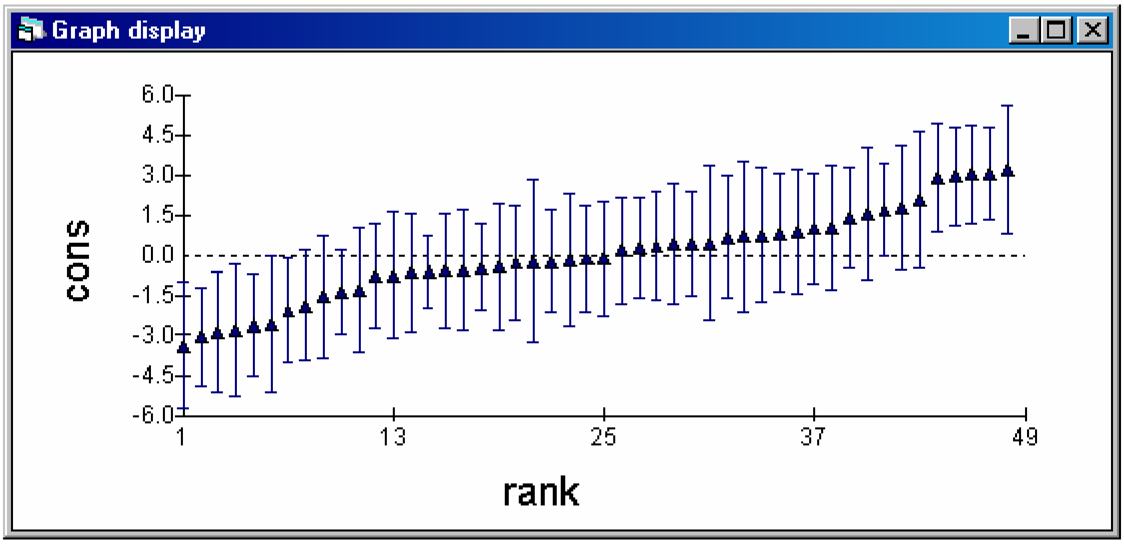
After taking account of pre-test (math3), what is the school with the most progress, and the least?

Values of 3.2 and -3.3 compared to ?

What has happened to particular schools and why?

School 6?
School 9?
School 10?

Close the command interface and output windows and return to the Plots tab of the residuals window, choose the same plot as before to get the following caterpillar plot:



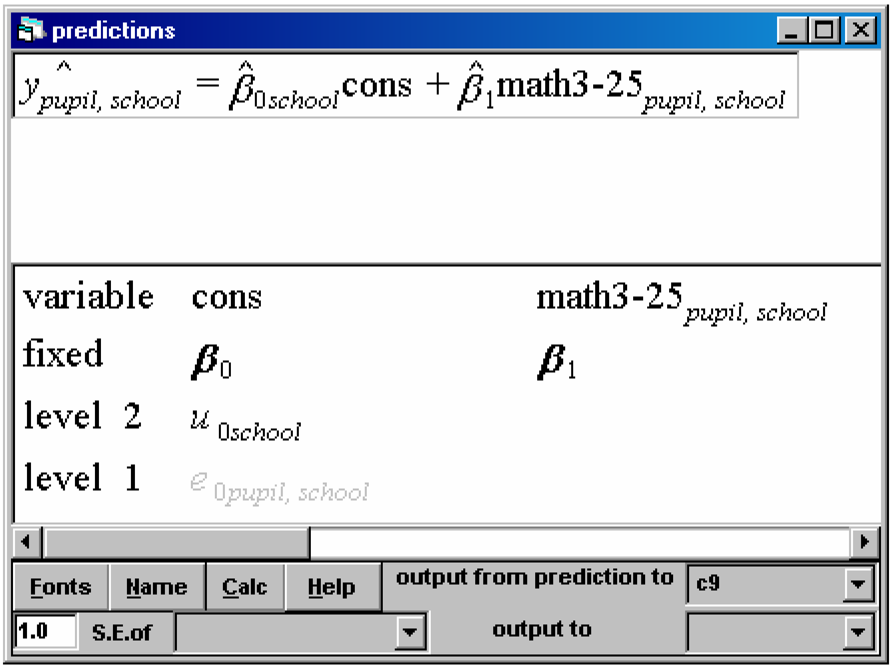
Some questions: 6

After taking account of pre-test, what can you say about the majority of schools? And hence league tables?

Making Predictions and drawing varying relations plots

Model on Main Menu
    Predictions

Complete the window as follows putting the revised school estimates to c9



The residuals at level 1 **must** remain greyed out

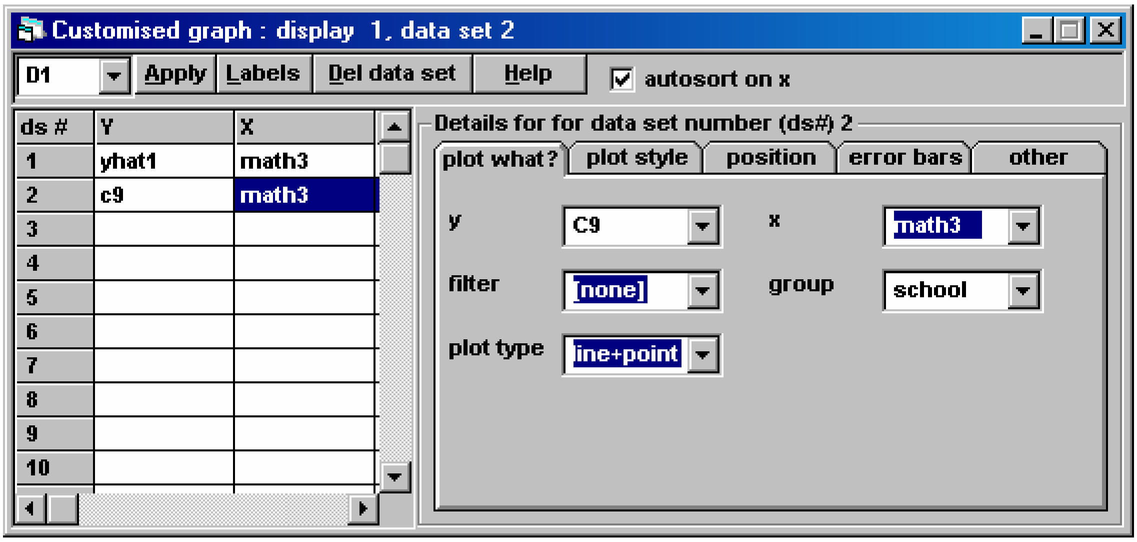
Graphics on Main Menu

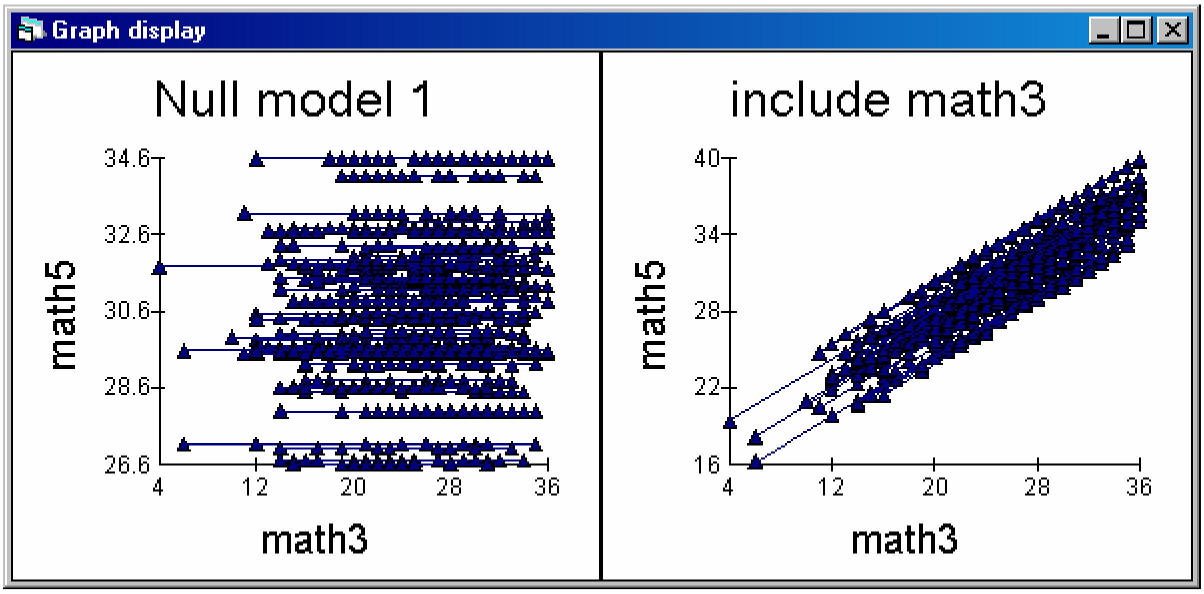
Customised graphics	
Switch to D1	[display graph set D1]
Click on right side to ds#2	[sub-graph not to overwrite ds#1]
y is c9	[type in c9 if not on list; revised math5 predictions on y-axis]
x is math3	[plot against math3]
Group is school	[to plot school lines]
Plot type is line+point	

Also on Plot' position' tab select position of graph ds#2

Mark X in col 2 and row 1	[original plot for model 1 in col 1 row 1]
Apply	

The Plot what screen should show that there are two sub-graphs in graph set display D1


Customised graph : display 1, data set 2
D1
Apply
Labels
Del data set
Help
autosort on x
ds #
Y
X
1
yhat1
math3
2
c9
math3
3
4
5
6
7
8
9
10
Details for for data set number (ds#) 2
plot what?
plot style
position
error bars
other
y
C9
x
math3
filter
[none]
group
school
plot type
line+point

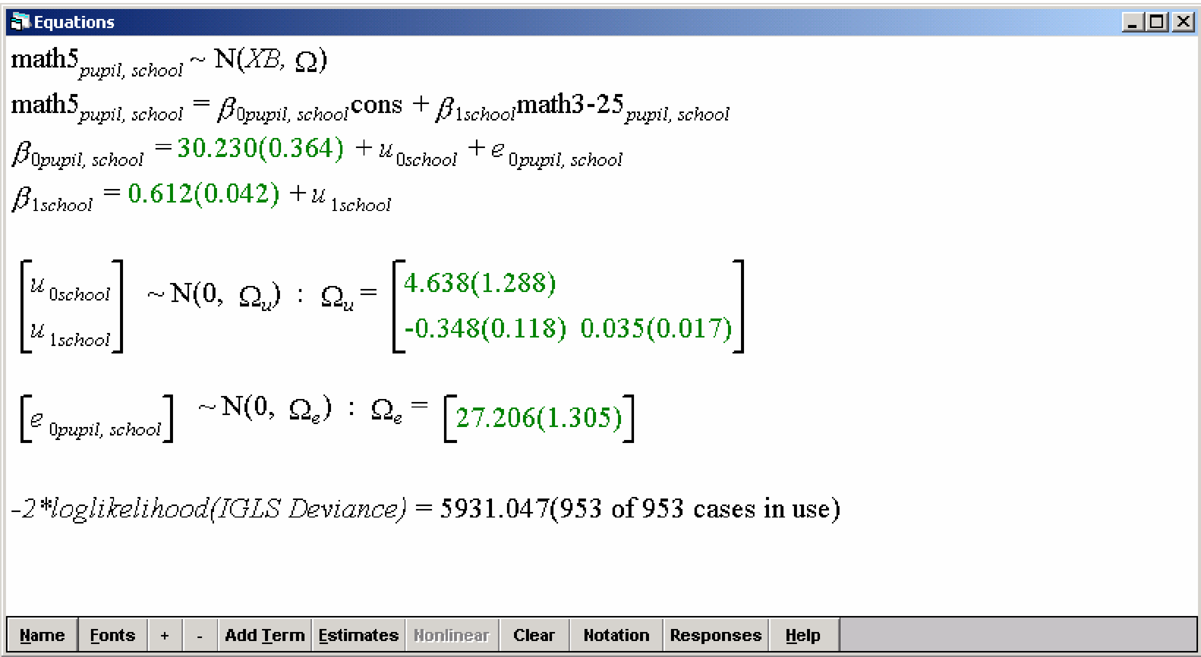

Graph display
Null model 1
math5
34.6
32.6
30.6
28.6
26.6
4
12
20
28
36
math3
include math3
math5
40
34
28
22
16
4
12
20
28
36
math3

Model 3: a random-intercepts and random-slope model (fully random at level 2)

Specifying and estimating the model

Return to the equations window

Click on Math3-25	[to get math3-25 variable pop-up menu]
Tick school as well as fixed	[to allow parameter to vary across schools]
Click Done	[to close window]
Click More	[continue estimation, blue to green]



Some questions: 7

What do the estimates represent?

30.23
0.61
4.638
-0.348
0.035
27.206

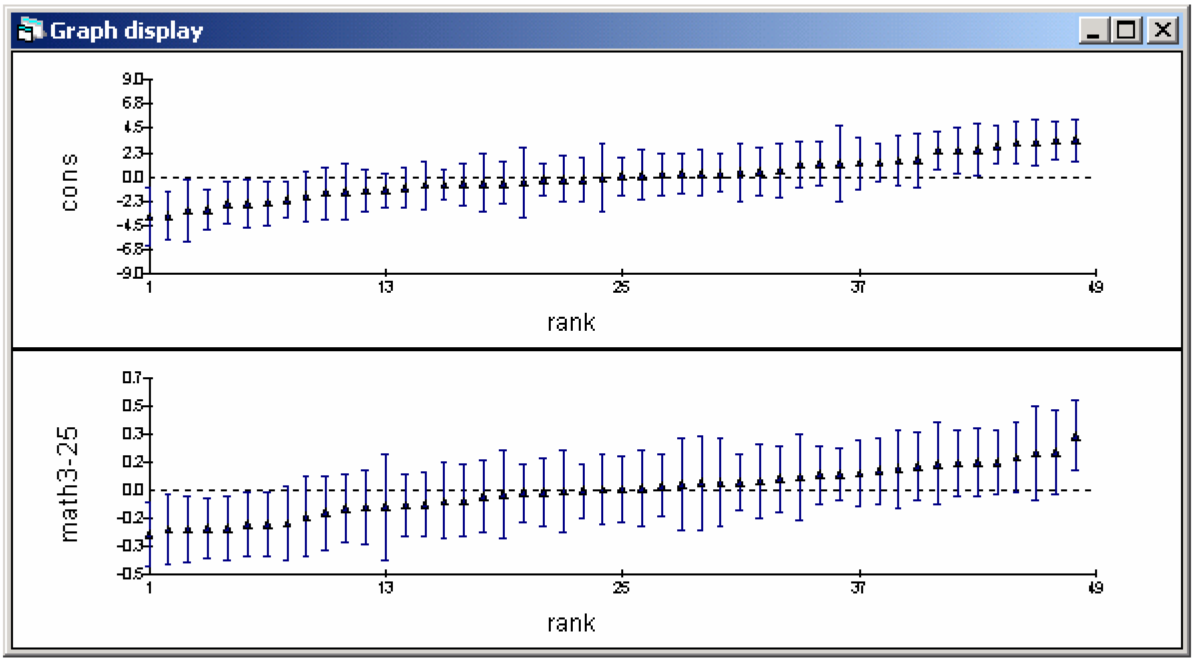
What do you think the school lines will look like? Do you anticipate fanning in or out?

Residuals and graph of residuals

Model on Main Menu
Residuals
Start output at c320 [not to overwrite existing columns/residuals]
Change 1.0 to 1.96 standard errors [to get 95% confidence bands]
Tick all types of residuals
Level 2: schools [replace 1:pupil; to get school level residuals]
Click Set columns [to get all output columns]
Click Calculate [to estimate]

Return to residuals window
Plot tab
Click residuals +/- 1.96 SD x rank [on single plots pane]
Apply [to get two plots in D10]

Two plots produced automatically



Some questions: 8

What does top graph show?	Intercepts
And bottom?	Slopes

Notice that there are two columns for each and every residual and that residuals are in c320 – c321 and ranks in c330 – c33. To print out residuals:

Data manipulation
Command interface
Print c320 c330 c321 c331

	C320	C330	C321	C331
N =	48	48	48	48
1	-2.0744	8.0000	0.098642	35.000
2	0.15082	26.000	-0.038014	18.000
3	0.55224	33.000	-0.0068282	22.000
4	-2.4376	5.0000	0.20626	45.000
5	-0.23633	22.000	0.015039	26.000
6	0.33606	29.000	-0.067354	17.000
7	1.2918	37.000	-0.067740	16.000
8	0.067890	25.000	0.029201	27.000
9	-3.1320	3.0000	0.23388	46.000
10	0.47875	32.000	-0.010702	21.000
11	-1.0391	14.000	0.10915	37.000
12	0.31135	28.000	-0.083578	15.000
13	-0.75296	15.000	0.057591	32.000
14	1.2599	36.000	-0.10221	13.000
15	-0.65434	17.000	0.074478	33.000
16	-0.079283	24.000	0.033971	28.000
17	-0.20206	23.000	0.0082753	25.000
18	1.4933	39.000	-0.13783	10.000
19	-0.59096	18.000	0.047515	30.000
20	-3.0173	4.0000	0.17771	43.000
21	-1.7869	9.0000	0.13736	39.000
22	-1.2481	12.000	0.17126	42.000
23	3.1448	45.000	-0.24154	3.0000
24	2.4185	42.000	-0.23345	5.0000
25	1.2384	35.000	-0.11504	11.000
26	0.20623	27.000	0.0060090	24.000
27	-2.4239	7.0000	0.15366	40.000
28	-3.7080	1.0000	0.23911	47.000
29	1.3084	38.000	-0.093209	14.000
30	3.3874	47.000	-0.21041	7.0000
31	0.36306	30.000	-0.015948	20.000
32	-1.2331	13.000	0.051546	31.000
33	2.9468	44.000	-0.21368	6.0000
34	-1.4890	10.000	0.16521	41.000
35	3.4022	48.000	-0.24268	2.0000
36	1.1556	34.000	-0.10572	12.000
37	3.1609	46.000	-0.27432	1.0000
38	-1.3364	11.000	0.081526	34.000
39	-2.4344	6.0000	0.17968	44.000
40	2.5131	43.000	-0.20469	8.0000
41	-0.29547	21.000	-0.0029025	23.000
42	-0.49893	20.000	0.046242	29.000
43	-3.5870	2.0000	0.34222	48.000
44	0.45211	31.000	-0.028513	19.000
45	2.4009	41.000	-0.23636	4.0000
46	-0.74351	16.000	0.10310	36.000
47	-0.57301	19.000	0.12434	38.000
48	1.5333	40.000	-0.16021	9.0000

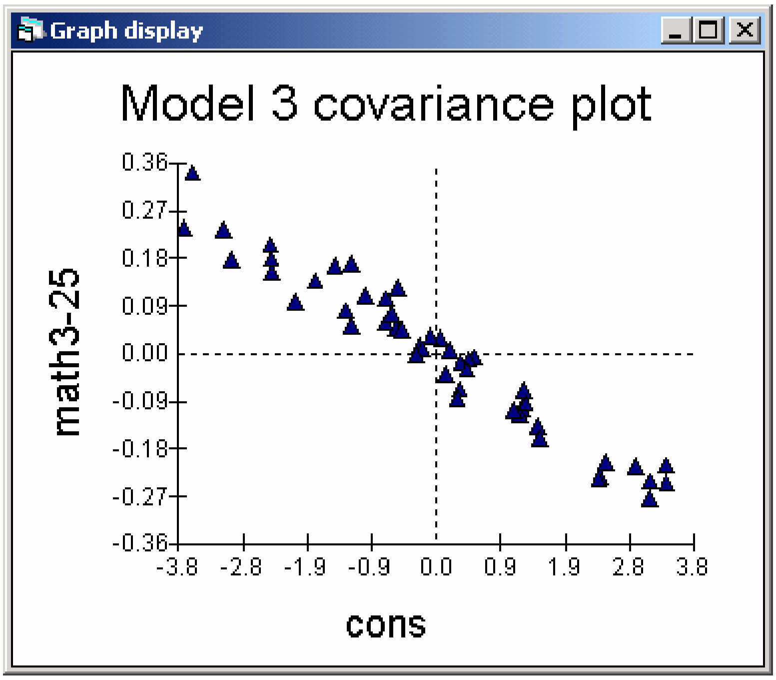
Some questions: 9

What does a pupil with a score of 25 achieve in school 8?	30.23 + 0.07
In school 28?	30.23– 3.71
In school 35?	30.23 + 3.40
What does a pupil with a score of 35 achieve in school 8?	30.23 + 0.07 + 10* (0.61 + 0.03)
In school 28?	30.23 – 3.71 + 10* (0.61 + 0.24)
In school 35?	30.23 + 3.40 + 10* (0.61 – 0.24)

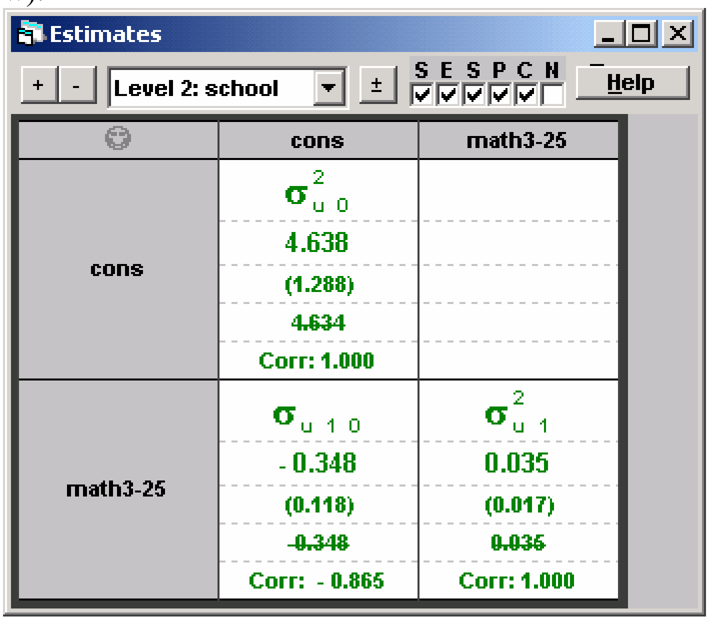
Close output and command interface. Save revised worksheet as model3.ws.

To get a covariance plot

Return to residuals window
    Plots tab
        Tick Residuals on pairwise pane [to get covariance plot]
        Click Apply
        Click in graph
        Graph title model 3: covariance plot



Strong tendency for schools that are good for the average pupil (right in horizontal axis) to make comparatively less progress for higher ability pupils (bottom on vertical axis), estimated correlation is -0.87 (via Main menu Model -> Estimates table). But picture most clearly summarised by varying relation plots (see below).


Estimates

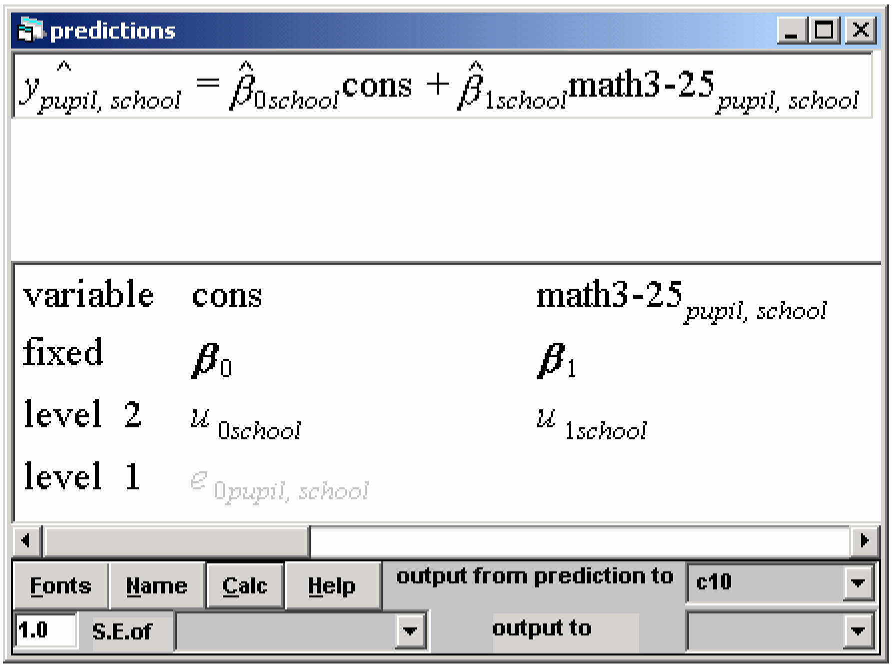
Level 2: school

	cons	math3-25
cons	σ^2_{u0} 4.638 (1.288) ~~4.634~~ Corr: 1.000	
math3-25	σ_{u10} - 0.348 (0.118) ~~-0.348~~ Corr: - 0.865	σ^2_{u1} 0.035 (0.017) ~~0.035~~ Corr: 1.000

Predictions and varying relations plots

Model on Main Menu
    Predictions
        Click on Cons   [to get all terms associated with Constant included]
        Click on Math3-25   [to get all terms associated with Math3-5 included]
        Click on level-1 residuals associated with Cons to exclude
        Output to c10   [free column]
        Click Calculate   [to estimate]

Name c10 as 'Yhat3' and save the revised worksheet.



To get two different kinds of varying relation graphs for model 3 follow instructions below.

Plot 3:
Graphics on Main Menu
    Customised graphs
    Switch to D1   [display graph set D1]
    Click on right side to ds#3   [sub-graph not to overwrite ds#1 & ds#2]
    y is Yhat3
    x is math3   [plot against math3 on x axis]
    Group is school   [to plot predicted school lines]
    Plot type is line+point

Also on Plot' position' tab select position of graph ds#3
    Mark X in col 1 and row 2   [original plot for model 1 in col 1 row 1]
    Apply

Plot 4:

Still on graph set D1

Click on right side to ds#4	[sub-graph not to overwrite ds#1, ds#2 & ds#3]
y is math5	[math5 observed scores on y axis]
x is math3	[plot against math3 on x axis]
Group is none	[to plot raw data]
Plot type is point	

Also on Plot' position' tab select position of graph ds#4

Mark X in col 2 and row 2	[original plot for model 1 in col 1 row 1]
Apply	

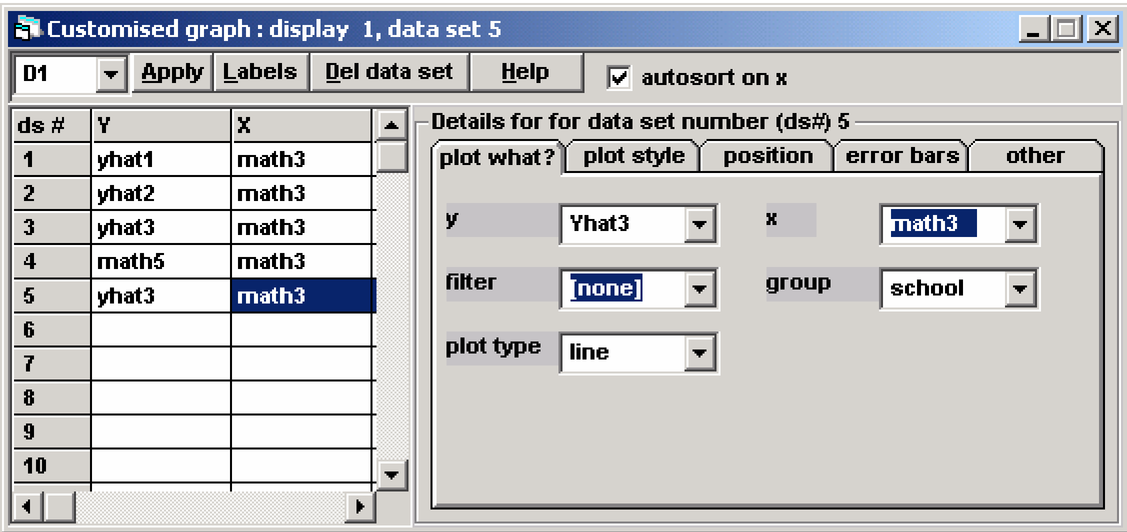
Still on graph set D1

Click on right side to ds#5	[sub-graph ds#5 not to overwrite ds#1-4]
y is Yhat3	[revised math5 predictions on y-axis]
x is math3	[plot against math3 on x axis]
Group is school	[to plot predicted school lines]
Plot type is line	

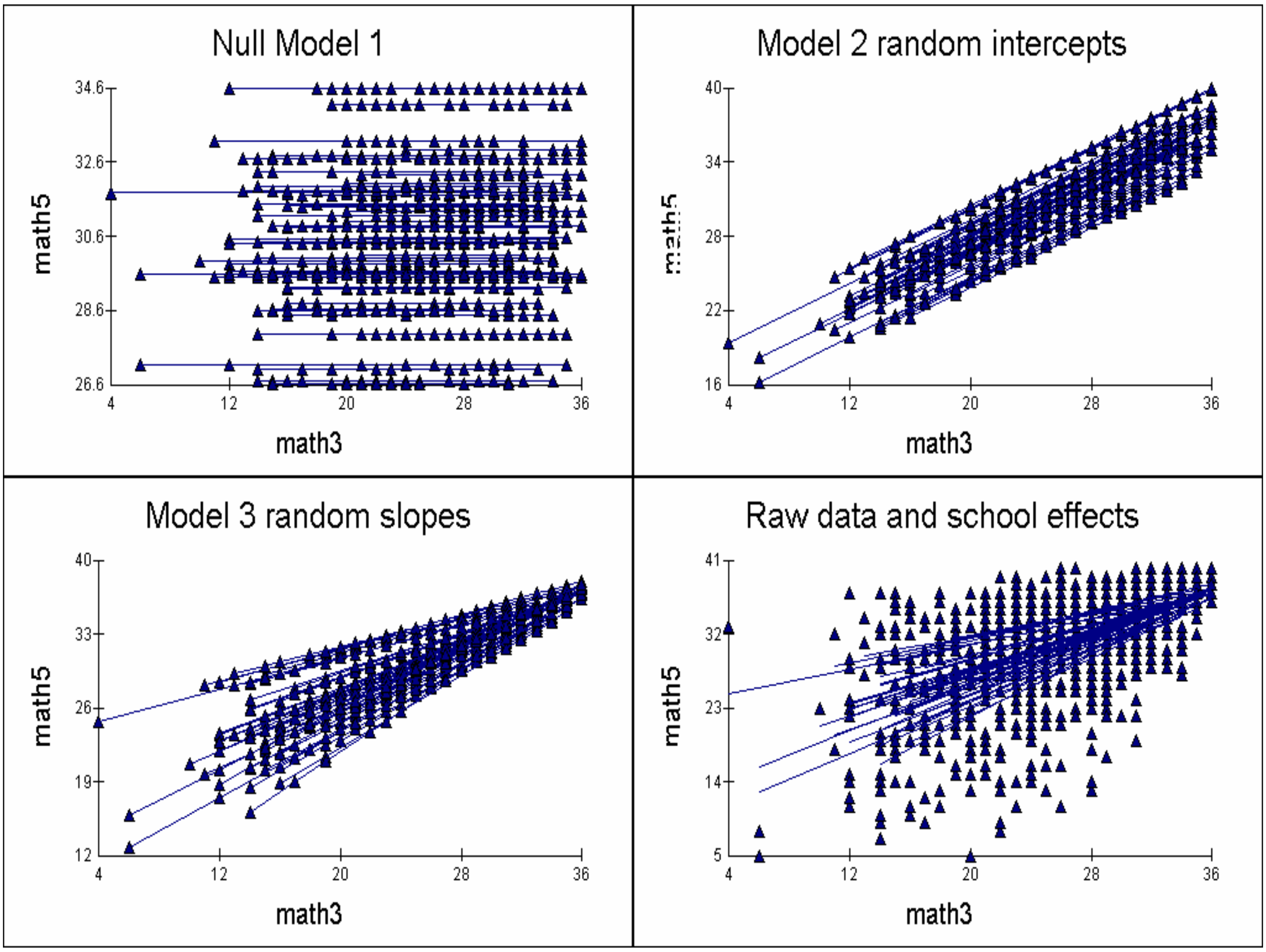
Also on Plot' position' tab select position of graph ds#5 (overlaid on ds#4)

Mark X in col 2 and row 2	[original plot for model 1 in col 1 row 1]
Apply	

The Plot what screen should show that there are five sub-graphs in graph set display D1 (sub-graphs 4 & 5 are overlaid on top of each other to superimpose model 3 school lines on raw data for Plot 4):



We can see from the four summary plots shown below that schools matter for the least able!



Some questions: 10

But …………..is there a ceiling effect?
But……………are boy girl differences important?

What about school-level variables? Etc etc

References

Mortimore, P., Sammons, P., Stoll, L., Lewis, D., and Ecob, R. (1989) School Matters, London, Open Books.

Sally Thomas
Graduate School of Education
University of Bristol

25th Oct 2007