Multilevel Masterclass

University of Stirling

Kelvyn Jones

Session Three: Fitting a Two-Level Model in MLwiN (used Beta9)

Introduction

This chapter aims to provide a straightforward example of fitting a two-level model with a continuous response and a continuous predictor. It is concerned with the practicalities of model specification and estimation. In essence it is a very short guide to the GUI of *MLwiN*. As always with this program there are several ways of doing the same thing and we will try and guide you through a convenient route. We will consider the following models:

1. a random intercepts null model with Price as the response; no predictor variables (apart from the Constant) and with the levels defined as houses in districts; the so-called empty or null RI model
2. a model which additionally includes the Size of house;
3. a model in which the parameter associated with Size is allowed to vary over District; that is random slopes as well as intercepts;
4. a model in which a particular district is treated as an outlier.

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

- data input; sorting, creating the constant term;
- model specification: response, predictors, 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 ills 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.

Data input and manipulation (new version reads SPSS, Minitab, Stata, Excel files)

Here is a recommended sequence to read an ASCII file:

Data input
File on Main Menu
ASCII text file input
Columns: c1-c5
File: c:\kjtemp\house.dat (change to all files to see this one)
OK

Name columns
The Names window will open automatically; highlight each column in turn and click on Edit names to give the following names

Names

C1:	House	enter
C2:	District	enter
C3:	Price	enter
C4:	Size	enter
C5:	Type	enter

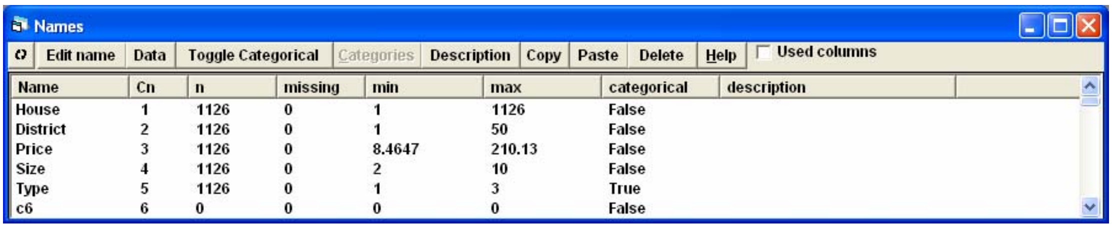
Naming categories
Highlight Type and Toggle Categorical which will change the categorical heading from false to true. Keeping Type highlighted , click on Categories; which will bring up the Set categories dialog box; highlight each name in turn, click Edit and give the categories as shown

Categories

1:	Terrr
2:	Semi
3:	Det

OK to complete

The completed Names window should be as follows



Name	Cn	n	missing	min	max	categorical	description
House	1	1126	0	1	1126	False	
District	2	1126	0	1	50	False	
Price	3	1126	0	8.4647	210.13	False	
Size	4	1126	0	2	10	False	
Type	5	1126	0	1	3	True	
c6	6	0	0	0	0	False	

Save the worksheet
File on Main Menu
Save worksheet as
c:\kjtemp\house2.ws
Remember to write down the complete filename you have used.

Saving the worksheet will save the data, the names, the categories, the equations the model specification, the current estimates and the commands to re-draw any graphs

Sorting the data: houses within districts

The program requires that the data are sorted so that all lower level units are grouped by higher level units; this is achieved by sorting. It is very importnat that all other relevant data are 'carried' in this sort ; otherwise, the data will get out of order and incorrect results will arise.

Data Manipulation on Main Menu
- Sort
  - Increase number of keys to 2
    - Choose District as the highest key [slowest changing]
    - Choose House as the lowest key [fastest changing]
      - Highlight House to Type
        - Same as Input
          - Add to Action List
            - Execute
            - Close Sort window

Check data and save sorted worksheet

In the Names window (you can use the tabs at the bottom of the main MLwiN window to navigate between currently opened windows)

Highlight the columns names House to Type inclusive and click on Data which should bring up the data extract

	House(1126)	District(1126)	Price(1126)	Size(1126)	Type(1126)
1	1.000	1.000	77.368	5.000	Semi
2	2.000	1.000	63.287	5.000	Semi
3	3.000	1.000	59.928	5.000	Terr
4	4.000	1.000	102.940	6.000	Semi
5	5.000	1.000	66.846	4.000	Semi
6	6.000	1.000	71.189	5.000	Terr
7	7.000	1.000	80.391	6.000	Terr
8	8.000	1.000	62.173	4.000	Terr
9	9.000	1.000	71.998	5.000	Semi
10	10.000	1.000	84.147	5.000	Terr

If it looks correct, save the revised data (it is good practice to this as you go along)

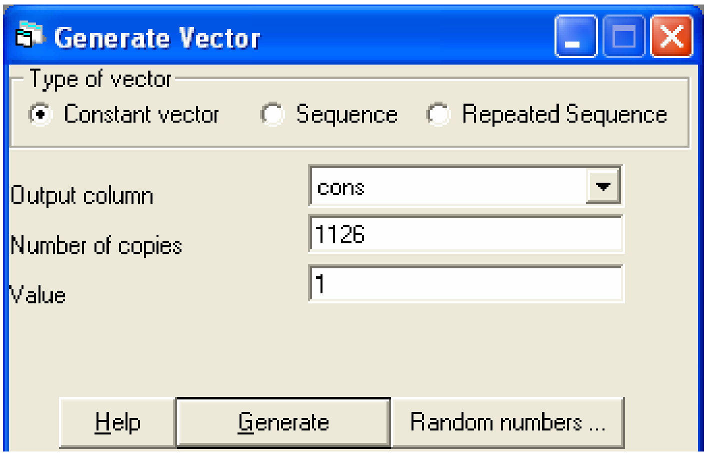
File on Main Menu
- Save (as House 2.ws)
- Yes to overwrite

There is a final variable we have to create before beginning modeling: the constant; that is a set of 1's. There are many ways of doing this but you must ensure that there is a 1 for each and every house. The simplest way to achieve this is:

Data Manipulation on The Main Menu
- Generate Vector
  - Constant Vector
    - Output Column: 6
      - Number of copies: 1126
        - Value: 1
          - Generate

Close window

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



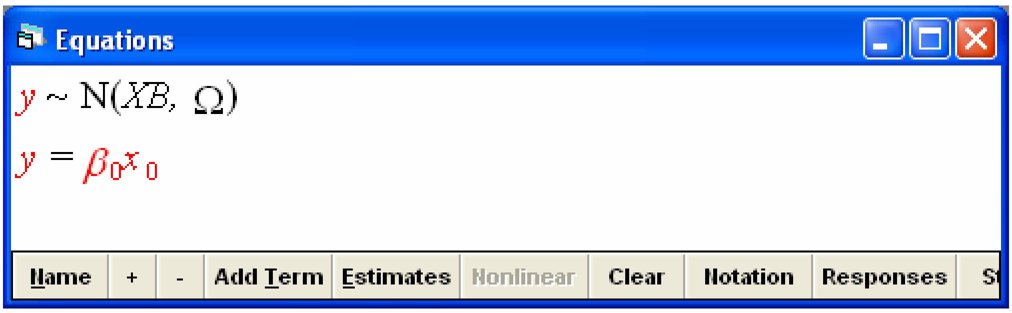
Edit the name c6 so that it is called 'cons'. After saving the revised data, you are ready for modeling; close the View data windows.

Model 1: two-level null random intercepts

Specifying the model

Go to Model on the main menu. Clicking on Equations will bring up the following screen which is the heart of the program. Here models are specified and estimates displayed. (It is also possible to specify models in the command window and to see the equations displayed there).

Ignoring the bottom tool bar for the moment; there are two equations:



- y is the response;
- N indicates a normal distribution for a fixed part Xβ 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 (or probably a paler grey in these notes!) is important as it indicates that the variable and the parameter associated with it have

not yet been specified.

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

y	price	[replaces none]
N levels to:	ij	[that is 2 levels]
Level 2 (j):	District	[j is higher level unit]
Level 1(i)	House	[i is lower level unit]
Done		

To specify the predictor to be a constant in the null random intercepts 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 district	[allows β_0 parameter to vary at level 2]
Tick	i house	[allows β_0 parameter to vary at level 1]
Done		

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 even more detail. You should now see the full algebraic specification of the model. Pressing - reduces the detail, clicking on the Zoom button allow the fonts size to be varied. You can copy this specification and paste into as graphic into a word-processor

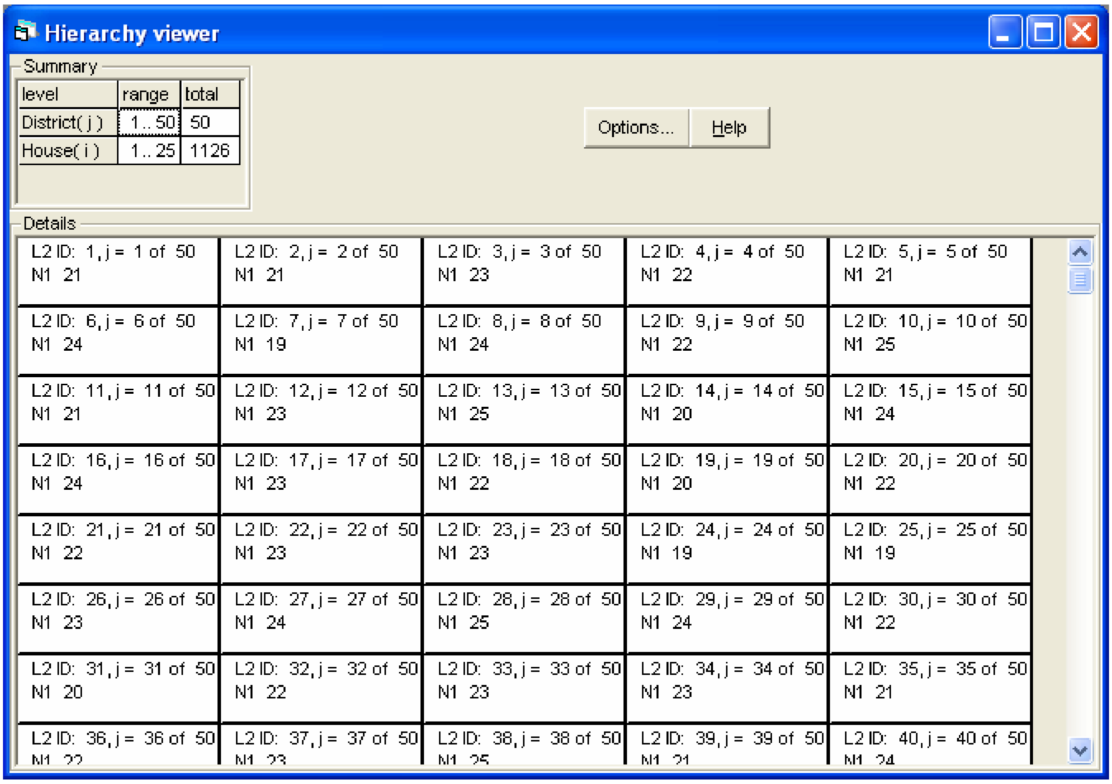
$$\mathrm{Price}_{ij} \sim \mathrm{N}(XB,\ \Omega)$$
$$\mathrm{Price}_{ij} = \beta_{0ij}\mathrm{cons}$$
$$\beta_{0ij} = \beta_0 + u_{0j} + e_{0ij}$$

$$\begin{bmatrix} u_{0j} \end{bmatrix} \sim \mathrm{N}(0,\ \Omega_u)\ :\ \Omega_u = \begin{bmatrix} \sigma^2_{u0} \end{bmatrix}$$

$$\begin{bmatrix} e_{0ij} \end{bmatrix} \sim \mathrm{N}(0,\ \Omega_e)\ :\ \Omega_e = \begin{bmatrix} \sigma^2_{e0} \end{bmatrix}$$

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

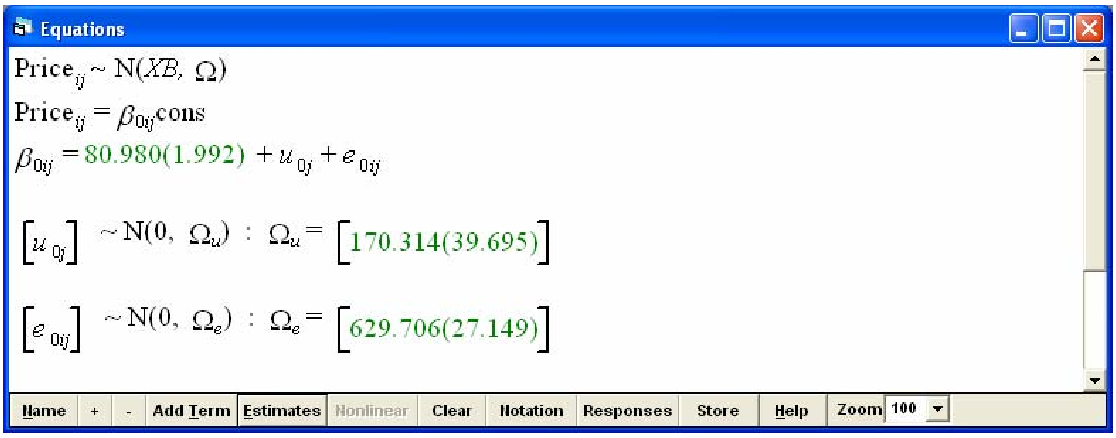
Model on main Menu
Hierarchy viewer



It is possible to see the summary of the number of houses in each and every higher-level district. Close the windows when you have examined the structure and it is as given here. Any problems are likely to be a result of incorrect sorting. Here there are 50 districts and they are numbered from 1 to 50, and there is a maximum of 25 houses in a single district.

Estimating the model

Before estimation begins, click on estimates in the lower tool bar twice. 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 district by district and the ‘gauge’ tanks are filled, 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 estimation procedure is being used: iterative generalized least squares. When all the estimates are green, the overall model has converged, and these are the estimates you want. (Unlike single-level models estimated by ordinary least squares; the multilevel model does not have a simple analytical exact solution; rather the IGLS algorithm performs an iterative sequence of Fixed-Random Fixed until a stable solution is reached.) For model 1, the following estimates are derived:



The terms in the Equations window represent parameter estimates with their estimated standard errors in brackets. We will discuss the log-likelihood later, 1126 out of 1126 cases in use means that there are no missing values in our data.

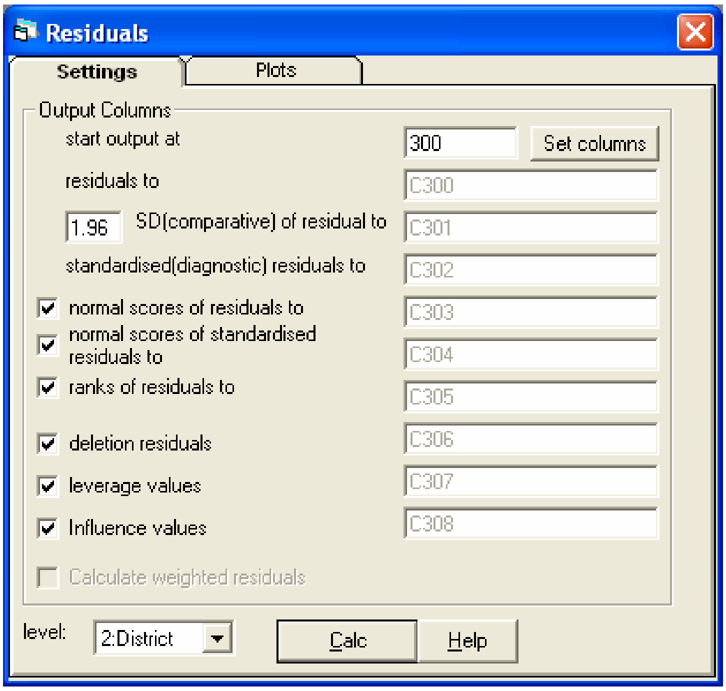
What does 80.98 represent?
And 170.3; and is it significantly different from zero?
And 629.7?
Does it appear that house prices vary between districts?

- 80.98 is the mean house price across all the districts and all the houses;
- 170.3 is the between district variance and as it is more than 2* the standard error, we can informally say that there is ‘significant’ between-district variance; we need a multilevel model to model these data adequetedly;
- 629.7 is the within district between house variation.

Estimating 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 average. The sequence is:

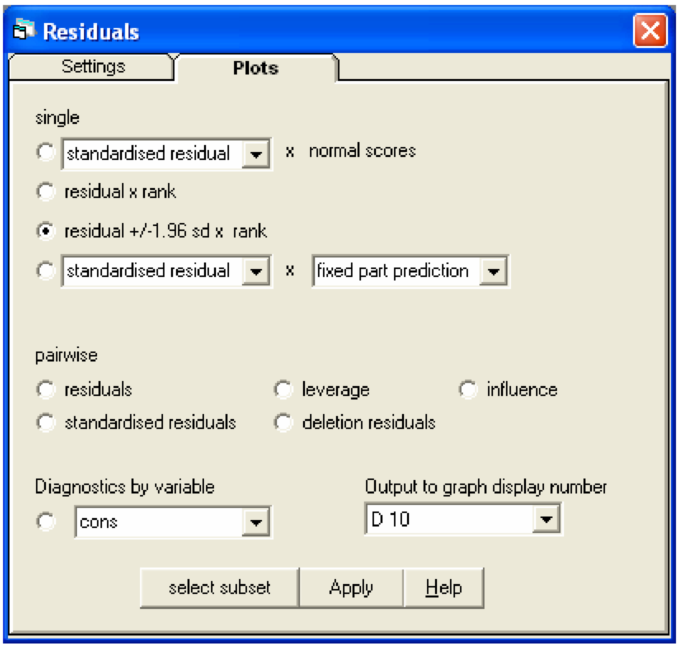
Model on Main Menu
Residuals
Change 1.0 to 1.96 SD (comparative) of residual to
Level 2 : district [replace 1 house]
Click Set Columns
Calculate

The completed screen should look like:



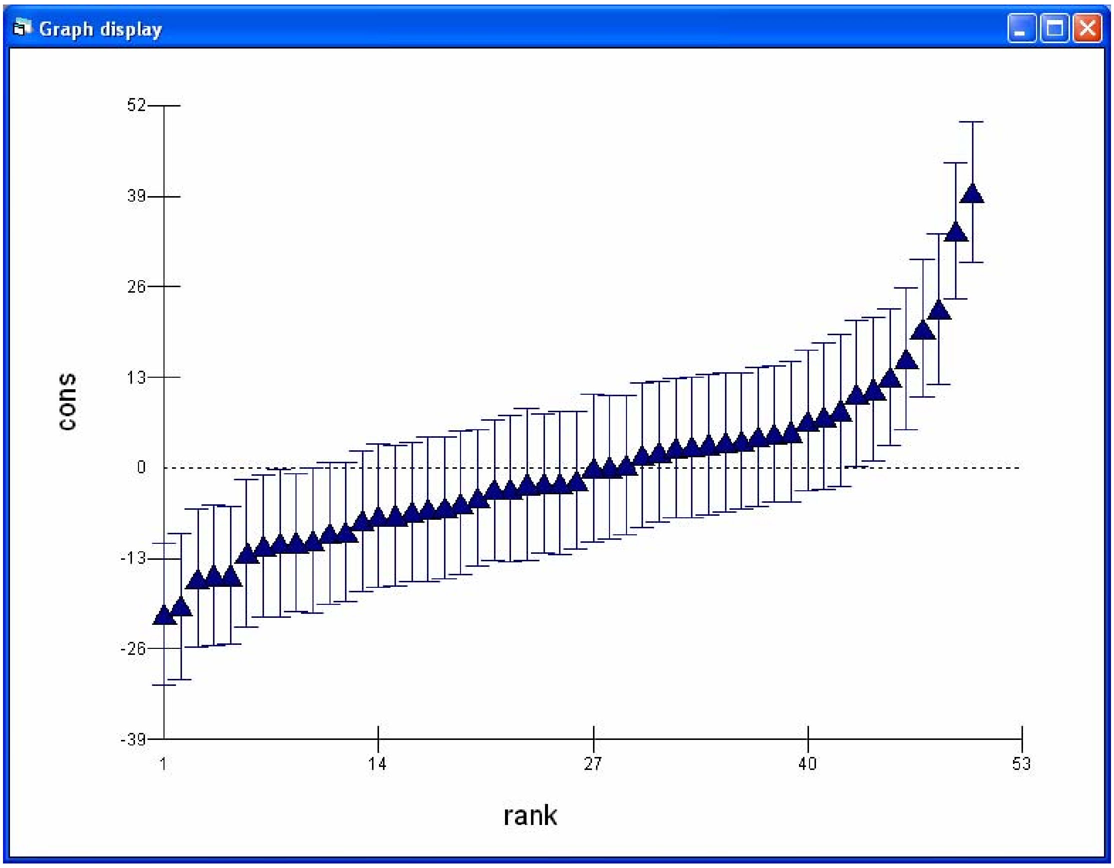
giving the columns where the requested values are to be stored; eg the residuals are in C300 and their ranks in c305. To view the values you can either use the View data window, or use the command interface to print them out.

Return to the residuals window and select the plots tab, and on the single pane at the top of the screen, select the 'residual +/- 1.96 SD x rank button and then Apply (Notice that D10 is the default graph display for this plot; ie the commands to execute the graph will be stored in Display 10.)



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

By clicking on the graph we can identify the cheapest and dearest districts.

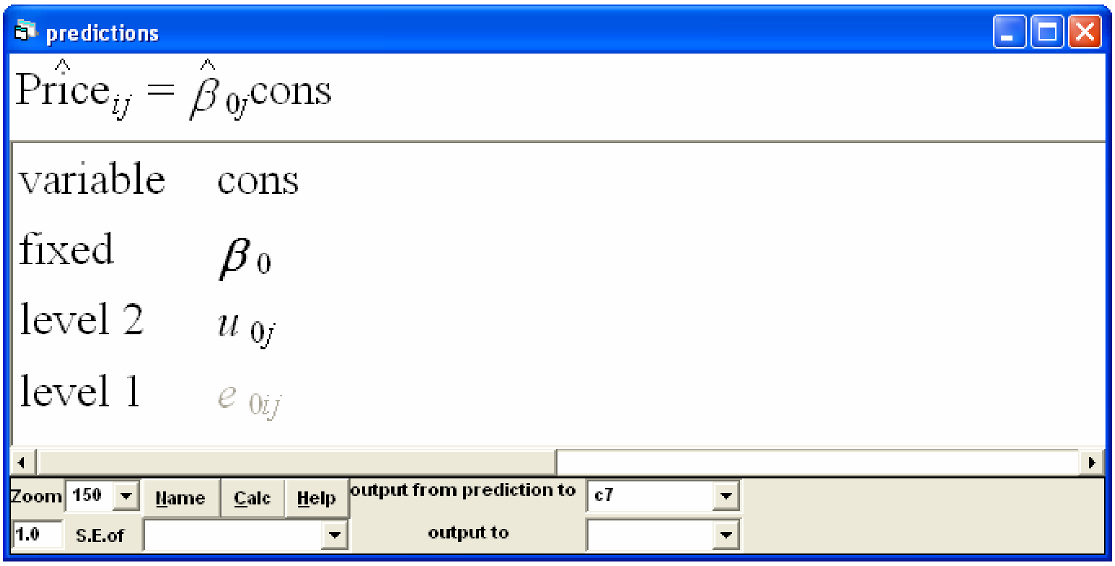


The dearest district is district 34 and houses cost some 39k more than generally across the city; the cheapest district is 21 and houses cost some 21k less than the all London average.

Making predictions and drawing varying relation plots
The next task is to make predictions of houses prices in each district and then to plot them in a customized 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 stored. Below is the completed screen to derive the predicted mean prices for each district; 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 predictions has been made.



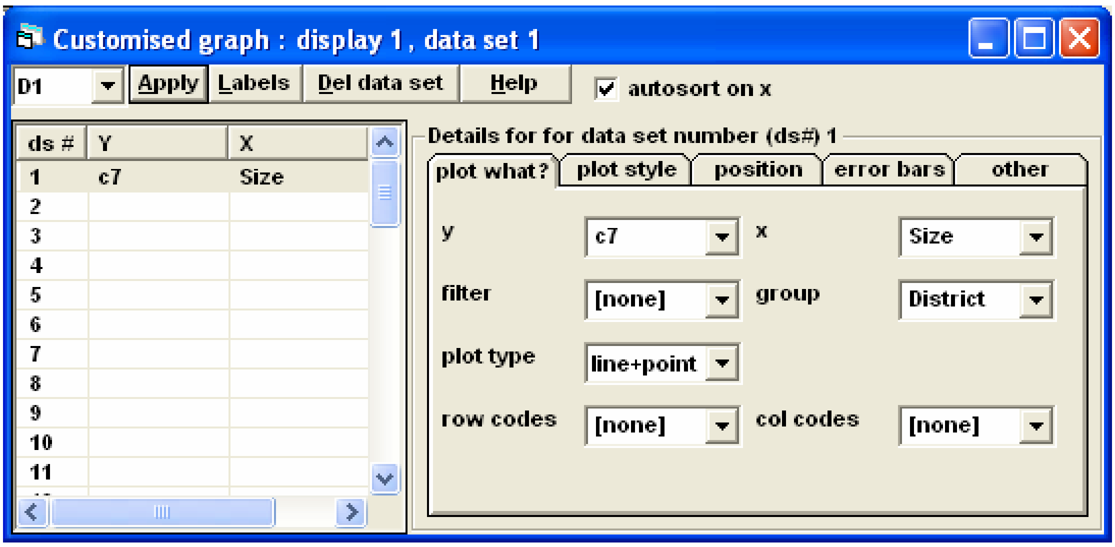
Next bring up the Customised graphics window

Graphs Main menu
Customised graphs

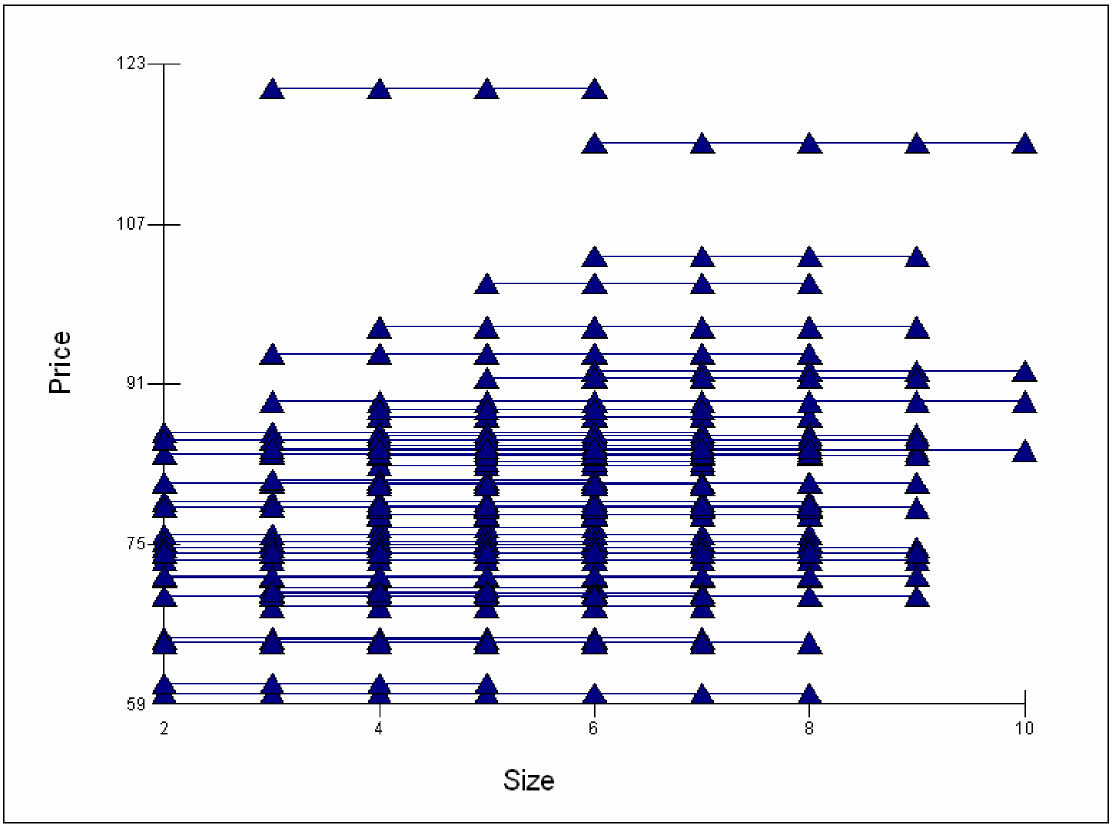
Currently the D10 graphic display is in operation as this was used to produce the caterpillar residual plot. Change this to D1

Choose	y is c7	
	x is size	[this is not yet in the model]
	Group is district	[to get a line of predictions for each district]
	Plot type is line and point	
	Apply	

The completed window is



The resultant graph after titles have been added and without the surrounding box is



Click the points to identify the two most expensive districts as districts 34 and 43.

That completes the first model, save the worksheet, model equations, graphs and estimates to a file called model1.ws, after giving the name Yhat1 to column 7. Close all windows except the Equations and Names window.

Model 2: 2-level random intercepts with a predictor centred on a specific value

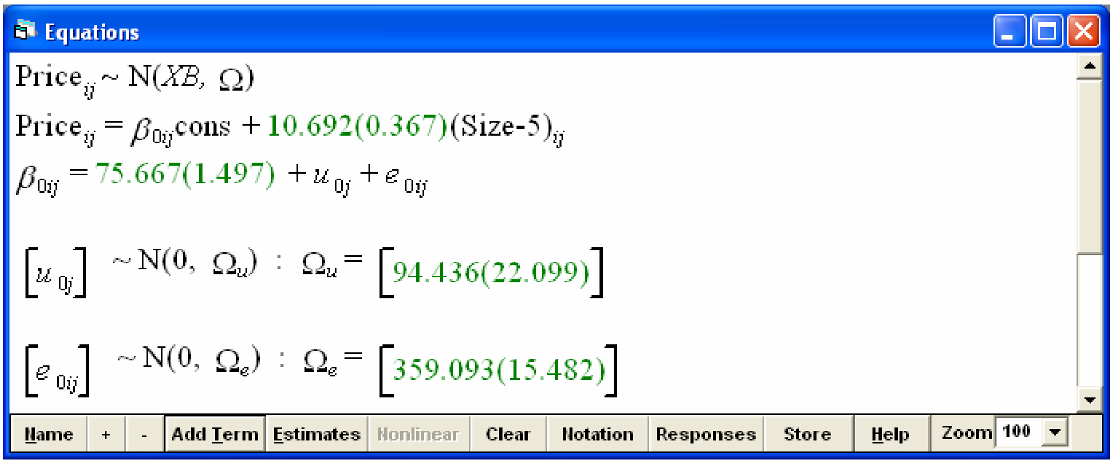
Specifying and estimating the model

To include the new variable in the fixed part of the model, click on Add Term on the bottom toolbar of the equations window

In the Specify term pop-up window

> Leave order to be 0 (this can be used to create 1st, end order etc interactions)
> Specify variable to be Size
> Because it is not a categorical variable you will be asked what should be done about centering
> Choose centering around the value 5 which is the median house size; this will give an interpretable intercept
> 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.



What do these values represent, and have the values in the random part altered from model 1?

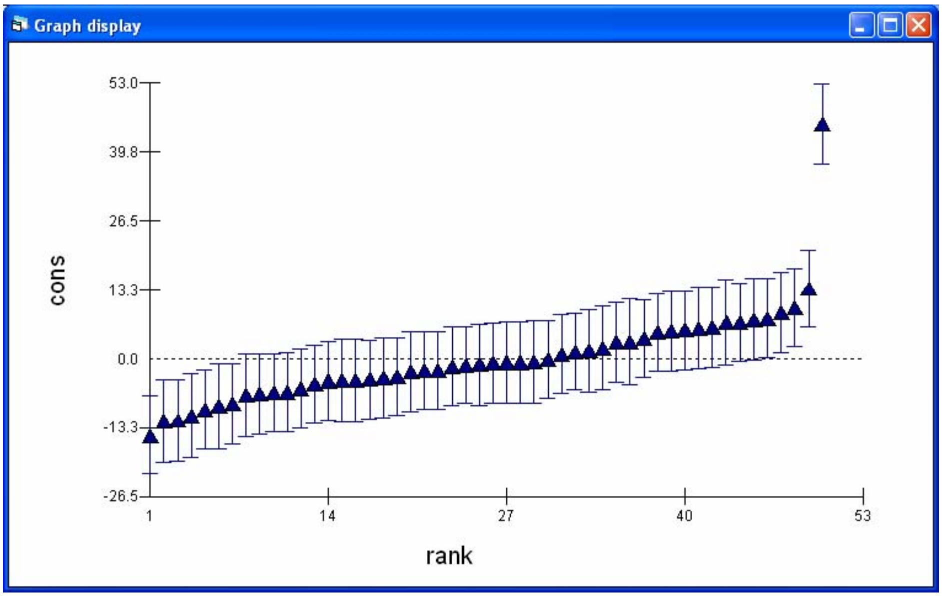
- 75.667 is the grand mean house prices for a 5-roomed house across all districts
- 10.692 is the grand mean slope, the cost of an additional room across all districts
- 94.436 is the between district variance; which although still significant has been substantially reduced; that must mean that size of houses varies between areas
- 359.093 is the within-district, between house variation; this has also decreased, there is a lot less unexplained between houses when account is taken of their size.

Calculating and graphiclevel-2 residuals

Model on Main Menu
Residuals
Start Output at C310 [not to overwrite existing]
Change 1.0 to 1.96 standard errors [to get 95% confidence bands]
Tick all types of residuals
Level 2: district
Set Columns [to get all output columns]
Calc [to estimate]

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

To get a caterpillar plot of the revised level-2 district residuals



Comparing the plot with last time there has been quite a lot of change, with one district now clearly differentiated from before. Use Identify points to verify that the outlying district is number 34. Why is it found to be so outlying (expensive) once size is taken into account? (Hint: think about house size in district 34.)

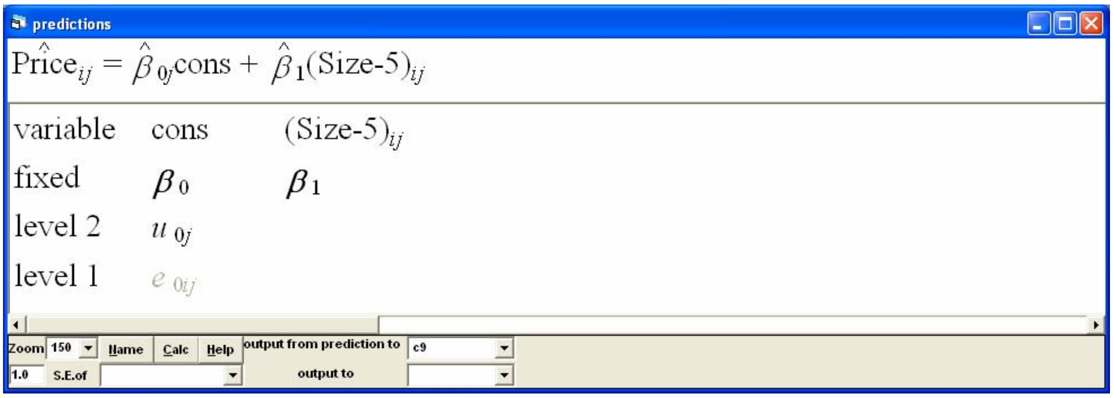
Predictions and varying relations plots

Model on Main Menu

Predictions

complete the window as follows putting the revised district estimates to c9

The residuals at level 1 **must** remain greyed out if you want to see the plot for districts



Graphics on Main Menu

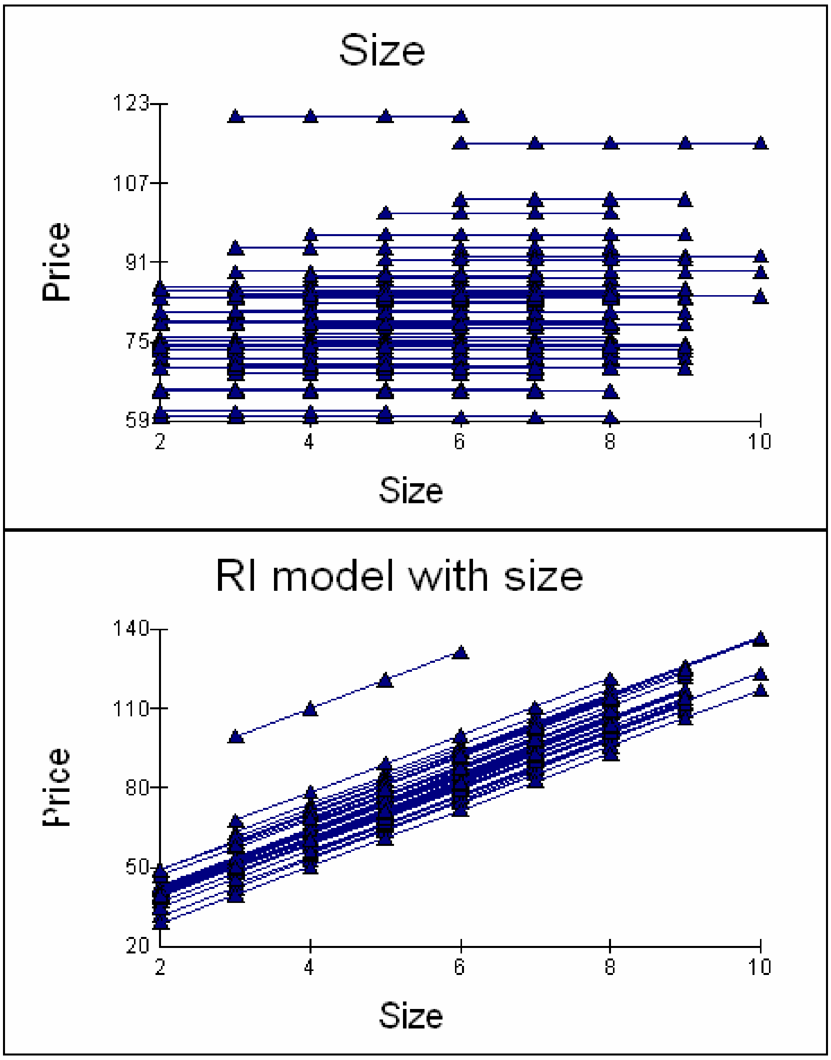
Customized graphics

Switch to D1	[display set D1]
Click on right-side to ds#2	[subgraph 2 not to overwrite ds#1]
Y c9	[revised predictions; not that Size-5 has been stored in the worksheet at col 8]

X size
group districts [to plot district lines]

Plot Position tab
choose col 1 and row 2 [original plot in col 1 row 1]
Apply

The Plot what screen should show that there are two subgraphs in display D1. The parallel lines assumption of the RI plot is clear.



We will come back to deal with the outlying district later.

Model 3: fully random model at level 2
Specifying and estimating a random-intercepts, random-slope model
Return to the equations window

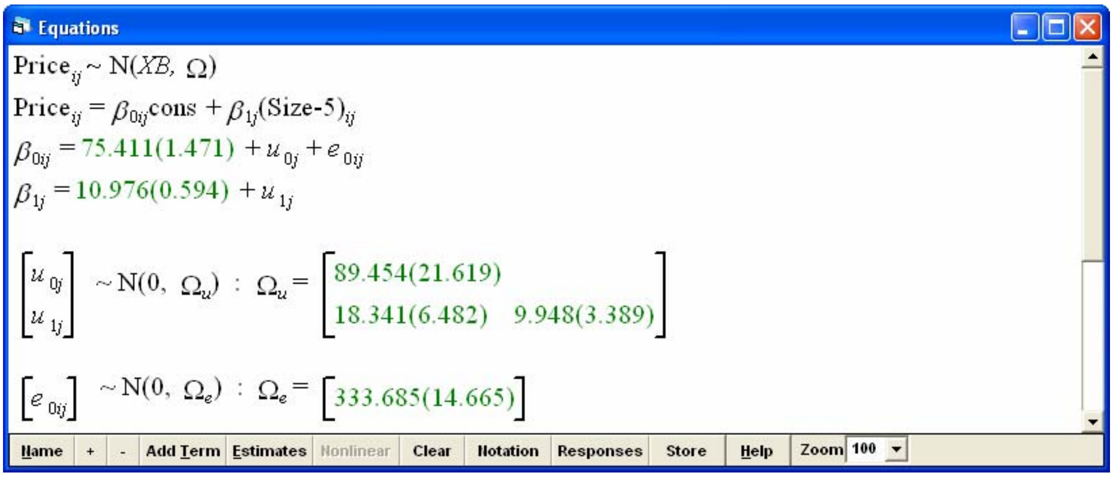
Click on Size-5 [to get X variable pop-up menu]

Tick District as well as fixed [to allow the associated slope parameter to vary over district]

Click Done [to close window]

Click More [continue estimation, blue to green]

Save revised model as model3.ws



There are now three terms at level 2 representing the variance-covariance for districts

- 89.454: there is significant between district-variance for 5 roomed house; the cost of a 5-romm house varies from place to place;
- 9.948: the variance for the slopes is also significant; while generally the cost of an extra room is 10.976, this varies from place to place
- 18.341: the covariance between the random intercepts and slopes is positive and significant; this means that districts which are expensive for a 5 room house will also have a steeper marginal relationship between price and size.

Calculating and graphic residuals

Model on Main Menu

Residuals

Start Output at C320 [not to overwrite existing]

Change 1.0 to 1.96 standard errors [to get 95% confidence bands]

Tick all types of residuals

Level 2: district

Set Columns [to get all output columns]

Calc [to estimate]

Return to Residuals window

Plot Tab

Click residuals +/- 1.96 SD x rank [on single plots pane]

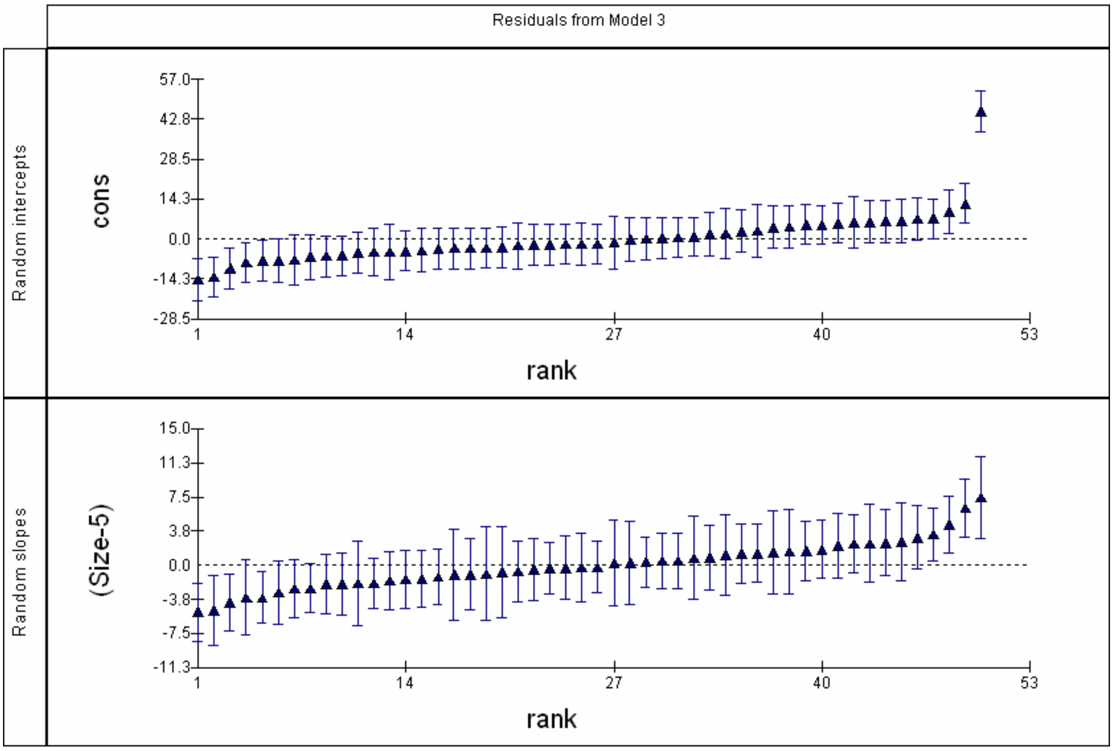
Apply [to get two plots in D10]

Two plots produced automatically.

Click in top graph

Titles tab

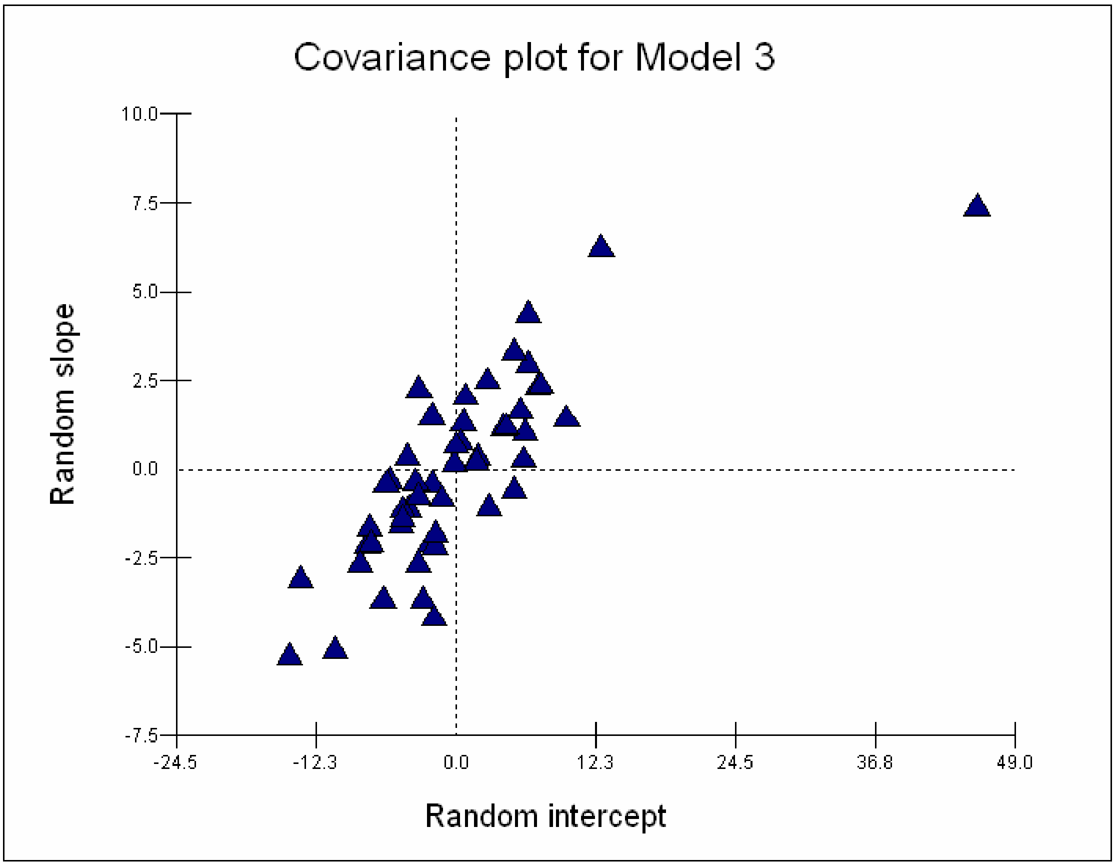
Margin (top) type Residuals from Model 3
Margin (left) type Random Intercepts
Tick box to show margins
Apply
Click in bottom graph
Titles tab
Margin (left) type Random Slopes
Tick box to show margins
Apply



Use Identify points to verify that the outlying district in terms of the random intercept is number 34,and that it is also the place with the steepest slope.

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

The positive covariance is very clear, as is the outlying nature of district 34.

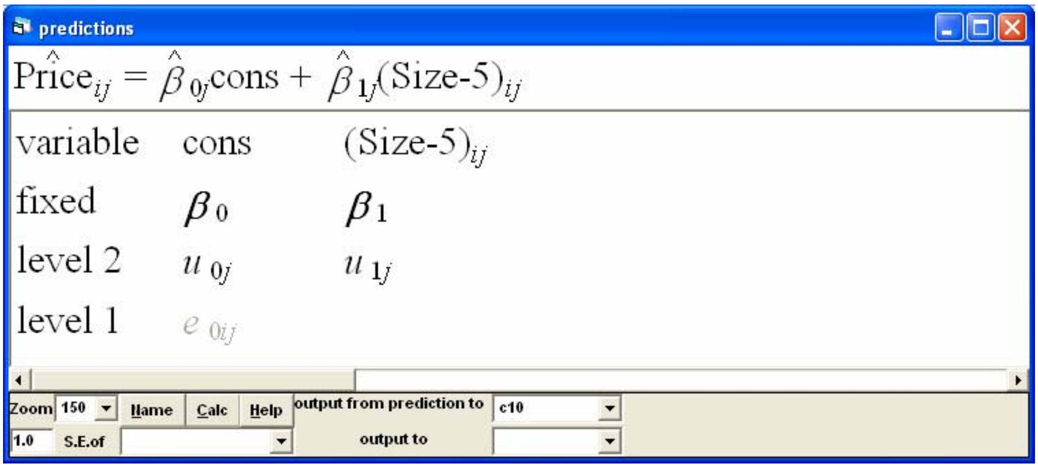


Predictions and varying relation plots

Model on Main Menu

Predictions

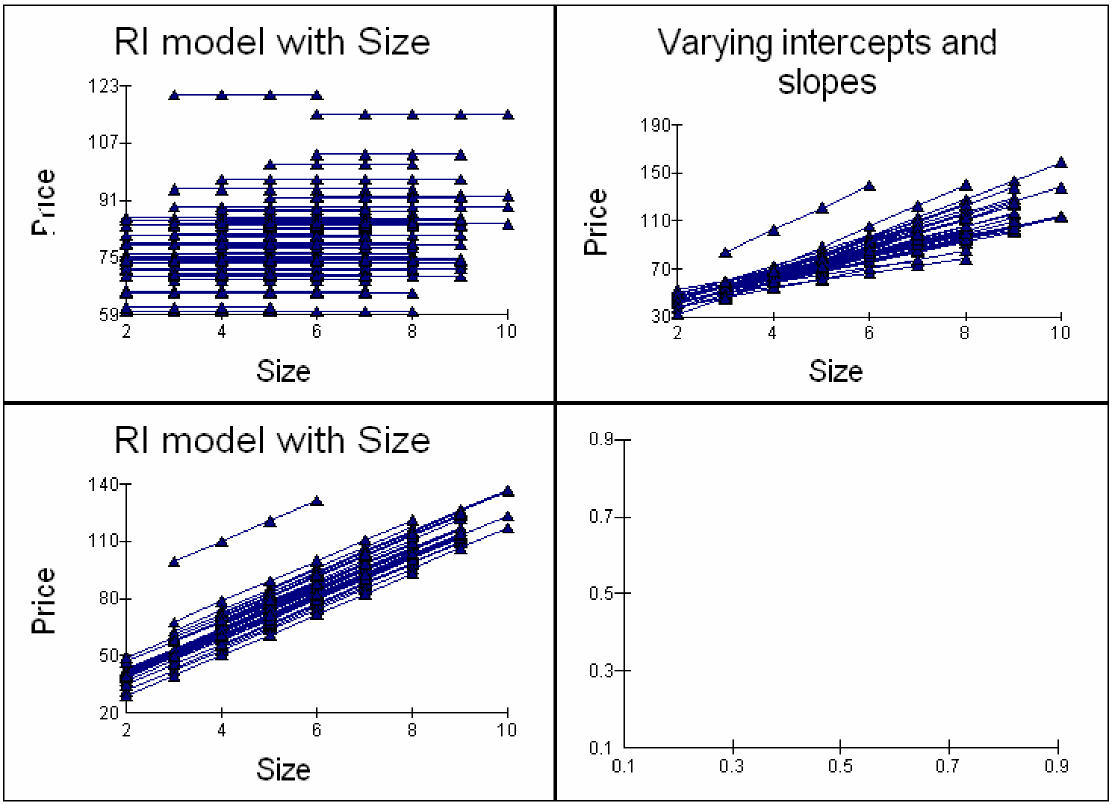
Click on Cons	[to get all terms associated with Constant included]
Click on Size-5	[to get all terms associated with Size-5 included]
Click on Level-1 residuals associated with Cons to **exclude**	
Output to c10	[free column]
	Calc



Name C10 as ‘Yhat3' and save the revised worksheet.

To get varying relation graph
- Graphs on Main Menu
  - Customized graphics
    - D1 [for graph display]
    - ds#3 [for third subgraph on display]
      - y: yhat3 [predicted values for each district]
      - x: Size
      - Plot type Line+point
      - Group District [to draw a line for each district]

Position tab
- Choose Column 2 and Row 1
- Apply



The fanning out associated with model 3 is clearly seen, there are bigger differences in price between districts for larger properties.

Model 4: Treating district as an outlier in the fixed part of the model
We now want to deal with district 34 as the marked outlier. We want to do this because it breaks the assumption that the district residuals follow a multivariate Normal distribution. We do so by including separate terms for district 34 in the fixed part of the model; it will automatically be removed from the level-2 random part.

Specifying and estimating the model
Click on the line for District 34 in the top; right-hand graph of the varying relations plot

Identify point
in Multilevel Filtering , highlight Level 2 district, idcode = 34
In model pane highlight Absorb in to dummy
Apply

In Absorb outliers into dummy variables pop-up menu

Tick interaction with Cons	[to get dummy with1 for District 34]
Tick interaction with Size-5	[to get interaction between dummy and Size-5]
Done	

This will create two new variables and include them in the model,. Click on each in turn and remove the centering around 5, (this is a current bug and workaround)

Return to the Equations window

More iterations

To get the estimated model as

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

$$\text{Price}_{ij} = \beta_{0ij}\text{cons} + \beta_{1j}(\text{Size-5})_{ij} + 54.373(6.401)\text{D\_District(34).cons}_j + 2.961(5.477)\text{D\_District(34).(Size-5)}_{ij}$$

$$\beta_{0ij} = 74.327(0.935) + u_{0j} + e_{0ij}$$

$$\beta_{1j} = 10.883(0.577) + u_{1j}$$

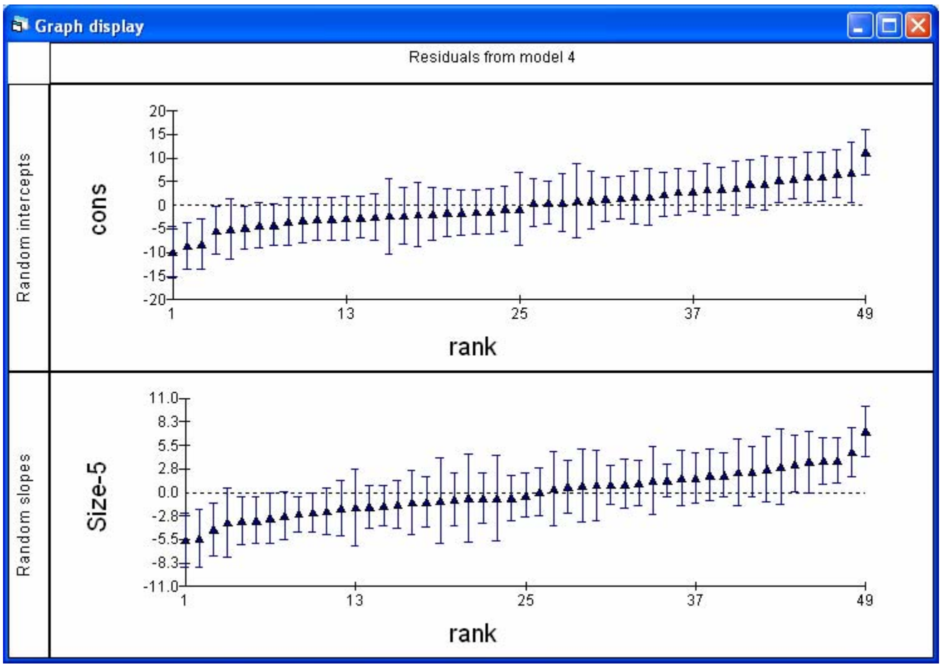
$$\begin{bmatrix} u_{0j} \\ u_{1j} \end{bmatrix} \sim \text{N}(0,\ \Omega_u) \ :\ \Omega_u = \begin{bmatrix} 25.256(8.512) & \\ 13.497(4.022) & 9.572(3.202) \end{bmatrix}$$

$$\begin{bmatrix} e_{0ij} \end{bmatrix} \sim \text{N}(0,\ \Omega_e) \ :\ \Omega_e = \begin{bmatrix} 333.630(14.630) \end{bmatrix}$$

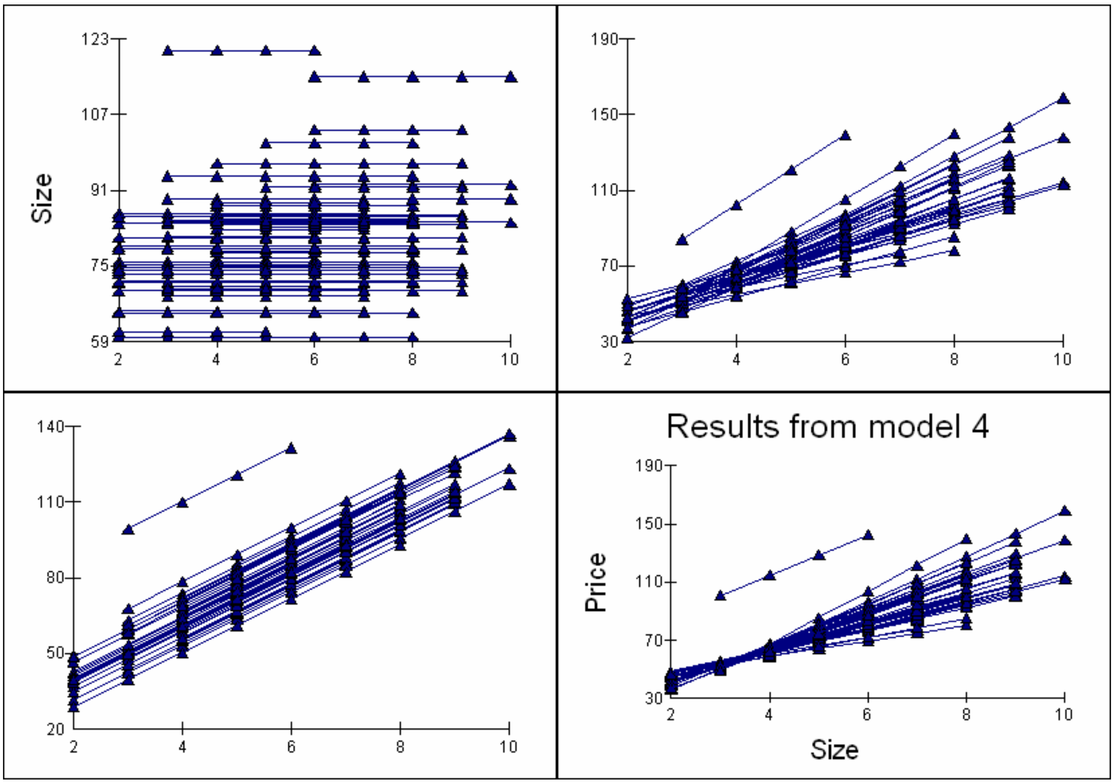
*-2*loglikelihood(IGLS Deviance)* = 9824.449(1126 of 1126 cases in use)

- 74.3 is the grand mean house prices for a 5-roomed house across all districts, except in district 34, where a 5 room house is 54.37 dearer (this difference is highly significant);
- 10.88 is the grand mean slope, the cost of an additional room across all districts, except in district 34, where an additional room is additionally 2.96 dearer (this difference is not very significant);
- 25.25 is the between district variance for a 5 roomed house; which although still significant has been substantially reduced now district 34 is not treated as part of the London distribution;
- 9.948: the variance for the slopes is also significant; this has not changed a great deal as the residual plot showed that district 34 while having the steepest slope was not an outlying value;
- 13.497: the covariance between the random intercepts and slopes is positive and significant; this means that districts which are expensive for a 5 room house will also have a steeper relationship between price and size
- 333.63 is the within-district, between house variation; this has hardly changed.

You should now be able to plot the residuals from this model and draw the varying relations plot.



There are now only 49 residual estimates as District 34 has been dummied out. There are also now no distinct outliers in the level-2 residuals



The distinctive nature of district 34 is seen as is the fanning out, so that the biggest differences in Price between districts are for larger properties.

Session 5: Logit modeling of proportions

Retrieve the data

File on main menu

Open worksheet

C:\talks\strirling\employ.ws

Name	Cn	n	missing	min	max	categorical
postcode	1	401	0	280101	310126	False
cell	2	401	0	1	4	False
gender	3	401	0	0	1	True
qualif	4	401	0	0	1	True
employed	5	401	0	0	20	False
total	6	401	0	1	27	False
adunemp	7	401	0	0	26.08	False
proportion	8	401	0	0	1	False
Code	9	401	0	1	4	True
constant	10	401	0	1	1	False

=

Note

Postcode is neighbourhood in Glasgow

Cell is element of the table for each postcode

Gender is male or female

Qualif is unqualified or qualified

Employed is count of number of employed teenagers in cell

Total is number of employed and unemployed teenagers in cell

Adunemp is adult unemployment in neighbourhood

Proportion is employed/total

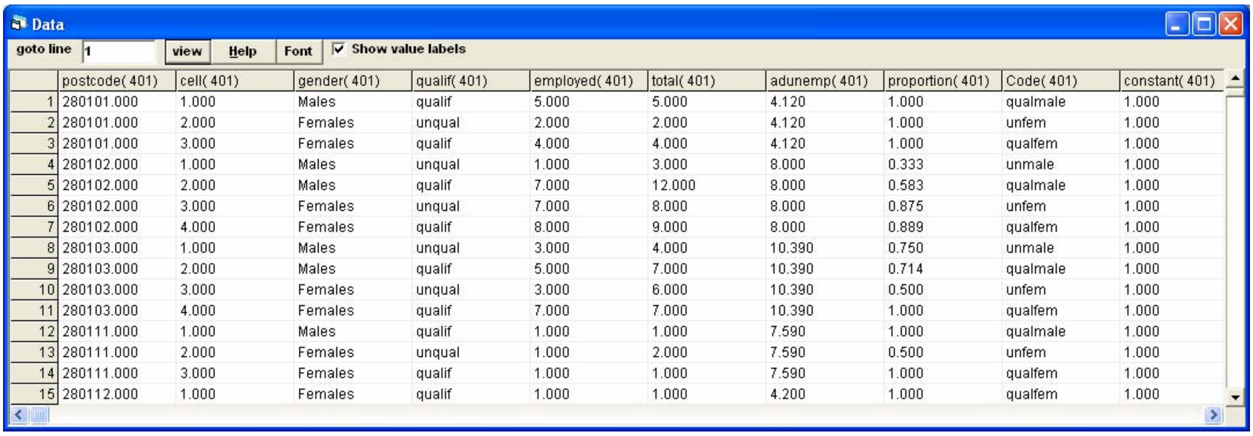
Code is categorical variable

1 = unqualified male   2 = unqualified females

3 = qualified males   4 = qualified females

Highlight the Names of the data; all variables

Press Data button



Data

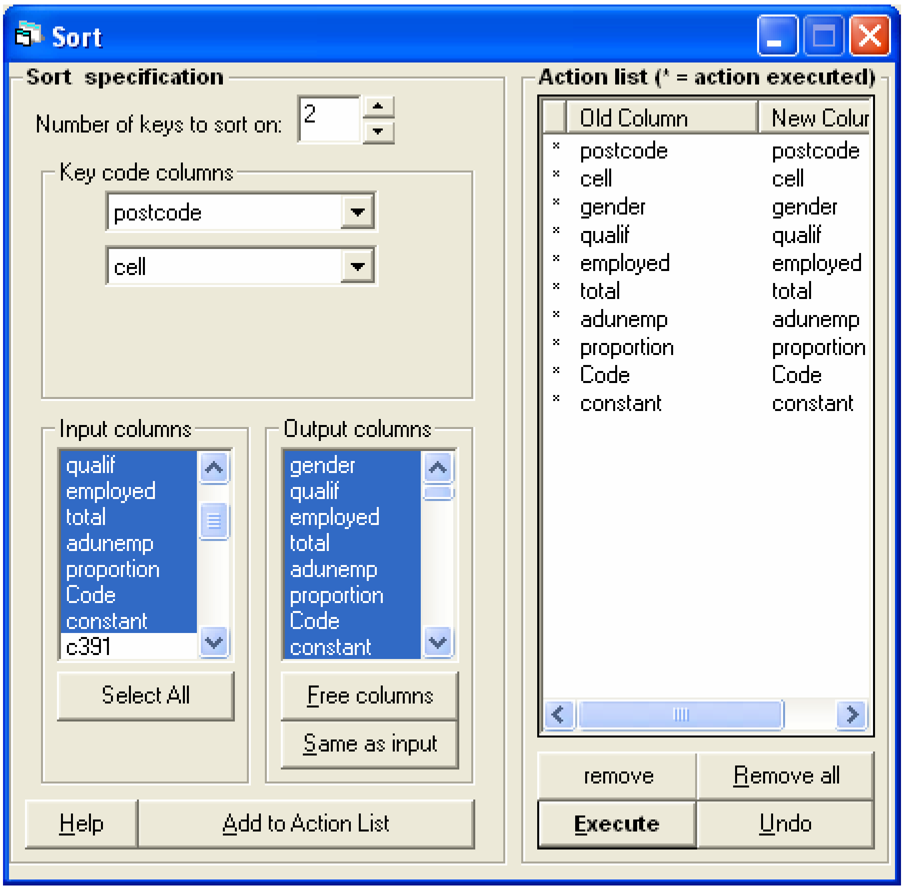
goto line 1 | view | Help | Font | Show value labels

	postcode(401)	cell(401)	gender(401)	qualif(401)	employed(401)	total(401)	adunemp(401)	proportion(401)	Code(401)	constant(401)
1	280101.000	1.000	Males	qualif	5.000	5.000	4.120	1.000	qualmale	1.000
2	280101.000	2.000	Females	unqual	2.000	2.000	4.120	1.000	unfem	1.000
3	280101.000	3.000	Females	qualif	4.000	4.000	4.120	1.000	qualfem	1.000
4	280102.000	1.000	Males	unqual	1.000	3.000	8.000	0.333	unmale	1.000
5	280102.000	2.000	Males	qualif	7.000	12.000	8.000	0.583	qualmale	1.000
6	280102.000	3.000	Females	unqual	7.000	8.000	8.000	0.875	unfem	1.000
7	280102.000	4.000	Females	qualif	8.000	9.000	8.000	0.889	qualfem	1.000
8	280103.000	1.000	Males	unqual	3.000	4.000	10.390	0.750	unmale	1.000
9	280103.000	2.000	Males	qualif	5.000	7.000	10.390	0.714	qualmale	1.000
10	280103.000	3.000	Females	unqual	3.000	6.000	10.390	0.500	unfem	1.000
11	280103.000	4.000	Females	qualif	7.000	7.000	10.390	1.000	qualfem	1.000
12	280111.000	1.000	Males	qualif	1.000	1.000	7.590	1.000	qualmale	1.000
13	280111.000	2.000	Females	unqual	1.000	2.000	7.590	0.500	unfem	1.000
14	280111.000	3.000	Females	qualif	1.000	1.000	7.590	1.000	qualfem	1.000
15	280112.000	1.000	Females	qualif	1.000	1.000	4.200	1.000	qualfem	1.000

Ensure data is sorted; cells within postcodes

Data Manipulation on main menu

Sort on Postcode and cell carry the rest and put back into original variables



Model 1: null random intercepts model

Model on main menu

Equations

Click on y and change to Proportion

Choose 2 levels

postcode as level 2

cell as level 1

Done

Click on N (for Normal theory model) and change to
Binomial distribution , then choose
Logit Link distribution

Click on red (ie unspecified) n_{ij} inside the Binomial brackets and
choose ‘total’ to be the binomial denominator (= number of trials)

Click on B_0 and choose the Constant, tick fixed effect; tick the j(postcode) to allow to vary over postcode (it is not allowed to vary at cell level, as we are assuming that all variation at this level is pure binomial variation)

At this point, the equations window should look like

$$\text{proportion}_{ij} \sim \text{Binomial}(\text{total}_{ij}, \pi_{ij})$$

$$\text{logit}(\pi_{ij}) = \beta_{0j}\text{constant}$$

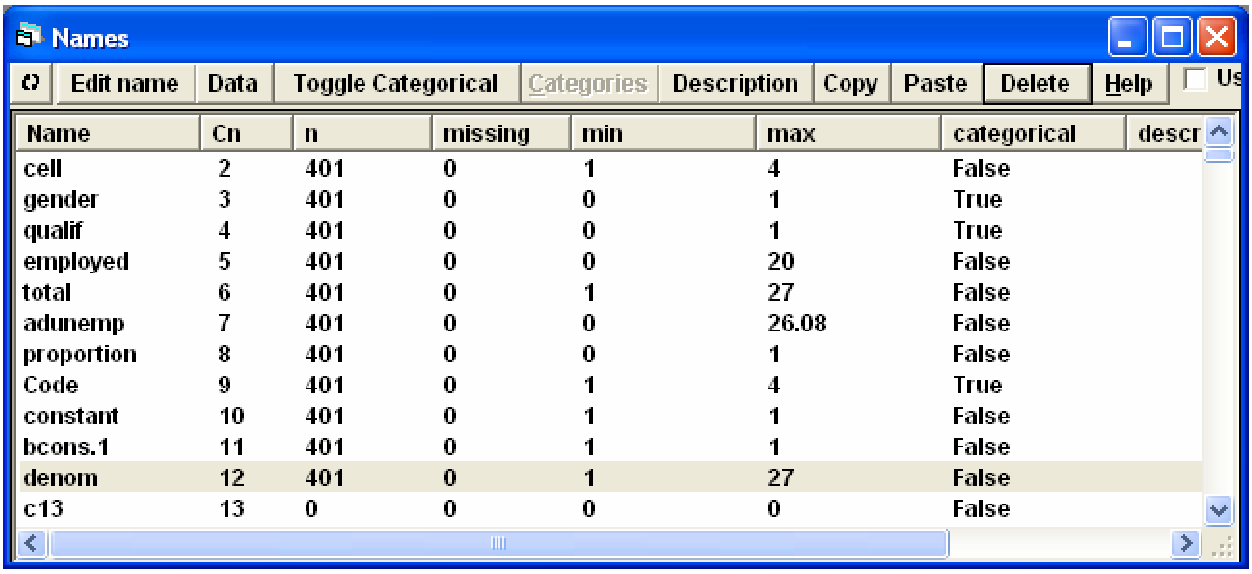
$$\beta_{0j} = \beta_0 + u_{0j}$$

$$\begin{bmatrix} u_{0j} \end{bmatrix} \sim \text{N}(0, \Omega_u) \ : \ \Omega_u = \begin{bmatrix} \sigma^2_{u0} \end{bmatrix}$$

$$\text{var}(\text{proportion}_{ij}|\pi_{ij}) = \pi_{ij}(1 - \pi_{ij})/\text{total}_{ij}$$

The variable ‘proportion employed’ in cell i of postcode j is specified to come from a Binomial distribution with and underlying probability, π_{ij}. The logit of the underlying probability is related to a Fixed effect, B_0 and an allowed to vary effect u_{0j} which as usual is assumed to come from a Normal distribution. The level 1, cell variation is assumed to be pure binomial variation in that it depends on the underlying probability and the total number of teenagers in a cell; it is not a parameter that has to be estimated.

It is worth looking at the worksheet as MLwiN will have created two variables in the background, Denom is the number of trials, that is Total in our case, while Bcons is a constant associated with the level 1 cell which is used in the calculation of the binomial weights; we can ignore this.

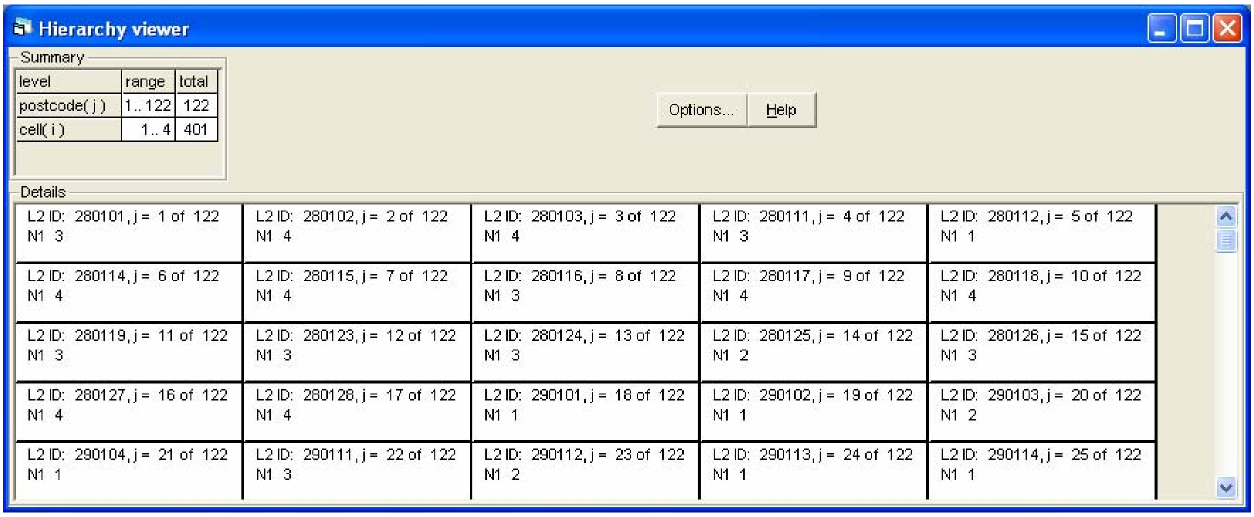


Name	Cn	n	missing	min	max	categorical	descr
cell	2	401	0	1	4	False	
gender	3	401	0	0	1	True	
qualif	4	401	0	0	1	True	
employed	5	401	0	0	20	False	
total	6	401	0	1	27	False	
adunemp	7	401	0	0	26.08	False	
proportion	8	401	0	0	1	False	
Code	9	401	0	1	4	True	
constant	10	401	0	1	1	False	
bcons.1	11	401	0	1	1	False	
denom	12	401	0	1	27	False	
c13	13	0	0	0	0	False	

Before estimating, it is important to check the hierarchy

Model on main menu

Hierarchy viewer



level	range	total
postcode(j)	1..122	122
cell(i)	1..4	401

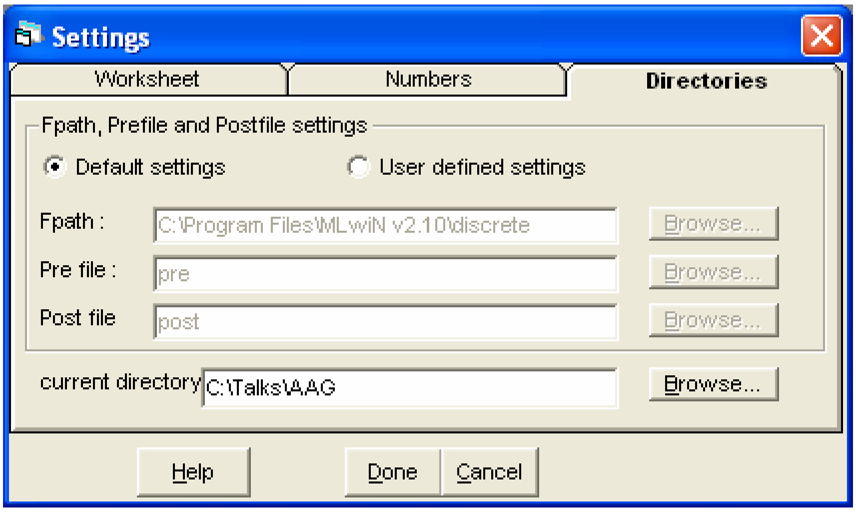
L2 ID: 280101, j = 1 of 122 N1 3	L2 ID: 280102, j = 2 of 122 N1 4	L2 ID: 280103, j = 3 of 122 N1 4	L2 ID: 280111, j = 4 of 122 N1 3	L2 ID: 280112, j = 5 of 122 N1 1
L2 ID: 280114, j = 6 of 122 N1 4	L2 ID: 280115, j = 7 of 122 N1 4	L2 ID: 280116, j = 8 of 122 N1 3	L2 ID: 280117, j = 9 of 122 N1 4	L2 ID: 280118, j = 10 of 122 N1 4
L2 ID: 280119, j = 11 of 122 N1 3	L2 ID: 280123, j = 12 of 122 N1 3	L2 ID: 280124, j = 13 of 122 N1 3	L2 ID: 280125, j = 14 of 122 N1 2	L2 ID: 280126, j = 15 of 122 N1 3
L2 ID: 280127, j = 16 of 122 N1 4	L2 ID: 280128, j = 17 of 122 N1 4	L2 ID: 290101, j = 18 of 122 N1 1	L2 ID: 290102, j = 19 of 122 N1 1	L2 ID: 290103, j = 20 of 122 N1 2
L2 ID: 290104, j = 21 of 122 N1 1	L2 ID: 290111, j = 22 of 122 N1 3	L2 ID: 290112, j = 23 of 122 N1 2	L2 ID: 290113, j = 24 of 122 N1 1	L2 ID: 290114, j = 25 of 122 N1 1

Question 1: Why the variability in the number of cells?

Before proceeding to estimation we can check location of non-linear macros for discrete data

Options on main menu

Directories



MLwiN creates a small file during estimation which has to be written temporarily to the current directory, this therefore has to be a place where files can be written; consequently you may have to change your current directory to something that can be written to.

After pressing start the model should converge to the following results, click on the lower Estimates button to see the numerical values

$$\text{proportion}_{ij} \sim \text{Binomial}(\text{total}_{ij}, \pi_{ij})$$
$$\text{logit}(\pi_{ij}) = \beta_{0j}\text{constant}$$
$$\beta_{0j} = 1.176(0.075) + u_{0j}$$

$$\left[u_{0j}\right] \sim \text{N}(0, \Omega_u) \;:\; \Omega_u = \left[0.270(0.079)\right]$$

$$\text{var}(\text{proportion}_{ij}|\pi_{ij}) = \pi_{ij}(1 - \pi_{ij})/\text{total}_{ij}$$

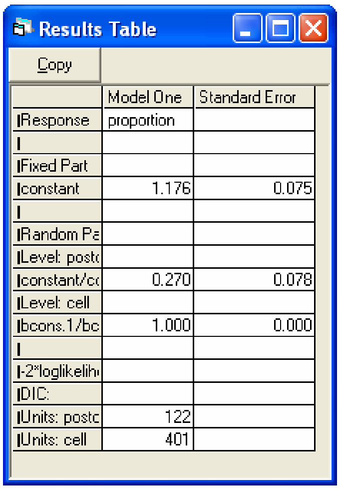
Question 2
Who is the constant?
What is 1.176 ?
What is 0.270 ?
Does the log-odds of teenage employment vary over the city?

We can store the estimates of this model as follows
Equations window
Click on Store model results
type in One in the pane
Ok

To see the results

Model on main menu
Compare stored models

This brings up the results in tabular form; these can be copied as a tab-delimited text file to the clipboard and pasted to Microsoft Word. Highlight the pasted text; Select Table, Insert, Table.

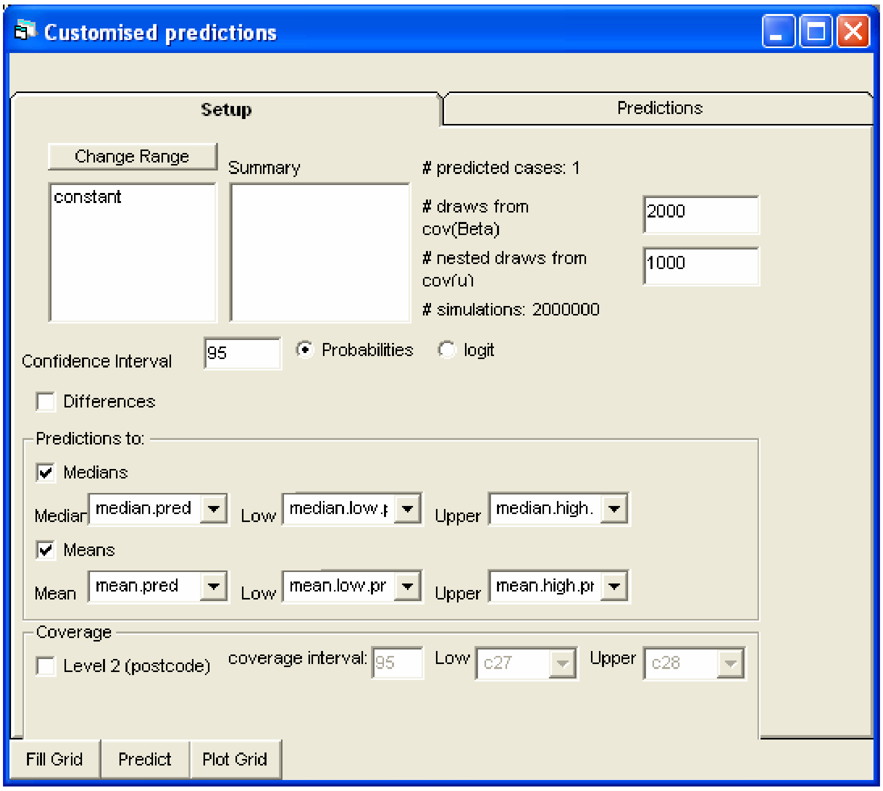


	Model One	Standard Error
Response	proportion	
Fixed Part		
constant	1.176	0.075
Random Pa		
Level: postc		
constant/cc	0.270	0.078
Level: cell		
bcons.1/bc	1.000	0.000
-2*loglikelih		
DIC:		
Units: postc	122	
Units: cell	401	

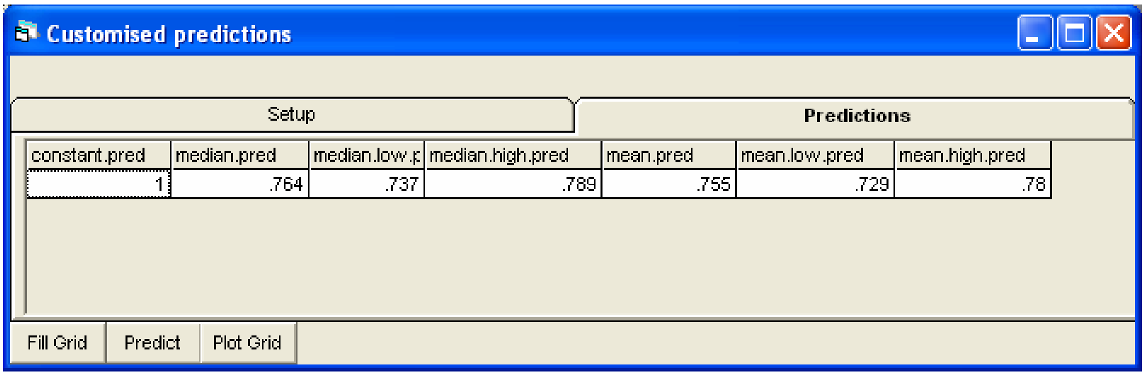
The log-odds are rather difficult to interpret, but we can change an estimate to a probability using the Customised predictions window:

Model on main menu
Customised predictions
In setup window
Confidence 95
Button on for Probabilities
Tick Medians
Tick Means
at bottom of pane: Fill grid
at bottom of pane: Predict
Switch toPredictions: all results have been stored in the worksheet.

The setup window should look like

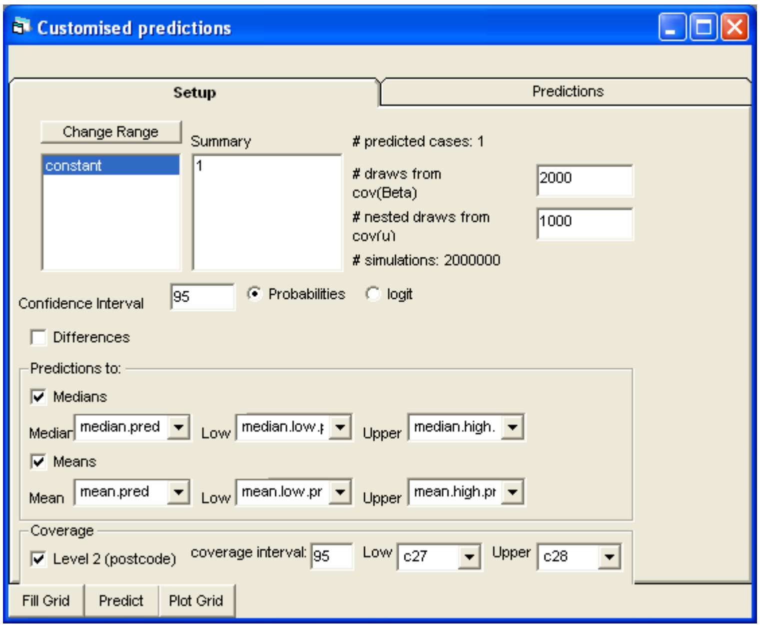


The predictions window should look like:

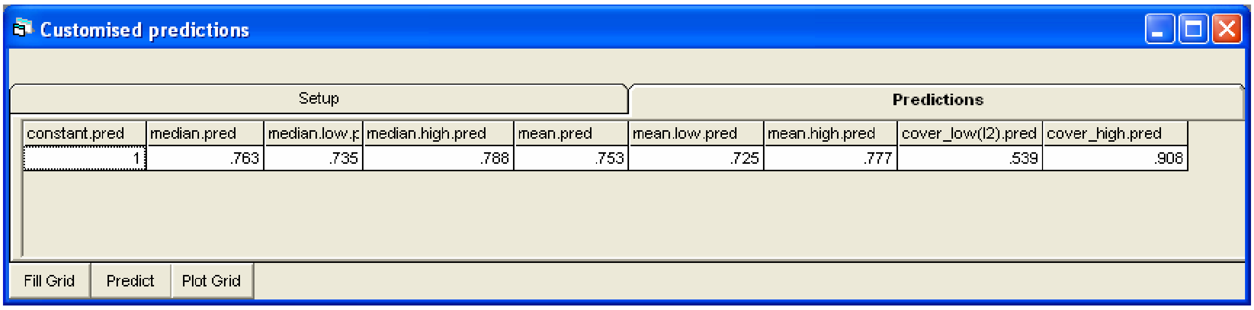


The cluster-specific estimated probability is given by the median of 0.764, with 95% confidence intervals of 0.737 and 0.789; while the population average values are very similar (0.755, CI: 0.73, 0.78) results. If we use Descriptive statistics on the main menu we find that the simple mean of the raw probabilities is 0.75.

Returning to the Setup window we can additionally tick for the coverage for level 2 postcodes and request the 95% coverage



The Predict window is now:



The estimated average teenage employment probability is 0.753, while the 95% coverage interval for Glasgow areas is between 0.539 and 0.908.

Returning to the equations window we can now distinguish between different types of teenagers

Model 2 with fixed part terms for qualifications and gender

Add term using Code with Unmale as the base or reference category, so that revised model after convergence is:

$$\text{proportion}_{ij} \sim \text{Binomial}(\text{total}_{ij}, \pi_{ij})$$
$$\text{logit}(\pi_{ij}) = \beta_{0j}\text{constant} + 0.149(0.148)\text{unfem}_{ij} + 0.996(0.149)\text{qualmale}_{ij} + 1.149(0.151)\text{qualfem}_{ij}$$
$$\beta_{0j} = 0.529(0.118) + u_{0j}$$

$$\left[u_{0j}\right] \sim \text{N}(0,\ \Omega_u)\ :\ \Omega_u = \left[0.237(0.075)\right]$$

$$\text{var}(\text{proportion}_{ij}|\pi_{ij}) = \pi_{ij}(1 - \pi_{ij})/\text{total}_{ij}$$

We can store the estimates of this model as Two using the Store button on the equations window

> Model on main menu
>
> > Compare stored models

This bring up the results in tabular form

Copy

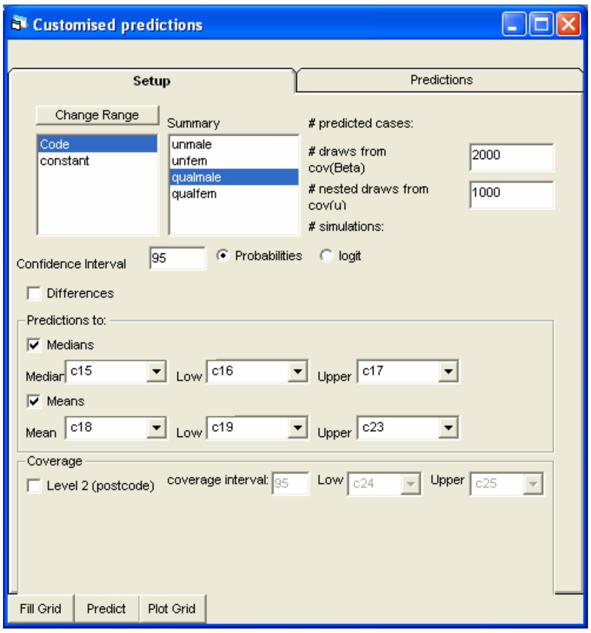
	Model One	Standard Error	Model Two	Standard Error
Response	proportion		proportion	
Fixed Part				
constant	1.176	0.075	0.529	0.118
unfem			0.149	0.148
qualmale			0.996	0.149
qualfem			1.149	0.151
Random Part				
Level: postcode				
constant/constai	0.270	0.078	0.237	0.075
Level: cell				
bcons.1/bcons.1	1.000	0.000	1.000	0.000
-2*loglikelihood:				
DIC:				
Units: postcode	122		122	
Units: cell	401		401	

We can now calculate the probability for all four types of teenager:

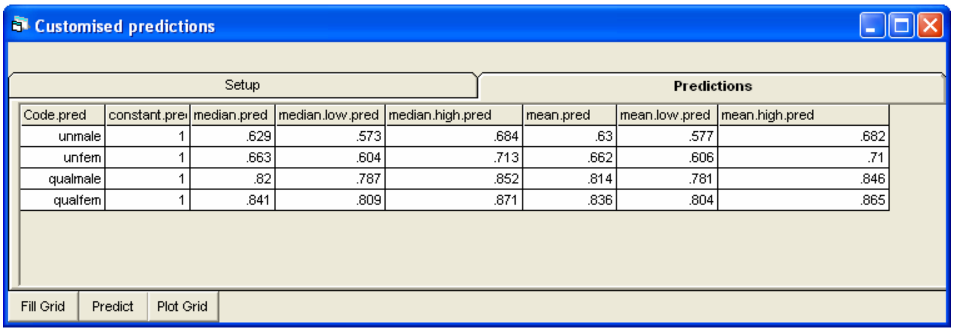
Model on main menu

- Customised predictions
  - In setup window
  - Clear (gets rid of previous choices)
    - Highlight Code and request Change Range
    - Tick all for each different type of teenager (unmale etc)
    - Confidence 95
    - Button on for Probabilities
    - Tick Medians
    - Tick Means
      - at bottom of pane: Fill grid
      - at bottom of pane: Predict
  - Predictions

The setup window is:



And the Predict window gives

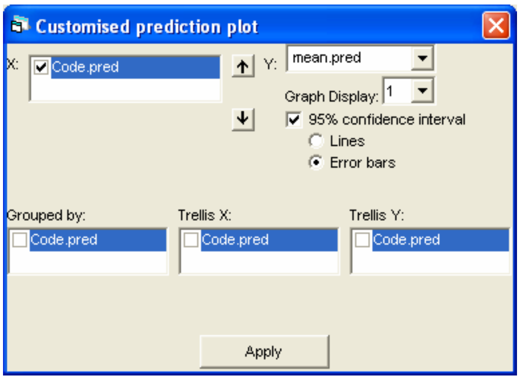


The values can be copied and pasted into Word to form a table

Code.pred	constant.	median.	median.low.	median.high.	mean.pred	mean.low.	mean.high.
unmale	1	0.6286546	0.57446176	0.68216538	0.61993444	0.56835616	0.67123282
unfem	1	0.66288924	0.61281633	0.71329087	0.65272975	0.60478234	0.70139945
qualmale	1	0.82063246	0.78834242	0.85116911	0.80792409	0.77535808	0.83902699
qualfem	1	0.84216172	0.81049794	0.87156969	0.82985991	0.79759902	0.86013967

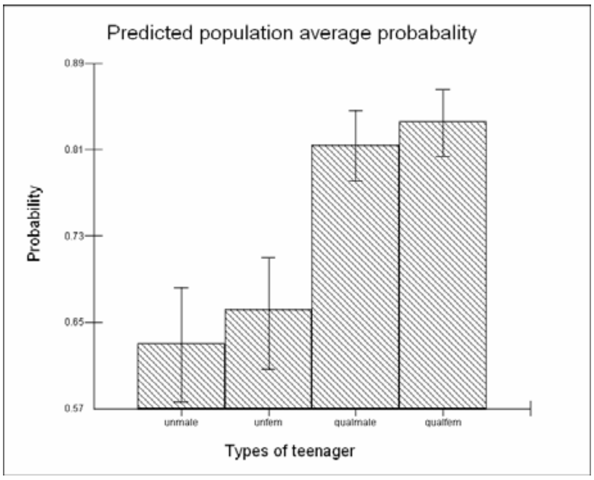
The higher employment is found for qualified teenagers, this is most easily seen by plotting the results

Customised predictions
  Plot Grid
    Y is Mean.pred, that is population averages
    Tick 95% confidence intervals
    Button error bars
    X variable: tick code.pred



Apply

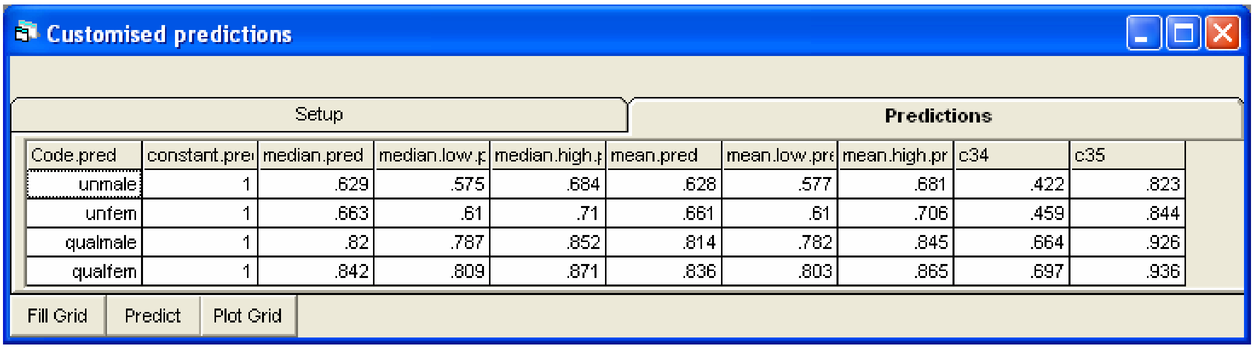
After some re-labeling of the graph we get (the plot is in customized windows D1)



The wider confidence bands for the unqualified reflect that there are fewer such teenagers.

Staying with this random-intercepts model, we can see the 95% coverage across Glasgow neighbourhoods for different types of teenagers:

Model on main menu
Customised predictions
In setup window
Tick coverage for postcode, and 95% coverage interval
Predict
Predictions



Customised predictions

Setup | Predictions

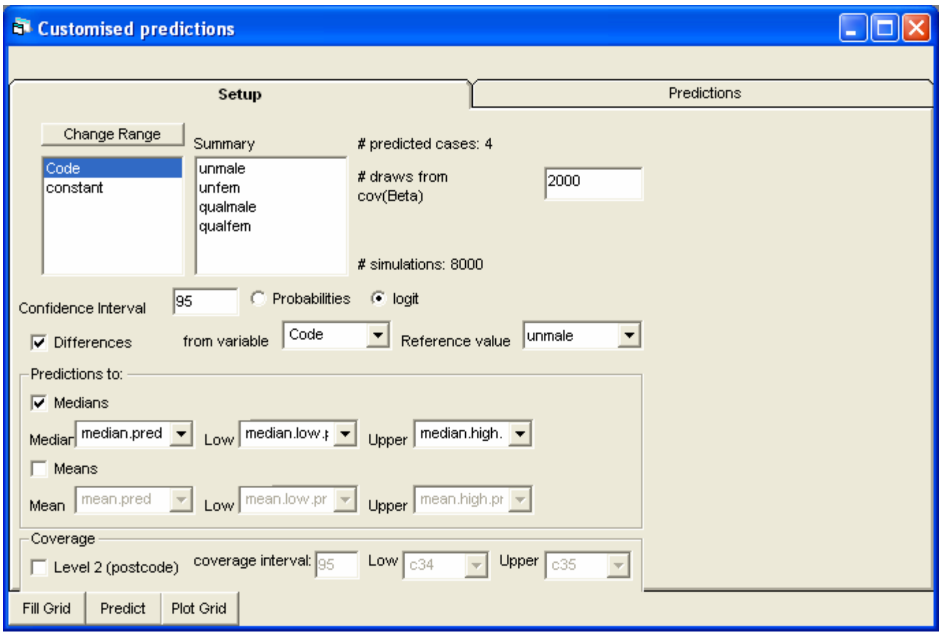
Code.pred	constant.pre	median.pred	median.low.p	median.high.p	mean.pred	mean.low.pre	mean.high.pr	c34	c35
unmale	1	.629	.575	.684	.628	.577	.681	.422	.823
unfem	1	.663	.61	.71	.661	.61	.706	.459	.844
qualmale	1	.82	.787	.852	.814	.782	.845	.664	.926
qualfem	1	.842	.809	.871	.836	.803	.865	.697	.936

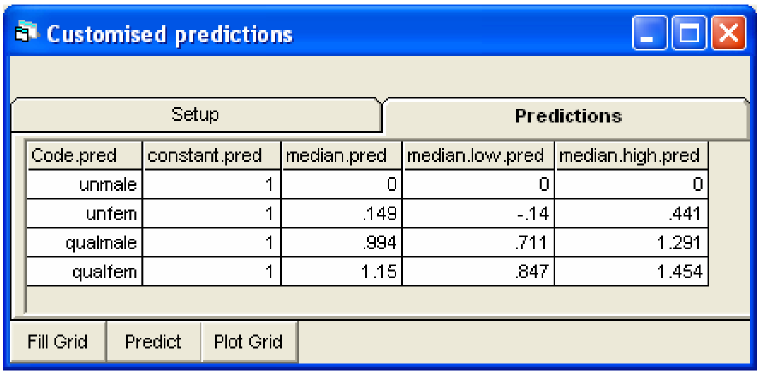
Fill Grid | Predict | Plot Grid

Across Glasgow the average probability of employment for unqualified males is estimated to be 0.628; in the 95% worst and best areas the probabilities are 0.422 and 0.823 respectively.

Sometimes it is preferred to interpret results from a logit model as relative odds, that is relative to some base or reference group. This can also be achieved in the customized predictions window. First we have to estimate differential logits by choosing a base category for our comparisons, and then we can exponentiate these values to get the relative odds of being employed. Here we choose unqualified males as the base category so that other teenagers will be compared to that group.

Customised predictions
    In setup window
    Button logit (instead of probabilities)
    Tick differences from variable Code, reference value Unmale
    Untick means
    Untick coverage
    Predict
Predictions




Customised predictions

Setup | Predictions

Code.pred	constant.pred	median.pred	median.low.pred	median.high.pred
unmale	1	0	0	0
unfem	1	.149	-.14	.441
qualmale	1	.994	.711	1.291
qualfem	1	1.15	.847	1.454

Fill Grid | Predict | Plot Grid

This gives the estimated differential cluster-specific logits which we can plot:

Plot Grid

Y is median.pred (not mean.pred)
X is code.pred
Tick 95% confidence interval
Button error bars

This will at first give the differential logits; to get odds we need to exponentiate the median and the 95% low and high values (from the Names window we see these are stored in c18-c20)

Data manipulation

Command interface

Expo c18-c20 c18-c20

After some re-labelling of the graph

In a relatively simple model with only one categorical predictor generating four main effects, we can achieve some of the above calculations by just using the Calculate command and the Expo and Alogit functions. Here are some illustrative results of doing this 'by hand':

Data manipulation

Command interface

calc b1 = 0.529 stores the logit unqualified male in a Box (that is a single value in comparison to a variate in a Column)

calc b2 = alogit b1 derives the cluster-specific probability: unqualified males

0.62925

calc b1 = 0.529+ 1.149 stores the logit for qualified female (base + differential)

1.6780

calc b2 = alogit b1 derives the c-s probability for qualified females

0.84264

To calculate the odds of being employed for any category compared to the base we simply exponeniate the differential logit (do not include the term associated with the constant)

calc b1 = 1.149 differential logit for qualified females

calc b2 = expo b1 odds for qualified females

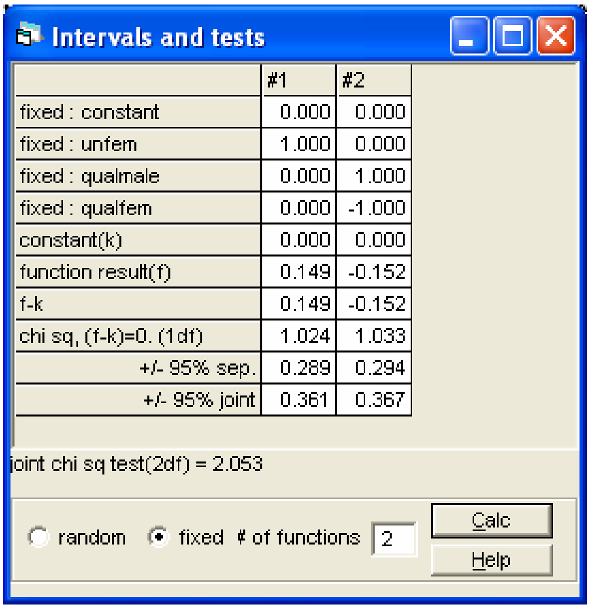
3.1550

The full table is as follows which agrees with minor rounding error with the simulated values

Who?	*Logit*	*Probability*	*Differential Logit*	*Odds*
Unqual Males	0.529	0.63	0	1*
Unqual Females	0.529 + 0.149 = 0.678	0.66	0.149	1.16
QualMale	.529 + 0.996 = 1.525	0.82	0.996	2.71
QualFemale	0.529 + 1.149 = 1.678	0.84	1.149	3.12

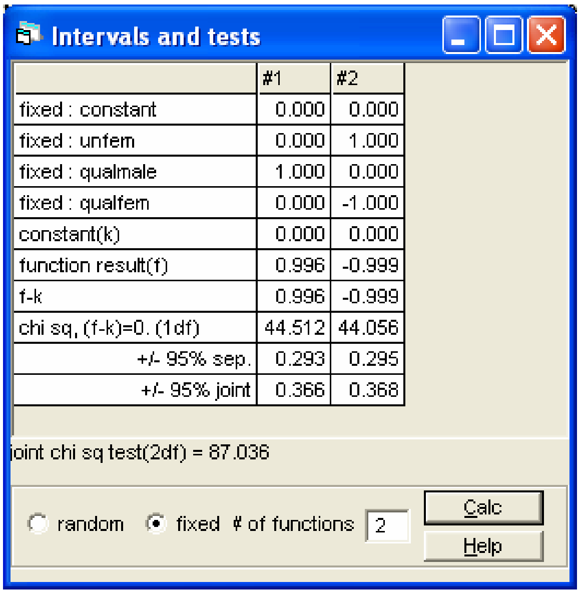
* the odds for the base category must always be 1

We can use the Intervals and tests window to test for the significance of difference between gender for qualified and unqualified teenagers. NB for unqualified teenagers it is given directly; for qualified it is not, and it has to be derived as a difference (note the -1)



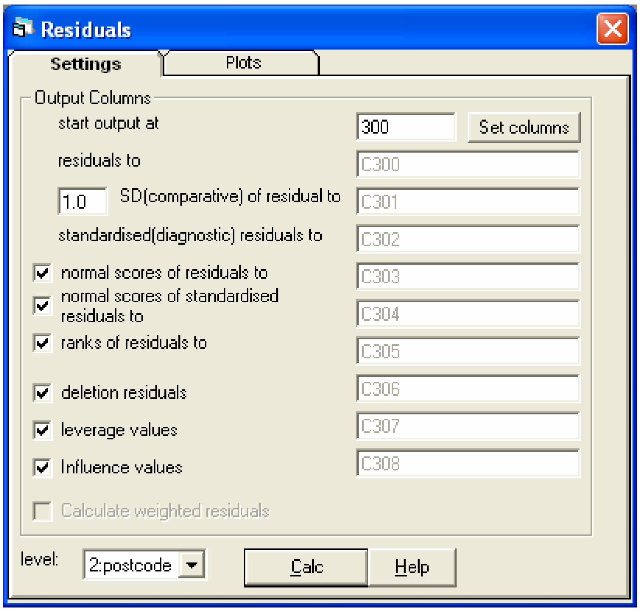
	#1	#2
fixed : constant	0.000	0.000
fixed : unfem	1.000	0.000
fixed : qualmale	0.000	1.000
fixed : qualfem	0.000	-1.000
constant(k)	0.000	0.000
function result(f)	0.149	-0.152
f-k	0.149	-0.152
chi sq, (f-k)=0. (1df)	1.024	1.033
+/- 95% sep.	0.289	0.294
+/- 95% joint	0.361	0.367

The chi-square statistics are all small; indicating that there is little difference between the genders. In contrast the differences between the levels of qualification for both males and females are highly significant



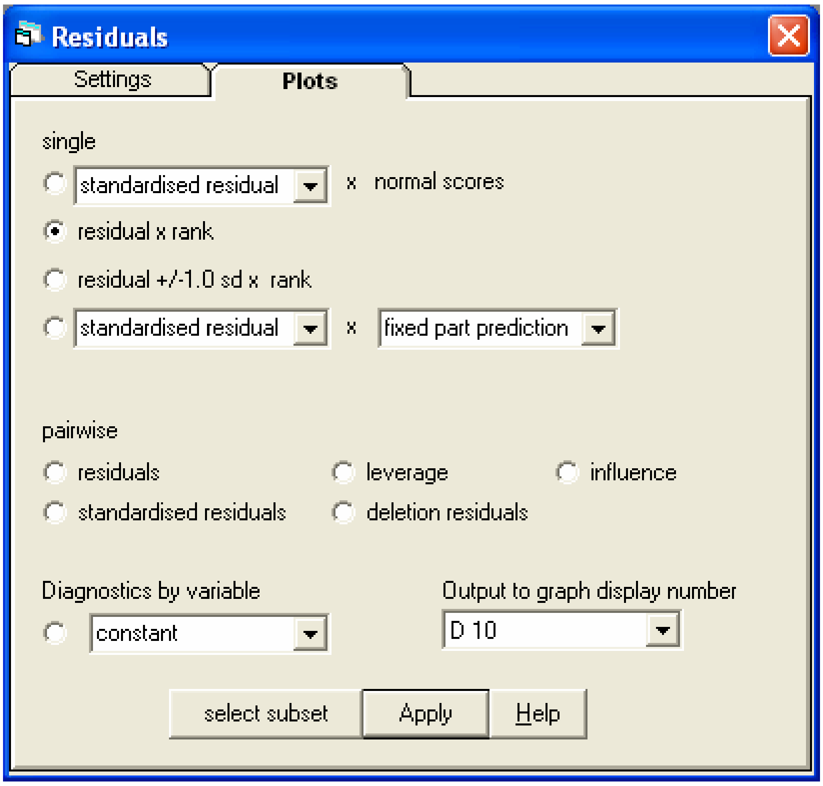
	#1	#2
fixed : constant	0.000	0.000
fixed : unfem	0.000	1.000
fixed : qualmale	1.000	0.000
fixed : qualfem	0.000	-1.000
constant(k)	0.000	0.000
function result(f)	0.996	-0.999
f-k	0.996	-0.999
chi sq, (f-k)=0. (1df)	44.512	44.056
+/- 95% sep.	0.293	0.295
+/- 95% joint	0.366	0.368

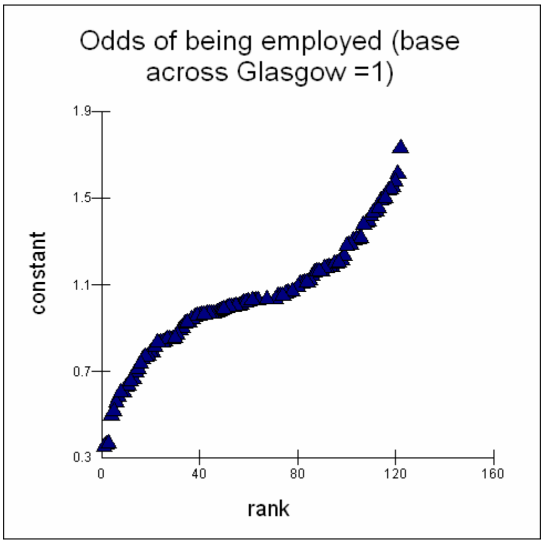
Turning now to the random effects, an effective way of presenting these is to calculate the odds of being employed against an all Glasgow average of 1. First calculate the level-2 residuals



and store in c300, then exponeniate these values (using the command interface)and plot them against their rank

Command interface
    Expo c300 c300





At the extremes some places only have 0.4 of the city wide odds, at the other extreme, the odds are increased by 1.8.

Model 2b: changing estimation

We have so far used the default non-linear options of mql, 1st order and exact binomial distribution; clicking on the non-linear button on the equations window we can change that to pql, 2nd order and allow extra-binomial variation, after more iterations the model converges to

$$\text{proportion}_{ij} \sim \text{Binomial}(\text{total}_{ij}, \pi_{ij})$$

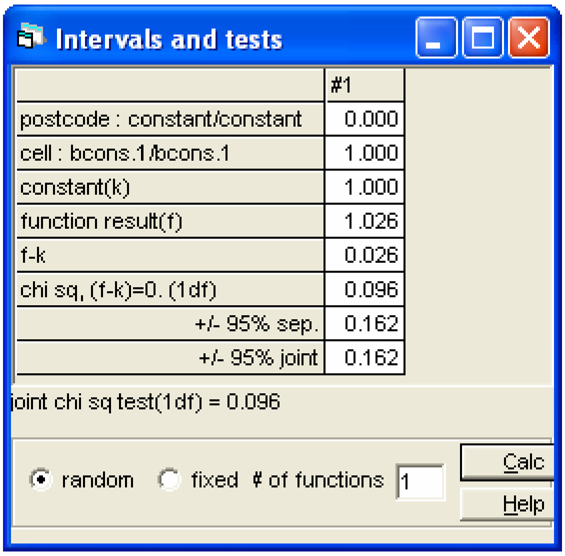
$$\text{logit}(\pi_{ij}) = \beta_{0j}\text{constant} + 0.157(0.150)\text{unfem}_{ij} + 1.044(0.153)\text{qualmale}_{ij} + 1.205(0.154)\text{qualfem}_{ij}$$

$$\beta_{0j} = 0.556(0.121) + u_{0j}$$

$$\left[u_{0j}\right] \sim \text{N}(0, \Omega_u) : \Omega_u = \left[0.240(0.081)\right]$$

$$\text{var}(\text{proportion}_{ij}|\pi_{ij}) = 1.025(0.083)\pi_{ij}(1 - \pi_{ij})/\text{total}_{ij}$$

Question 3:
Have the results changed a great deal?
Is there significant over-dispersion for the extra-binomial variation?



Note that we have tested the over-dispersion parameter (associated with the binomial weight bcons) against 1, and that there is no significant overdispersion as shown by the very low chi-square value. Use the non-linear button to set the distributional assumption back to an exact Binomial.

Model 3: modelling the cross-level interaction between gender, qualifications and adult unemployment

To estimate the effects of adult unemployment on teenage employment, add term to the model and choose to centre this variable around a mean of 8% which is the rounded, across-Glasgow average. This gives the main effect for adult unemployment. We want to see whether this interacts with the individual characteristics of qualification and gender.

In equations window

Order 1	first order interactions
Code	choose unmale as base
Adunemp	the continuous variable (the software takes account of centering)

After more iterations to convergence the results are:

$$\text{proportion}_{ij} \sim \text{Binomial}(\text{total}_{ij}, \pi_{ij})$$

$$\text{logit}(\pi_{ij}) = \beta_{0j}\text{constant} + 0.048(0.168)\text{unfem} + 0.866(0.160)\text{qualmale}_{ij} + 1.078(0.165)\text{qualfem}_{ij} + -0.111(0.025)(\text{adunemp-8})_j + 0.054(0.030)\text{unfem.(adunemp-8)}_{ij} + 0.071(0.033)\text{qualmale.(adunemp-8)}_{ij} + 0.028(0.033)\text{qualfem.(adunemp-8)}_{ij}$$

$$\beta_{0j} = 0.705(0.127) + u_{0j}$$

$$\begin{bmatrix} u_{0j} \end{bmatrix} \sim \text{N}(0,\ \Omega_u) \ : \ \Omega_u = \begin{bmatrix} 0.153(0.062) \end{bmatrix}$$

$$\text{var}(\text{proportion}_{ij}|\pi_{ij}) = \pi_{ij}(1 - \pi_{ij})/\text{total}_{ij}$$

The interactions have been created, labeled and placed in the model.

Store the model as three
    Mstore "three"

This bring up the results in tabular form

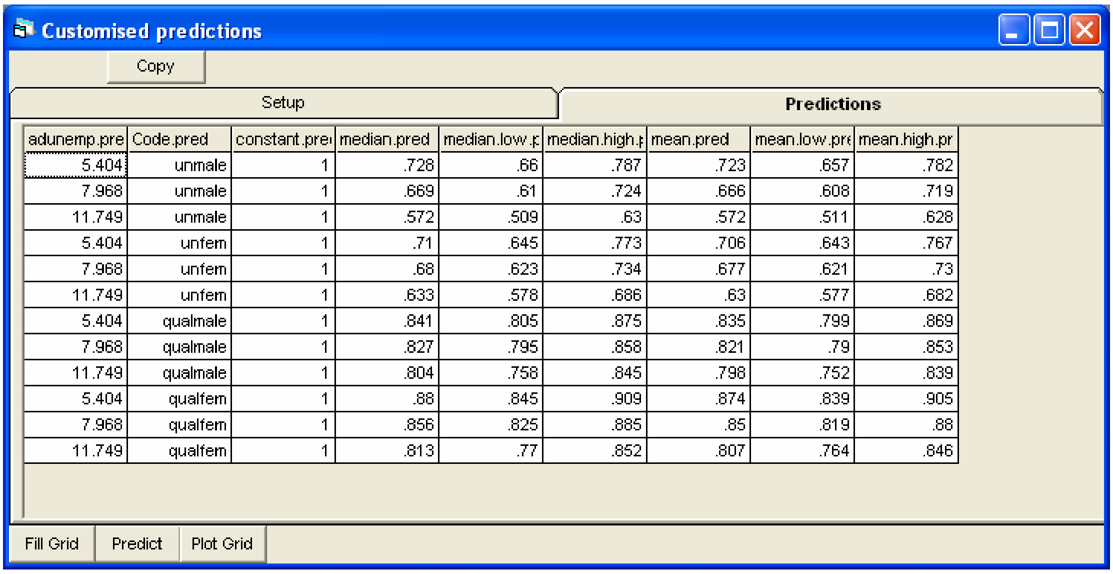
	Model One	**Standard Error**	**Model Two**	**Standard Error**	**Model Three**	**Standard Error**
Response	**proportion**		**proportion**		**proportion**	
Fixed Part						
constant	1.176	0.075	0.529	0.118	0.705	0.127
unfem			0.149	0.148	0.048	0.168
qualmale			0.996	0.149	0.866	0.160
qualfem			1.149	0.151	1.078	0.165
(adunemp-8)					-0.111	0.025
unfem.(adunemp-8)					0.054	0.030
qualmale.(adunemp-8)					0.071	0.033
qualfem.(adunemp-8)					0.028	0.033
Random Part						
Level: postcode						
constant/constant	0.270	0.079	0.237	0.075	0.153	0.062
Level: cell						
bcons.1/bcons.1	1.000	0.000	1.000	0.000	1.000	0.000
-2*loglikelihood:						
DIC:						
Units: postcode	122		122		122	
Units: cell	401		401		401	

The results are most perhaps most easily appreciated as the probability of being employed in a cross-level interaction plot (adunemp is a level 2 variable; code is a level-1 one variable)

Model on main menu
    Customised predictions (this automatically takes account of interactions)
        In setup window
        Clear (gets rid of previous choices)
        Highlight Adunemp and request Change Range
        Nested means; level of nesting 1 (repeated calc of means to get 3 characteristic values of the un-centred variable)
        Highlight Code and request Change Range
        Tick all boxes for each different type of teenager (unmale etc)
        Confidence 95
        Button on for Probabilities

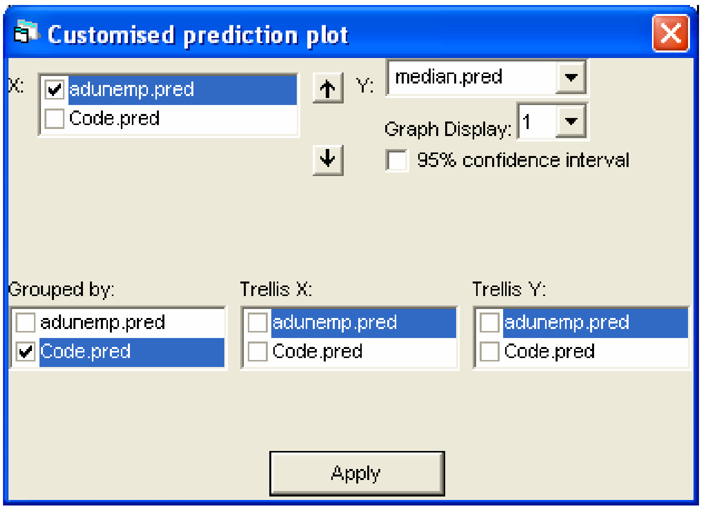
Tick Medians
Tick Means
at bottom of pane: Fill grid
at bottom of pane: Predict
Predictions

The predictions are for 12 rows (4 types of teenager for each of 3 characteristic values of adult unemployment):

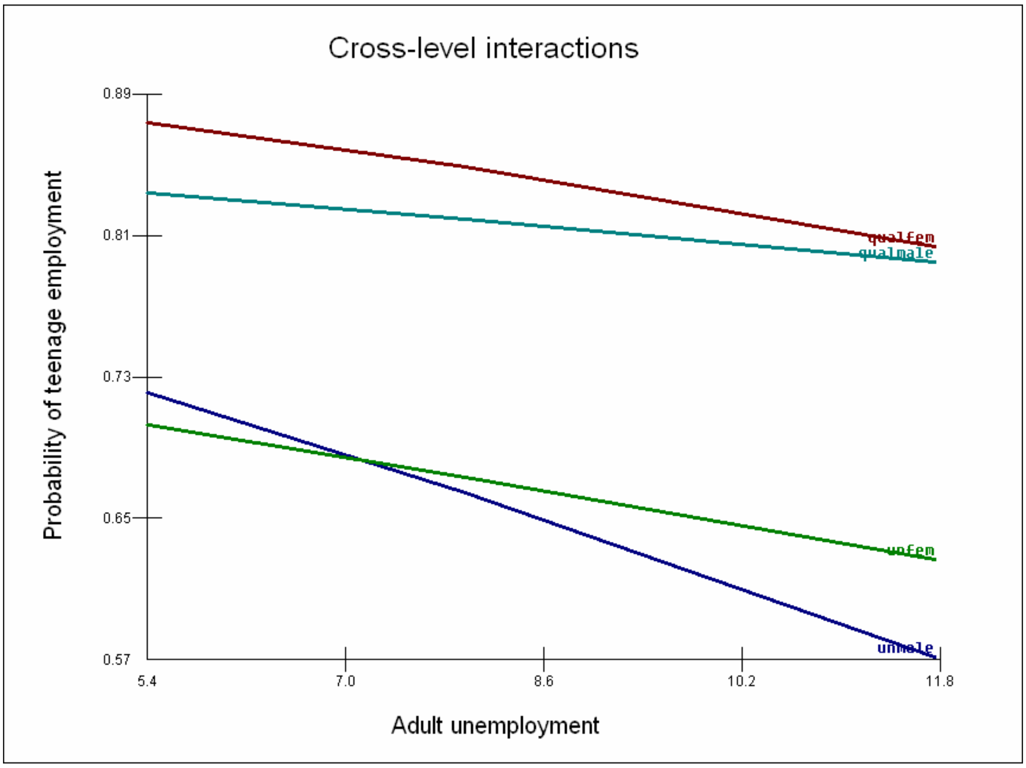


adunemp.pre	Code.pred	constant.pre	median.pred	median.low.p	median.high.p	mean.pred	mean.low.pre	mean.high.pr
5.404	unmale	1	.728	.66	.787	.723	.657	.782
7.968	unmale	1	.669	.61	.724	.666	.608	.719
11.749	unmale	1	.572	.509	.63	.572	.511	.628
5.404	unfem	1	.71	.645	.773	.706	.643	.767
7.968	unfem	1	.68	.623	.734	.677	.621	.73
11.749	unfem	1	.633	.578	.686	.63	.577	.682
5.404	qualmale	1	.841	.805	.875	.835	.799	.869
7.968	qualmale	1	.827	.795	.858	.821	.79	.853
11.749	qualmale	1	.804	.758	.845	.798	.752	.839
5.404	qualfem	1	.88	.845	.909	.874	.839	.905
7.968	qualfem	1	.856	.825	.885	.85	.819	.88
11.749	qualfem	1	.813	.77	.852	.807	.764	.846

To get a plot
Plot Grid
Y is median pred (cluster specific)
X is adunemp (the continuous predictor)
Grouped by code.pred
Tick off the 95% CI's



Thickening the lines and putting labels on the graph:



Estimating the VPC

The next thing that we would like to do for this model is to partition the variance to see what percentage of the residual variation still lies between postcodes. This is not as straightforward as in the normal-theory case.

$$\text{proportion}_{ij} \sim \text{Binomial}(\text{total}_{ij}, \pi_{ij})$$

$$\text{logit}(\pi_{ij}) = \beta_{0j}\text{constant} + 0.048(0.168)\text{unfem}_{ij} + 0.866(0.160)\text{qualmale}_{ij} + 1.078(0.165)\text{qualfem}_{ij} + \text{-}0.111(0.025)\text{adunemp}_j + 0.054(0.030)\text{adunemp.unfem}_{ij} + 0.071(0.033)\text{adunemp.qualmale}_{ij} + 0.028(0.033)\text{adunemp.qualfem}_{ij}$$

$$\beta_{0j} = 0.705(0.127) + u_{0j}$$

$$\left[u_{0j}\right] \sim \text{N}(0,\ \Omega_u)\ :\ \Omega_u = \left[0.153(0.062)\right]$$

$$\text{var}(\text{proportion}_{ij}|\pi_{ij}) = \pi_{ij}(1 - \pi_{ij})/\text{total}_{ij}$$

One simple method is to use a threshold approach (Snijders T, Bosker R, 1999 *Multilevel analysis: an introduction to basic and advanced multilevel modeling*, London, Sage) and to treat the level 1 between cell variation as having a variance of a standard logistic distribution which is 3.29. Then with this model, the proportion of the variance lying between postcode is

calc b1 = 0.153/ (0.153 + 3.29)
0.044438

But this ignores the fact that the level –1 variance is not constant, but is function of the mean probability which depends on the predictors in the fixed part of the model. There is a macro called VPC.txt that will simulate the values given desired settings for the predictor variables

Input values to c151 for all the fixed predictor values (Data manipulation and View)
EG 1 0 0 0 0 0 0 0 represents unqualified males in an area of average adult unemployment
EG 1 0 0 1 0 0 0 0 represents qualified females in an area of average adult unemployment

Input values in c152 for predictor variables which have random coefficients at level 2
EG c152 1 because this a random intercepts model

To run the Macro
File on main menu
Open macro vpc.txt then Execute

The result is obtained by printing B8
->prin b8
B8
0.033117

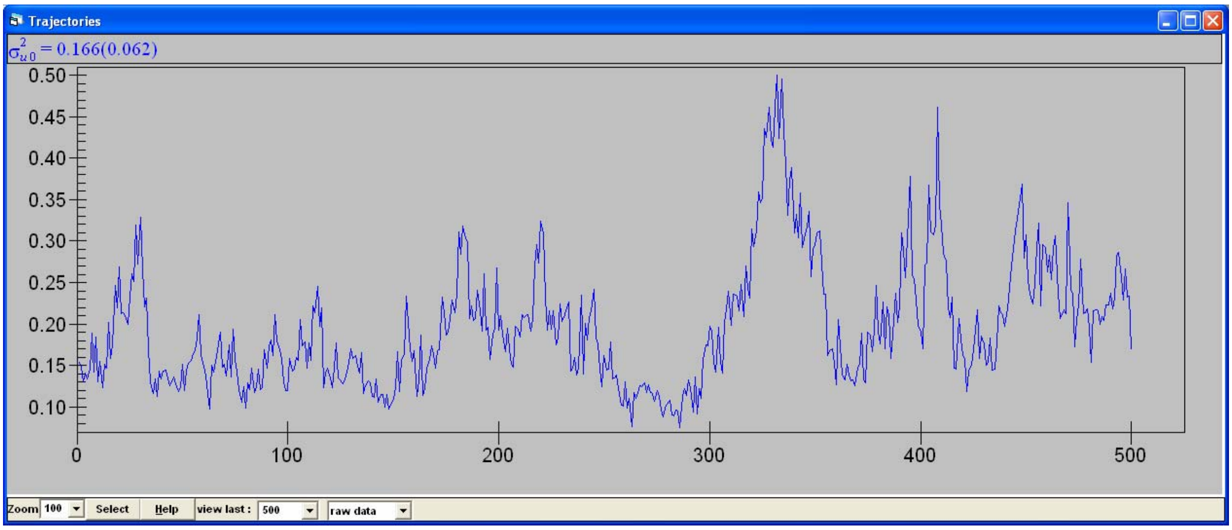
which is for unqualified males, while the result for qualified females is
->prin b8
B8
0.020085
So some 2 to 4% of the residual variance lies between postcodes.

Comparing models
Unfortunately because of the way that logit model are estimated in MLwiN through quasi-likelihood, it is not possible to use the deviance to compare models. One could use the Intervals and Tests procedures to test individual and sets of estimates for significance. But using MCMC methodology one can compare the overall fit of model using the DIC diagnostic

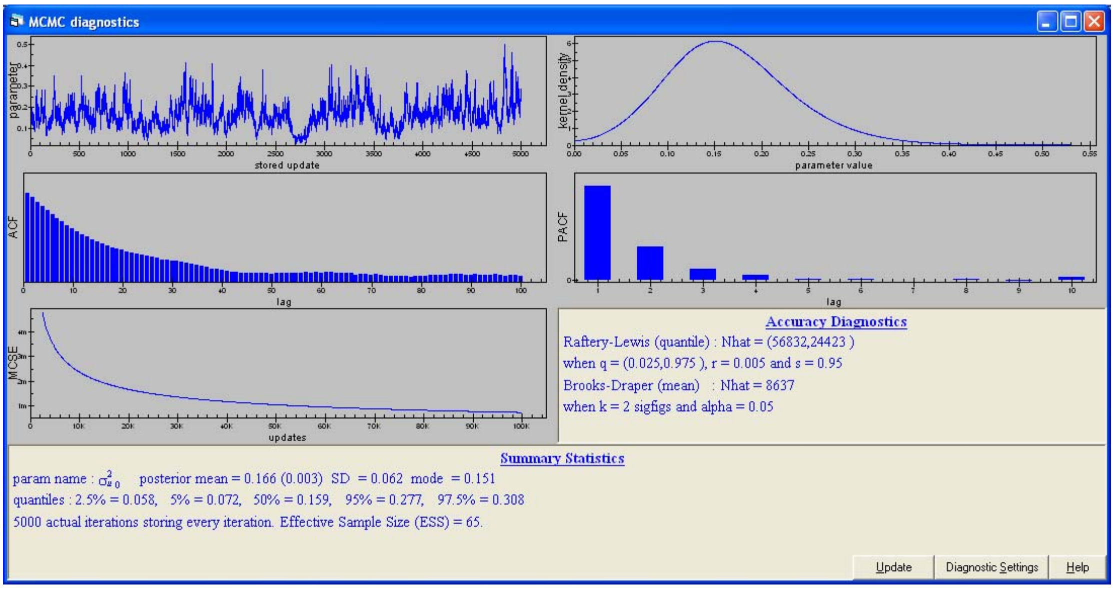
Using the IGLS/ RIGLS estimates as starting values
Estimation Control
Switch to MCMC and use the default values of a burn-in of 500, followed by a monitoring length of 5000
Start
To examine the estimates
Model on main menu
Trajectories

Select the level 2 variance (Postcode: Constant/Constant)
Change Structured graph layout to '1 graph per row'
Done

This gives the trajectory of the estimate for the last 500 simulated draws



Click in the middle of this graph to get the summary of these results:



You can see that the mean of the estimate for the level-2 variance is 0.166 and the 95% credible interval does not include zero in going from 0.058 to 0.308; the parameter distribution is positively skewed. Note however that both the Raftery-Lewis and Brooks-Draper statistics are suggesting that we have not ran the chain for long enough as the chain is highly auto-correlated; we have requested a run of 5000 simulations but they are behaving as an effective sample size of only 65. Ignoring this for the moment, we want to get the DIC diagnostic,

Model on main menu

MCMC

DIC diagnostic

produces the following results

```
Bayesian Deviance Information Criterion (DIC)
  Dbar    D(thetabar)    pD      DIC
  885.76     844.88    40.87   926.63
```

To increase the number of simulated draws

Estimation Control

MCMC

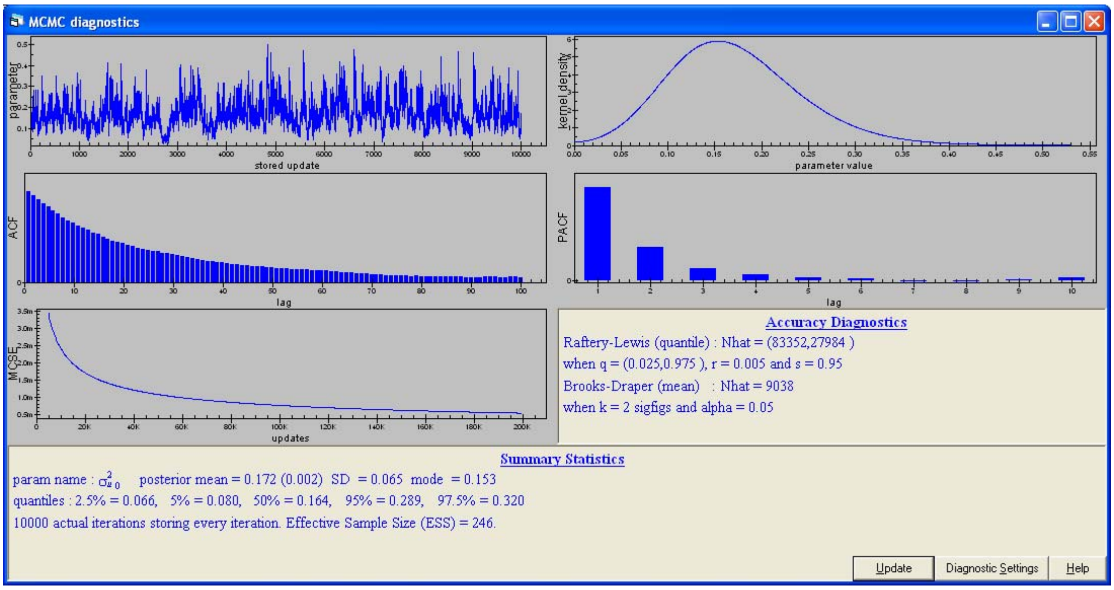
Change 5000 to 10000

Done

More iterations on top bar

The trajectories will be updated as the 5000 extra draws are performed (it makes good sense in large model to close this window down as it slows down the model, without being really informative)

Click Update on the MCMC diagnostics



To see that there are now effectively now 246 independent draws, the DIC diagnostic is

```
Bayesian Deviance Information Criterion (DIC)
  Dbar    D(thetabar)    pD      DIC
  885.27     843.67    41.59   926.86
```

Doubling the number of draws has changed the DIC diagnostic by only a small amount

There are two key elements to the interpretation of the DIC:

pD This gives the complexity of the model as the 'effective degrees' of freedom consumed in the fit, this takes into account both the fixed and random part; here we know there are 8 fixed terms and the rest of the effective degrees of freedom comes from treating the 122 postcodes as a distribution;

DIC Deviance Information Criterion (DIC), which is a generalisation of the Akaike Information Criterion (AIC); The AIC the Deviance + 2p, where p is the number of parameters fitted in the model and the model with the smallest AIC is chosen as the most appropriate. The DIC diagnostic statistic is simple to calculate from an MCMC run as it simply involves calculating the value of the deviance at each iteration, and the deviance at the expected value of the unknown parameters. Then we can calculate the 'effective' number of parameters, by subtracting from the average deviance from the complete set of iterations . The DIC diagnostic can then be used to compare models as it consists of the sum of two terms that measure the 'fit' and the 'complexity' of a particular model. Models with a lower DIC are therefore to be preferred as a trade-off between complexity and fit. Crucially this measure can be used in the comparison of non-nested models and non-linear models.

Here are the results for a set of models, all based on 10k simulated draws. To change a model specification, you have to use IGLS/ RIGLS estimation and then MCMC and with single models you cannot use mql and 2nd order IGLS. The results are ordered in terms of increasing DIC, the simplest and yet best fitting model at the top. The Mwipe command clears the stored models

Model	*Terms*	*PD*	*DIC*
4	2level,Cons+Code+Ad-Unemp	38.63	927.38
5	2level,Cons+Code*Ad-Unemp	41.39	927.39
3	2 level,Cons+Code	48.71	937.01
2	2 level,Cons	49.16	1025.44
1	1 level,Cons	1	1086.44

In terms of DIC, the chosen model is a two level one, with an additive effect for 3 categories of code and an additive effect for adult-unemployment, although there is no substantive difference to the model with the cross-level interactions

The plot for the final most parsimonious model is given below for logits and probabilities.

$\text{proportion}_{ij} \sim \text{Binomial}(\text{total}_{ij}, \pi_{ij})$

$\text{logit}(\pi_{ij}) = \beta_{0j}\text{constant} + 0.190(0.150)\text{unfem}_{ij} + 0.985(0.148)\text{qualmale}_{ij} + 1.161(0.153)\text{qualfem}_{ij} +$
$-0.072(0.015)(\text{adunemp-8})_j$

$\beta_{0j} = 0.608(0.116) + u_{0j}$

$\begin{bmatrix} u_{0j} \end{bmatrix} \sim \text{N}(0,\ \Omega_u) \ : \ \Omega_u = \begin{bmatrix} 0.161(0.064) \end{bmatrix}$

$\text{var}(\text{proportion}_{ij}|\pi_{ij}) = \pi_{ij}(1 - \pi_{ij})/\text{total}_{ij}$

Deviance(MCMC) = 889.887(401 of 401 cases in use)