

## Contents

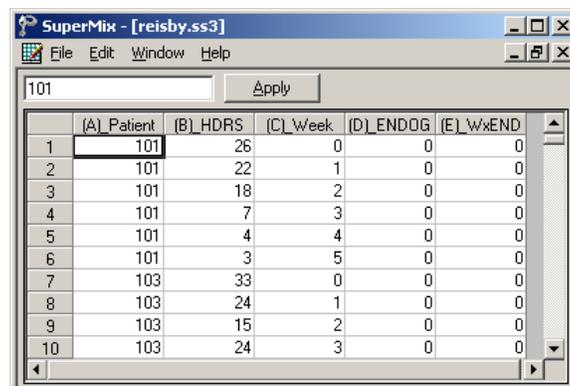
<i>Two-level models for continuous outcomes</i> .....	1
The data.....	1
The models.....	2
Example: Random intercept and slope model.....	3
Example: A random intercept and slope model with covariate and interaction effect.....	10

## Two-level models for continuous outcomes

### The data

The data set is from a study described in Reisby *et. al.*, (1977) that focused on the longitudinal relationship between imipramine (IMI) and desipramine (DMI) plasma levels and clinical response in 66 depressed inpatients (37 endogenous and 29 non-endogenous). Following a placebo period of 1 week, patients received 225 mg/day doses of imipramine for four weeks. In this study, subjects were rated with the Hamilton depression rating scale (HDRS) twice during the baseline placebo week (at the start and end of this week) as well as at the end of each of the four treatment weeks of the study. Plasma level measurements of both IMI and its metabolite DMI were made at the end of each week. The sex and age of each patient were recorded and a diagnosis of endogenous or non-endogenous depression was made for each patient.

Although the total number of subjects in this study was 66, the number of subjects with all measures at each of the weeks fluctuated: 61 at week 0 (start of placebo week), 63 at week 1 (end of placebo week), 65 at week 2 (end of first drug treatment week), 65 at week 3 (end of second drug treatment week), 63 at week 4 (end of third drug treatment week), and 58 at week 5 (end of fourth drug treatment week). The sample size is 375. Data for the first 10 observations of all the variables used in this section are shown below in the form of a SuperMix spreadsheet file, named **reisby.ss3**.



	(A) Patient	(B) HDRS	(C) Week	(D) ENDOG	(E) wxEND
1	101	26	0	0	0
2	101	22	1	0	0
3	101	18	2	0	0
4	101	7	3	0	0
5	101	4	4	0	0
6	101	3	5	0	0
7	103	33	0	0	0
8	103	24	1	0	0
9	103	15	2	0	0
10	103	24	3	0	0

The variables of interest are:

- Patient is the patient ID (66 patients in total).
- HDRS is the Hamilton depression rating scale.
- Week represents the week (0, 1, 2, 3, 4 or 5) at which a measurement was made.
- ENDOG is dummy variable for the type of depression a patient was diagnosed with (1 for endogenous depression and 0 for non-endogenous depression).
- WxENDOg represents the interaction between Week and ENDOG, and is the product of Week and ENDOG.

## The models

A general two-level model for a continuous response variable  $y$  depending on a set of  $r$  predictors  $x_1, x_2, \dots, x_r$  can be expressed as

$$y_{ij} = \mathbf{x}'_{ij}\boldsymbol{\beta} + \mathbf{z}'_{ij}\mathbf{u}_i + e_{ij}$$

where  $y_{ij}$  denotes the value of  $y$  for the  $j$ -th level-1 unit nested within the  $i$ -th level-2 unit for  $i = 1, 2, \dots, N$  and  $j = 1, 2, \dots, n_i$ . The scalar product  $\mathbf{x}'_{ij}\boldsymbol{\beta}$  is the fixed part of the model, and  $\mathbf{z}'_{ij}\mathbf{u}_i$  and  $e_{ij}$  denote the random part of the model at levels 2 and 1 respectively. For the fixed part of the model,  $\mathbf{x}'_{ij}$  is a typical row of a design matrix  $\mathbf{X}_i$  while the vector  $\boldsymbol{\beta}$  contains the fixed, but unknown, parameters to be estimated. In the case of the random part of the model at level 2,  $\mathbf{z}'_{ij}$  represents a typical row of a design matrix  $\mathbf{Z}_i$ , and  $\mathbf{u}_i$  the vector of random level-2 effects to be estimated. It is assumed that  $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_N$  are independently and identically distributed (i.i.d.) with mean vector  $\mathbf{0}$  and covariance matrix  $\boldsymbol{\Phi}_{(2)}$ . Similarly, the  $e_{ij}$  are assumed i.i.d., with mean vector 0 and variance  $\sigma^2$ . The elements of  $\mathbf{z}_{ij}$  are typically a subset of those of  $\mathbf{x}_{ij}$ .

## The random intercept and slope model

The random intercept and slope model for the response variable HDRS may be expressed as

$$\text{HDRS}_{ij} = \beta_0 + \beta_1 * (\text{Week})_{ij} + u_{i0} + u_{i1} (\text{Week})_{ij} + e_{ij}$$

where  $\beta_0$  denotes the average expected depression rating scale value,  $\beta_1$  denotes the coefficient of the predictor variable Week (slope) in the fixed part of the model,  $u_{i1}$  denotes the variation in the slopes over patients, and  $u_{i0}$  and  $e_{ij}$  denote the variation in the average expected HDRS value over patients and between patients respectively.

## The random intercept and slope with a covariate and an interaction model

The random intercept and slope model for the response variable HDRS with the variable ENDOG as a covariate and with an interaction effect between Week and ENDOG may be expressed as

$$\text{HDRS}_{ij} = \beta_0 + \beta_1 * (\text{Week})_{ij} + \beta_2 * (\text{ENDOG})_i + \beta_3 * (\text{WxENDOG})_{ij} + u_{i0} + u_{i1} (\text{Week})_{ij} + e_{ij}$$

where  $\beta_0$  denotes the average expected depression rating scale value,  $\beta_1$  and  $\beta_2$  denote the coefficients of the predictor variables Week and ENDOG in the fixed part of the model,  $\beta_3$  denotes the coefficient of the interaction between Week and ENDOG in the fixed part of the model,  $u_{i1}$  denotes the variation in the Week slopes over patients, and  $u_{i0}$  and  $e_{ij}$  denote the variation in the average expected HDRS value over patients and over measurements (*i.e.*, between patients) respectively.

### Example: Random intercept and slope model

#### Importing the data

The random intercept and slope model above is fitted to the data in **reisby.ss3**. The first step is to create the **ss3** file shown above from the Excel file **reisby.xls**. This is accomplished as follows.

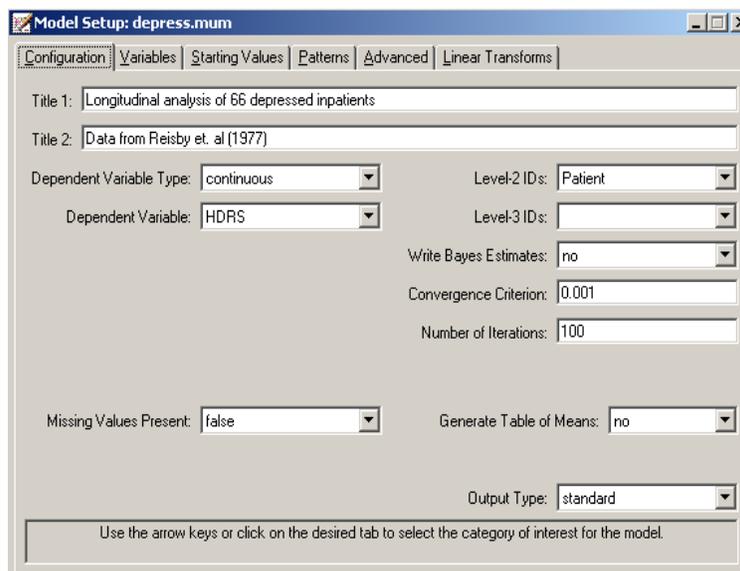
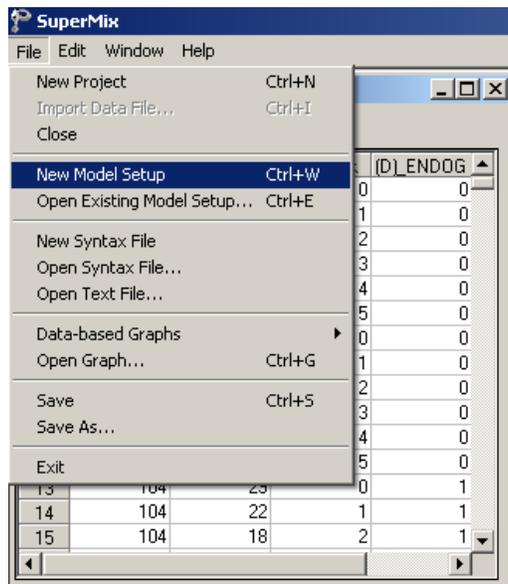
- Use the **Import Data File** option on the **File** menu to open the **Open** dialog box.
- Browse for the file **reisby.xls** in the **Examples\Primer\Continuous** folder.
- Select the file and click on the **Open** button to open the following SuperMix spreadsheet window for **reisby.ss3**.

	(A)_Patient	(B)_HDRS	(C)_Week	(D)_ENDOG	(E)_WxEN
1	101.00	26.00	0.00	0.00	0.00
2	101.00	22.00	1.00	0.00	0.00
3	101.00	18.00	2.00	0.00	0.00
4	101.00	7.00	3.00	0.00	0.00
5	101.00	4.00	4.00	0.00	0.00
6	101.00	3.00	5.00	0.00	0.00
7	103.00	33.00	0.00	0.00	0.00
8	103.00	24.00	1.00	0.00	0.00
9	103.00	15.00	2.00	0.00	0.00
10	103.00	24.00	3.00	0.00	0.00

After selecting the **File, Save** option, we are ready to fit the random intercept and slope model for HDRS to the data in **reisby.ss3**.

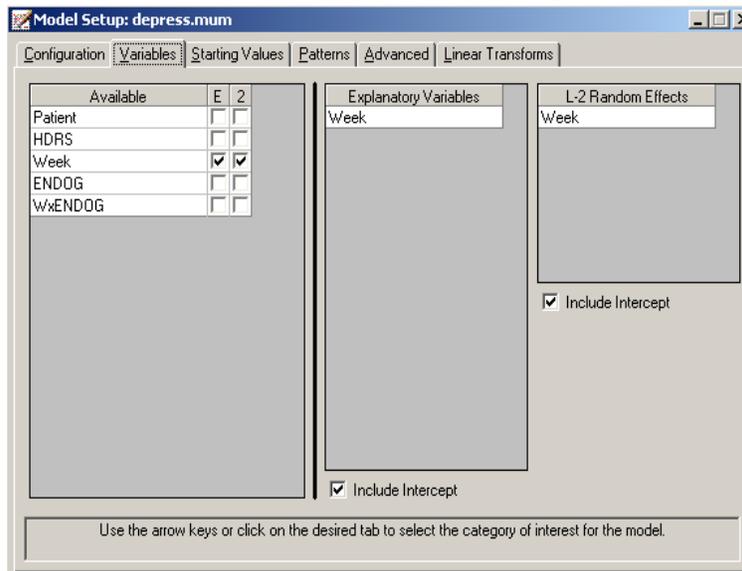
#### Setting up the analysis

Start by selecting the **New Model Setup** option on the **File** menu as shown below to open the **Model Setup** window. The **Model Setup** window has six tabs: **Configuration, Variables, Starting Values, Patterns, Advanced,** and **Linear Transforms**. In this example, only the **Configuration** and the **Variables** tabs are used.

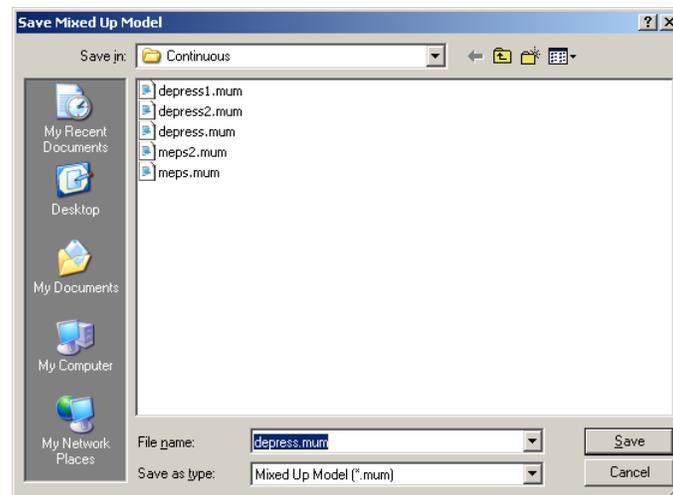


Starting with the **Configuration** screen, the titles Longitudinal analysis of 66 depressed inpatients and Data from Reisby et. al. (1977) are entered in the **Title 1** and **Title 2** text boxes respectively. The continuous outcome variable HDRS is selected from the **Dependent Variable** drop-down list box. The variable Patient, which defines the levels of the hierarchy, is selected as the Level-2 ID from the **Level-2 IDs** drop-down list box. We keep the default settings of all the other options in order to produce the **Configuration** screen as shown above.

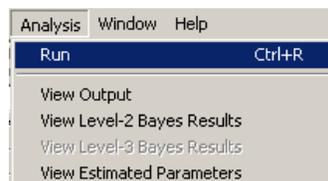
Click the **Variables** tab to proceed to the **Variables** screen of the **Model Setup** window. This screen shows the list of variables available for analysis and next to it two columns, with headings **E** (for explanatory variables) and **2** (for level-2 random effects). The variable Week is specified as the covariate of the fixed part of the model by checking the **E** check box for Week in the **Available** grid. We mark the **2** check box for Week in the **Available** grid to specify the random slope at level 2 of the model. After completion, the **Variables** screen should look as shown below.



Before the analysis can be run, the model specifications have to be saved to file. To accomplish this, we select the **Save As** option on the **File** menu to open the **Save Mixed Up Model** dialog box and then enter the name **depress.mum** in the **File name** text box to produce the following dialog box.



The analysis is run by selecting the **Run** option from the **Analysis** menu as shown below



to produce the corresponding output file **depress.out**.

## Discussion of results

Four separate sections of the output file **depress.out** are shown below. The summary of the hierarchical structure of the data below, which is given first, shows how the 375 measurements are nested within the 66 patients. It also indicates that the number of repeated measurements per patient varies from 4 to 6 observations.

```
Numbers of observations
-----
Level 2 observations =      66
Level 1 observations =     375

N2 :    1    2    3    4    5    6    7    8
N1 :    6    6    6    6    5    5    6    5

N2 :    9   10   11   12   13   14   15   16
N1 :    5    5    6    5    6    6    6    6

N2 :   17   18   19   20   21   22   23   24
N1 :    6    6    6    6    6    5    6    6

N2 :   25   26   27   28   29   30   31   32
N1 :    6    5    4    6    6    5    6    6

N2 :   33   34   35   36   37   38   39   40
N1 :    6    5    5    5    6    5    6    6

N2 :   41   42   43   44   45   46   47   48
N1 :    6    5    6    6    6    6    5    6

N2 :   49   50   51   52   53   54   55   56
N1 :    6    6    5    5    6    6    5    6

N2 :   57   58   59   60   61   62   63   64
N1 :    6    6    6    6    6    5    6    6

N2 :   65   66
N1 :    6    6
```

The following portion of the output file consists of a listing of selected descriptive statistics of the variables of the model. The descriptive results show, for example, that the mean HDRS score is 17.64 with a standard deviation of 7.19.

```
Descriptive statistics for all variables
-----
Variable           Minimum      Maximum      Mean      Stand. Dev.

Dependent
-----
HDRS                0.00000    39.00000    17.63733    7.19006

Random-Effects
-----
intcept (2)         1.00000    1.00000     1.00000     0.00000
Week (2)            0.00000    5.00000     2.48000     1.68320
intcept (1)         1.00000    1.00000     1.00000     0.00000

Fixed Regressor(s)
-----
intcept             1.00000    1.00000     1.00000     0.00000
Week                0.00000    5.00000     2.48000     1.68320
```

The next part of the output file contains the starting values for both fixed and random components.

SuperMix - [depress.out]

File Analysis Window Help

Parameter starting values  
-----

Fixed regressor(s)  
-----

Variable	Estimate	Std. Err.	Z-value	p-value
intercept	23.60263	0.54761	43.10128	0.00000
Week	-2.40536	0.18278	-13.15986	0.00000

Log Likelihood = -2474.9557  
Number of free parameters = 6

Variance/covariance components  
-----

Level 2		Estimate	Std. Err.	Z-value	p-value
intercept	/intercept	12.73430	0.27812	45.78734	0.00000
Week	/intercept	-1.50873	0.16119	-9.36004	0.00000
Week	/Week	2.18703	0.18554	11.78759	0.00000

Level 1		Estimate	Std. Err.	Z-value	p-value
intercept	/intercept	12.14019	0.09064	133.93833	0.00000

Save As... Close

The last part of the output shows the final estimates of the fixed and random coefficients included in the model, along with some goodness of fit measures. The results given below show that the  $p$ -values for the time effect, as represented by the variable Week, is highly significant. At the beginning of the study, when Week = 0, the average expected HDRS score is 23.57695. For each subsequent week, a decrease of  $-2.37707$  in average HDRS score is expected. At the end of the study period, the average expected HDRS score is  $23.57695 - 5(2.37707) = 11.6916$ .

The  $p$ -values for the estimates of the random coefficients are also significant, with the exception of that for the covariance between the intercept and slope. From the output above we have  $\hat{\text{var}}(u_{i0}) = 12.62930$ ,  $\hat{\text{var}}(u_{i1}) = 2.07899$ ,  $\hat{\text{cov}}(u_{i0}, u_{i1}) = -1.42093$ , and that  $\hat{\text{var}}(e_{ij}) = 12.21663$ . It is clear that there is considerably more variation in patients' intercepts than in their slopes (12.62930 vs. 2.07899). This indicates that there are significant differences in the initial HDRS scores, but that the patients' slopes over time do not vary as much. This seems to indicate that the pattern in HDRS scores over time may be similar for patients, although they start with markedly different initial HDRS scores. Typically, one would expect most of the variation in HDRS scores at the measurement level, and thus would expect  $\hat{\text{var}}(e_{ij})$  to be larger than any of the other variances/covariances. In this case, however, there is more variation in the random intercepts over patients than in the measurements nested within patients. Due to this, it may be of interest to take a closer look at the variation in HDRS scores at the two levels of the hierarchy.

In the case of a model with only a random intercept, there are two variances of interest: the variation in the random intercept over the patients, and the residual variation at level-1, over the measurements. By calculating the total variation in the HDRS score explained by such a model, obtained as  $\hat{\text{var}}(e_{ij}) + \hat{\text{var}}(u_{i0})$ , we can obtain an estimate of the intraclass correlation coefficient.

The intraclass coefficient is defined as

$$ICC = \frac{\hat{\text{var}}(u_{i0})}{\hat{\text{var}}(e_{ij}) + \hat{\text{var}}(u_{i0})}$$

and would, for a random intercept model for this data, represent the proportion in HDRS scores between patients.

```

depress.out
-----
Convergence attained in 5 iterations
-----

TITLE: Longitudinal analysis of 66 depressed patients
       Data from Resiby et. al (1977)

Maximum likelihood estimates
-----

Fixed regressor(s)
-----
Variable             Estimate      Std. Err.      Z-value      p-value
-----
intcept              23.57695      0.54555       43.21714     0.00000
Week                 -2.37707      0.20865      -11.39280    0.00000

Log Likelihood          =    -1109.5188
-2 Log Likelihood (Deviance) =    2219.0375
Akaike's Information Criterion =    2231.0375
Schwarz's Bayesian Criterion =    2244.1754
Number of free parameters =          6

Variance/covariance components
-----

Level 2
-----
Variable             Estimate      Std. Err.      Z-value      p-value
-----
intcept /intcept      12.62930      3.46653       3.64322     0.00027
Week /intcept         -1.42093      1.02595      -1.38500     0.16605
Week /Week            2.07899      0.50417       4.12363     0.00004

Level 1
-----
Variable             Estimate      Std. Err.      Z-value      p-value
-----
intcept /intcept      12.21663      1.10696      11.03615     0.00000

```

In the current model, which contains both a random intercept and a random slope, the situation is somewhat more complicated. This is due to the possible correlation between the level-2 random effects. When calculating an estimate of the total variation, the covariance(s) between random effects have to be taken into account in any attempt to estimate the proportion of variation in outcome at any level or for any random coefficient. In addition, the inclusion of a covariate such as ENDOG can affect the variance estimates.

The total variation in HDRS scores over patients is defined as

$$\text{Var}(\text{level } 2) = \text{var}(u_{i0}) + \text{var}(u_{i1})(\text{WEEK})_{ij}^2 + 2[\text{cov}(u_{i0}, u_{i1})](\text{Week})_{ij}$$

The total variation is a function of the value assumed by the predictor Week, which has a random slope. As such, the total variation at the beginning of the study is

$$\begin{aligned}\text{Var}(\text{level 2}) &= \text{var}(u_{i0}) + \text{var}(u_{i1})(0)^2 + 2[\text{cov}(u_{i0}, u_{i1})](0) \\ &= \text{var}(u_{i0})\end{aligned}$$

while at the end of the study we have

$$\begin{aligned}\text{Var}(\text{level 2}) &= \text{var}(u_{i0}) + \text{var}(u_{i1})(5)^2 + 2[\text{cov}(u_{i0}, u_{i1})](5) \\ &= \text{var}(u_{i0}) + 25 \text{var}(u_{i1}) + 10 \text{cov}(u_{i0}, u_{i1})\end{aligned}$$

An estimate of the total variation at this level can be obtained by using the estimates of the variances and covariance obtained under this model. By substituting  $\hat{\text{var}}(u_{i0})$ ,  $\hat{\text{var}}(u_{i1})$ , and  $\hat{\text{cov}}(u_{i0}, u_{i1})$  into the equations above, we obtain the estimated variation in HDRS scores over patients at different points during the study period.

At the beginning of the study, the estimated total variation in HDRS scores over patients is simply the estimated variation in the random intercept, *i.e.*,  $\hat{\text{var}}(u_{i0}) = 12.62930$ . At the end of the study, the total variation at level-2 is estimated as

$$\begin{aligned}\hat{\text{var}}(\text{level 2}) &= \hat{\text{var}}(u_{i0}) + 25 \hat{\text{var}}(u_{i1}) + 10 \hat{\text{cov}}(u_{i0}, u_{i1}) \\ &= 12.62930 + 25(2.07899) + 10(-1.42093) \\ &= 50.39475.\end{aligned}$$

At the beginning of the study we obtain

$$\begin{aligned}\frac{\hat{\text{var}}(\text{level 2})}{\hat{\text{var}}(\text{level 2}) + \hat{\text{var}}(\text{level 1})} &= \frac{12.62930}{12.62930 + 12.21663} \\ &= 0.5083\end{aligned}$$

and thus conclude that that 50.8% of the variation in HDRS scores at this time is over patients. At the end of the study, we find that

$$\begin{aligned}\frac{\hat{\text{var}}(\text{level 2})}{\hat{\text{var}}(\text{level 2}) + \hat{\text{var}}(\text{level 1})} &= \frac{50.39475}{50.39475 + 12.21663} \\ &= 0.8049,\end{aligned}$$

so that only 20% of the variation in HDRS scores are estimated to be at the measurement level, with 80% at the patient level. As mentioned before, the total variation in HDRS scores is a function of the time of measurement, as represented by the variable Week. The very different estimates of variation at a patient level show how the introduction of an important predictor, in this case at the measurement level, can have an impact on variance estimates at a different level of the hierarchy. By the end of the study period, the residual variation over measurements has been dramatically reduced, this being explained largely by the inclusion of the time effect. Most of the remaining unexplained variation is at the patient level. As a result of this finding and in the light of our original research question, whether the initial depression classification of a patient is also related to the HDRS scores over the time in which medication is administered, the model will be extended to include the covariate ENDOG. This dichotomous variable assumes a value of 1 when endogenous depression

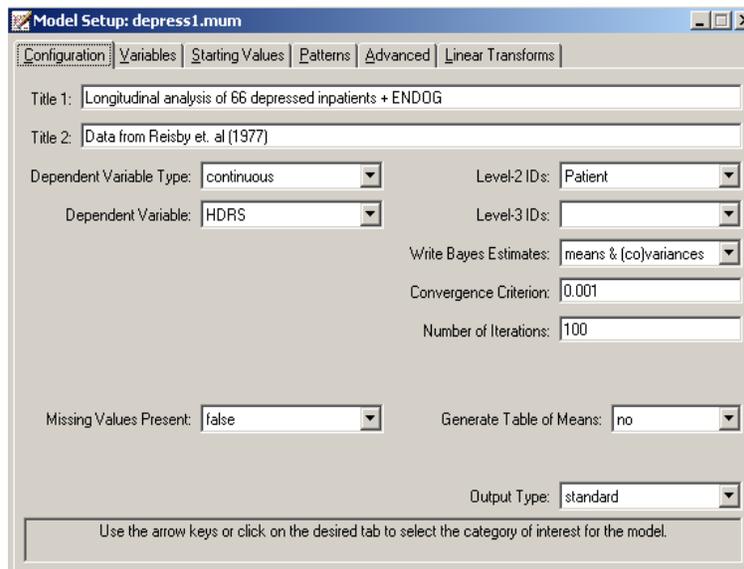
was observed, and 0 if not. In addition, we will make provision for a possible interaction between depression classification and the measurement occasion by including the interaction term  $W \times \text{ENDOG}$  in the model. While  $W \times \text{ENDOG}$  can be viewed as a cross-level interaction, as *Week* is a measurement-level variable and *ENDOG* a patient-level variable, the inclusion of the patient-level variable *ENDOG* may enable us to explain more of the remaining variation in the random intercepts and slopes at the patient level.

### Example: A random intercept and slope model with covariate and interaction effect

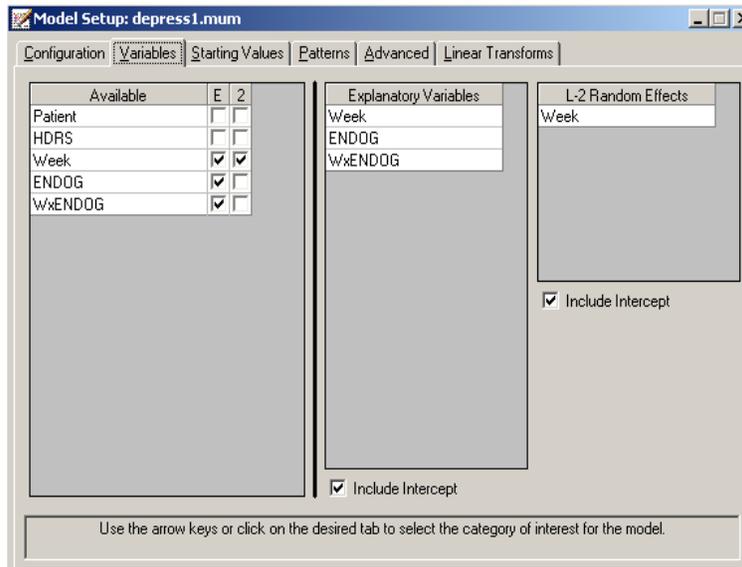
To fit the random intercept and slope model with a covariate and an interaction effect for HDRS, we need to include the variables *ENDOG* and  $W \times \text{ENDOG}$  as covariates in the model. Significant *ENDOG* and  $W \times \text{ENDOG}$  coefficients will imply that there are significant differences in depression value ratings between the two groups of patients, or that there is significant interaction between the type of depression and the time at which a measurement was made. One way to decide whether *ENDOG* and  $W \times \text{ENDOG}$  should be included is to fit the model with and without these predictors to the data. Then the difference in the deviance ( $-2 \log$ -likelihood) values can be used to obtain a  $\chi^2$ -square test statistic value. The difference in the number of parameters estimated in each of these models gives the corresponding degrees of freedom.

### Setting up the analysis

To create the model specifications for this model, we start by opening **reisby.ss3** in a SuperMix spreadsheet window. Then we use the **Open Existing Model Setup** option on the **File** menu to open the **Model Setup** window for **depress.mum**. We extend the string in the **Title 1** text box on the **Configuration** screen by adding the string "+ *ENDOG*". Since we would like to produce the Bayes estimates for the means, we select the **means & (co)variances** option from the drop-down list box next to **Write Bayes Estimates** to produce the following screen.



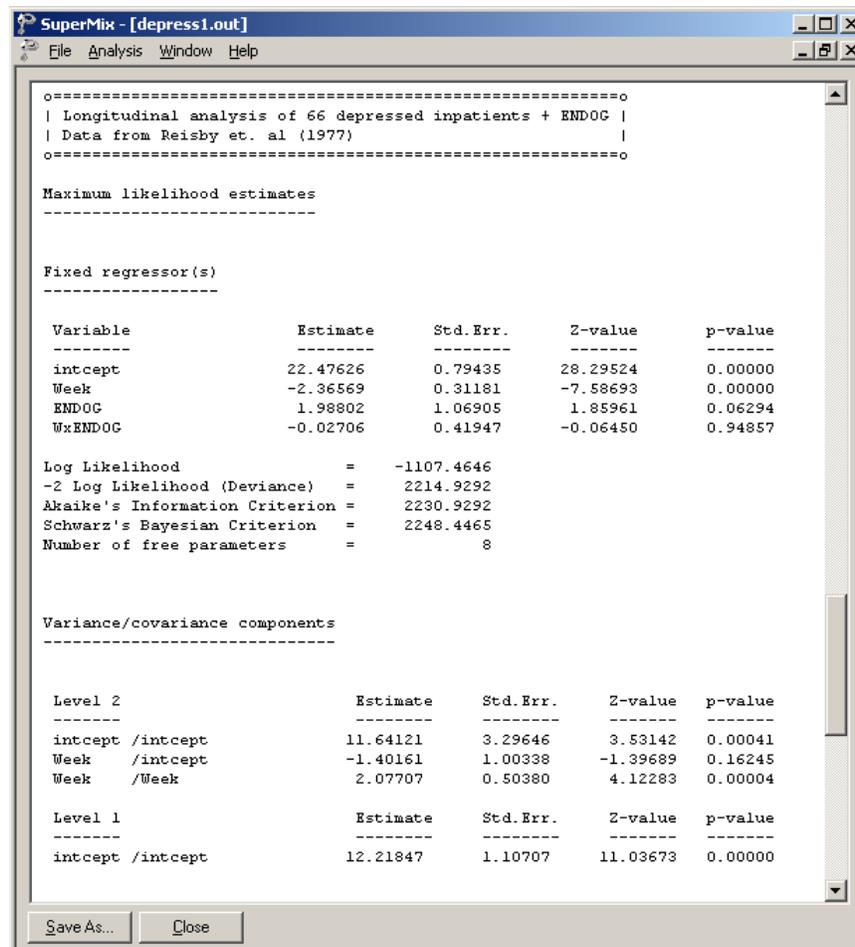
Next, click on the **Variables** tab to proceed to the **Variables** screen of the **Model Setup** window. The two covariates are specified by checking the **E** check boxes for *ENDOG* and  $W \times \text{ENDOG}$  in the **Available** grid respectively to produce the following **Variables** screen.



Save the changes to the file **depress1.mum** by using the **Save As** option on the **File** menu. To fit the revised model to the data, select the **Run** option on the **Analysis** menu to produce the output file **depress1.out**.

## Discussion of results

A portion of the output file **depress1.out** is shown below.

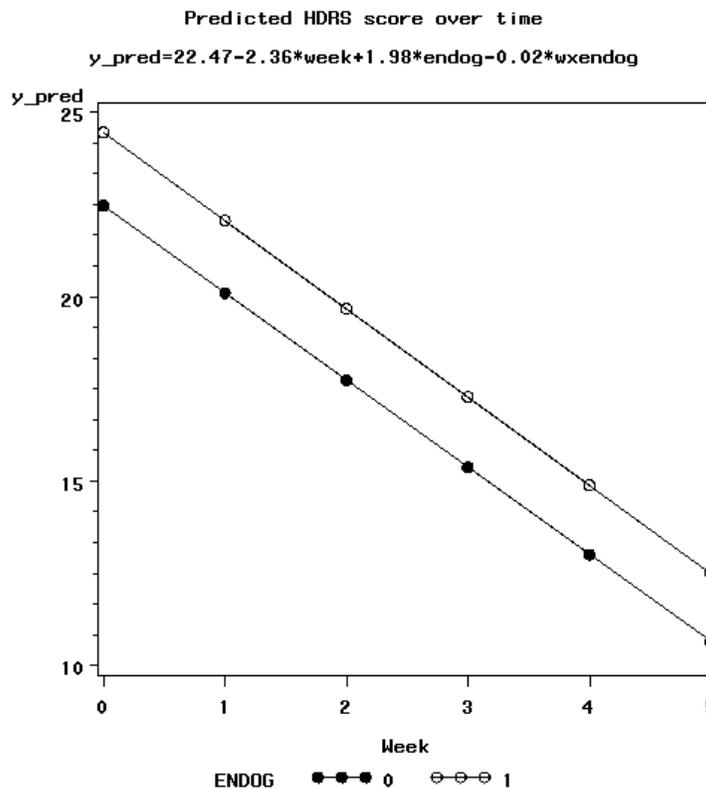


The interaction  $W \times \text{ENDOG}$  between the time variable  $\text{Week}$  and the depression classification of the patient, as represented by the variable  $\text{ENDOG}$ , is not significant. Given this, we can take a closer look at the estimated coefficients for the main effects  $\text{Week}$  and  $\text{ENDOG}$  respectively. Note, however, that the  $p$ -value for the  $\text{ENDOG}$  coefficient is larger than 0.05, and thus can only be considered significant at a 10% level of significance. The effect of time, on the other hand, is found to be highly significant. While the average HDRS score is predicted to decrease by  $-2.37$  score scale units each week, patients classified as having endogenous depression (*i.e.*,  $\text{ENDOG} = 1$ ) are predicted to have a HDRS score of 2 units higher at all occasions.

This is clear from a plot of the predicted HDRS scores over time for the two  $\text{ENDOG}$  groups, as shown below in Figure 3.1. For all practical purposes, the two lines are parallel to each other, again underscoring the absence of significant interaction between  $\text{Week}$  and  $\text{ENDOG}$ .

To obtain the predicted average HDRS scores as shown in these plots, the estimates obtained from the output are used:

$$\begin{aligned} \hat{y} &= \hat{\beta}_0 + \hat{\beta}_1(\text{Week}) + \hat{\beta}_2(\text{ENDOG}) + \hat{\beta}_3(W \times \text{ENDOG}) \\ &= 22.47626 - 2.36569(\text{Week}) + 1.98802(\text{ENDOG}) - 0.02706(W \times \text{ENDOG}) \end{aligned}$$



**Figure 3.1: Predicted HDRS scores over time for two groups**

To illustrate differences between significant and nonsignificant interaction, we examine two scenarios. In the first, suppose that a negative and significant estimate of  $-1.02$  was obtained for the interaction between  $\text{Week}$  and  $\text{ENDOG}$ . In this case, the graph (see Figure 3.2) would have looked somewhat different, as shown below. Under this scenario, patients with endogenous depression would have improved faster than their counterparts with non-endogenous depression.

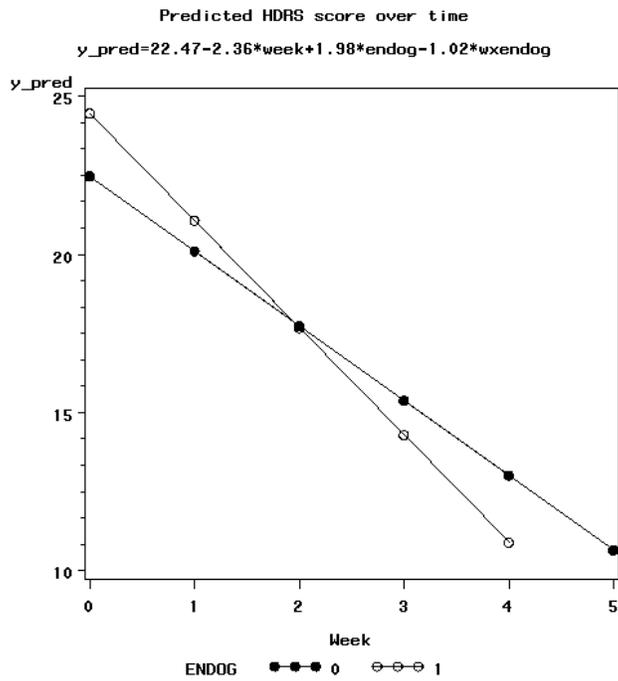


Figure 3.2: Predicted HDRS scores over time for two groups

For the second scenario, suppose that a significant and positive interaction term of 1.02 was obtained. In this case, patients with non-endogenous depression would have shown more marked improvement over time than the patients with endogenous depression. This scenario is illustrated in Figure 3.3.

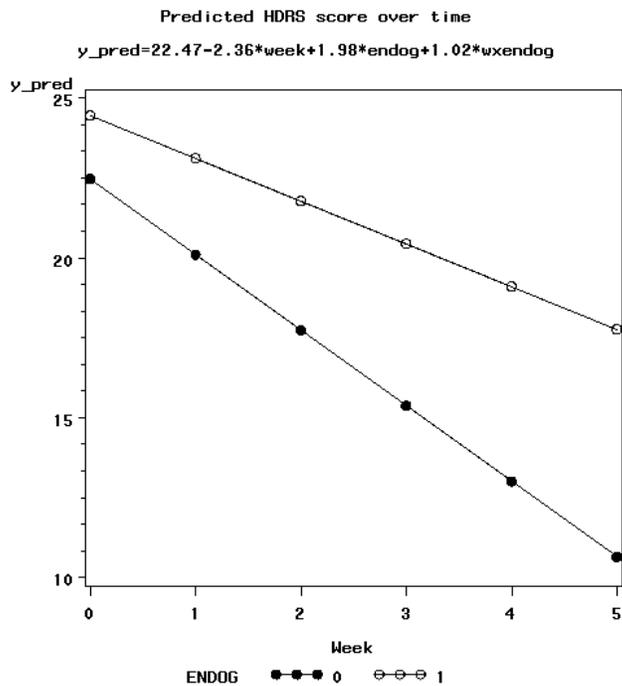


Figure 3.3: Predicted HDRS scores over time for two groups

A question that arises from inspection of the results is whether the interaction term contributes overall to the explanation of the variation in the HDRS scores. To test this, we can fit a model without the interaction term and use the deviance reported in the output to compare results for the model with interaction and the model without this term. The relevant output (**depress2.out**) from an analysis without the interaction term is shown below. We note that the deviance obtained for the simpler model is almost identical to that of the model considered in this section. Based on this, we conclude that a model without the interaction  $W \times \text{ENDOG}$  would fit the data as well as the one with the interaction term included.

Fixed regressor(s)				
Variable	Estimate	Std. Err.	Z-value	p-value
intcept	22.49344	0.74839	30.05592	0.00000
Week	-2.38064	0.20859	-11.41317	0.00000
ENDOG	1.95650	0.95083	2.05769	0.03962
Log Likelihood	=	-1107.4667		
-2 Log Likelihood (Deviance)	=	2214.9334		
Akaike's Information Criterion	=	2228.9334		
Schwarz's Bayesian Criterion	=	2244.2610		
Number of free parameters	=	7		

In addition, we can test the hypothesis that the model with covariate (ENDOG) fits the data better than the random intercept and slope model considered previously. To test this hypothesis, we calculate the difference between the  $-2$  log likelihood value obtained for the previous model (**depress1.out**) and the  $-2$  log likelihood value for the current model (**depress2.out**). It can be shown that this difference of  $2219.04 - 2214.93 = 4.11$  has a  $\chi^2$  distribution with associated degrees of freedom equal to the difference in the number of parameters estimated in the two examples, *i.e.*,  $8 - 7 = 1$  degree of freedom. Since the  $p$ -value for this test statistic is less than 0.05, it is concluded that the random intercept and slope model with ENDOG as a covariate does not provide a better description of the data than the original random intercept and slope model. This finding is supported by the fact that the  $p$ -value for ENDOG when the interaction effect between Week and ENDOG is excluded.

## Residual analysis

Up to this point, we have considered results averaged over all patients. We now turn our attention to the residual file **depress1.ba2**, which offers the opportunity to take a closer look at the results by individual patient. In the image below, the contents of this file are displayed for the first 10 patients.

Two lines of information are given for each patient, containing, in order of appearance,

- the patient ID,
- the number of the empirical Bayes coefficient,
- the empirical Bayes estimate,
- the estimated variance of the Bayes coefficient, and
- the name of the associated coefficient as used in the model.

Value	Week	Intercept	Slope	Label
101.00	1	1.7950	3.9520	intercept
101.00	2	-2.0413	0.45969	Week
103.00	1	4.3130	3.9520	intercept
103.00	2	-0.39169	0.45969	Week
104.00	1	1.9805	3.9520	intercept
104.00	2	-1.5243	0.45969	Week
105.00	1	-2.1804	3.9520	intercept
105.00	2	0.26636	0.45969	Week
106.00	1	-0.30569	4.1342	intercept
106.00	2	0.97334	0.69758	Week
107.00	1	-1.2044	3.9533	intercept
107.00	2	-0.47604	0.53483	Week
108.00	1	-2.5600	3.9520	intercept
108.00	2	-0.86300	0.45969	Week
113.00	1	0.31819E-01	4.1342	intercept
113.00	2	1.8643	0.69758	Week
114.00	1	-2.3599	5.8413	intercept
114.00	2	-0.46925	0.58433	Week
115.00	1	-2.9641	5.8413	intercept
115.00	2	0.42967	0.58433	Week

To obtain patient-specific predicted HDRS scores the empirical Bayes estimates for each patient have to be taken into account, as these estimates indicate the extent to which the random intercept or slope for that patient deviates from the intercept and slope over all patients. Patient-specific predicted HDRS scores are calculated as

$$\hat{y} | \mathbf{u}_i = 22.47626 - 2.36569(\text{Week}) + 1.98802(\text{ENDOG}) - 0.02706(\text{WxENDOG}) + \hat{u}_{i0} + \hat{u}_{i1}(\text{Week})$$

For the first patient shown in the residual file above, we have  $\hat{u}_{i0} = 1.7950$  and  $\hat{u}_{i1} = -2.0413$ . From this information, we can already tell that the intercept for the patient is higher than average, but that the Week slope for this patient is lower than average. This patient was classified as having non-endogenous depression, so that  $\text{ENDOG} = 0$ . At week 1, the predicted HDRS score for this patient (Patient = 101) is found to be

$$\begin{aligned} \hat{y} &= 22.47626 - 2.36569(1) + 1.7950 - 2.0413(1) \\ &= 19.8643 \end{aligned}$$

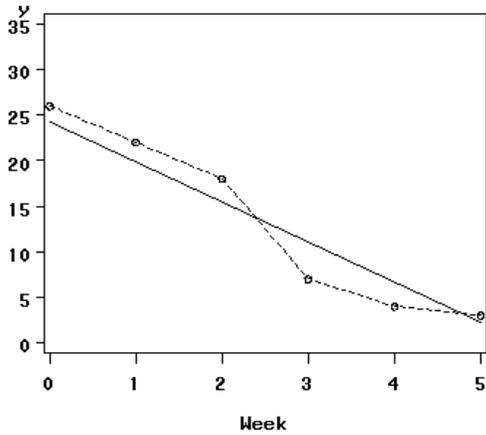
and at the end of the study period (Week = 5) the predicted HDRS score for Patient 101 is

$$\begin{aligned} \hat{y} &= 22.47626 - 2.36569(5) + 1.7950 - 2.0413(5) \\ &= 2.3631. \end{aligned}$$

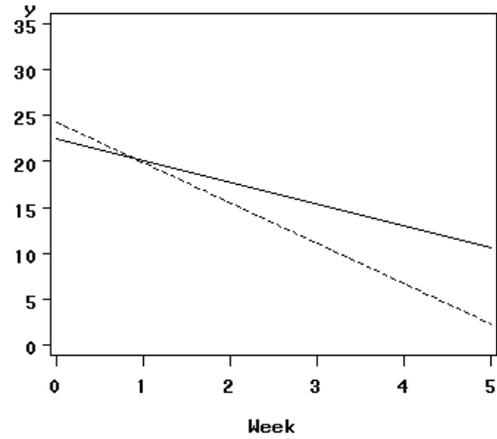
Figures 3.4 and 3.5 are graphical representations of the average and patient-specific regression lines of six patients, three of which were classified as non-endogenous and three as endogenous. The graphs on the left show the patient-specific regression lines (solid lines) and the observed HDRS trajectories (the dotted lines) for each of the patients. The graphs on the right show the average regression line (solid) and the patient-specific regression line (dotted).

For the patients with non-endogenous depression (Figure 3.4; graphs on the right), all the average regression lines are identical. Similarly, the average regression lines for patients with endogenous depression (Figure 3.5; graphs on the right) are identical. For the latter case, the average regression lines are higher. Recall that for the predictor ENDOG an estimated coefficient of 1.98802 was obtained. A patient classified as having endogenous depression is thus expected to have a higher average HDRS score than a patient with non-endogenous depression.

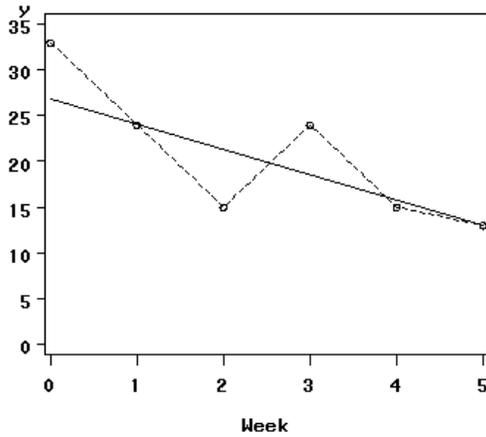
Regression line and observed trajectory  
Patient=101



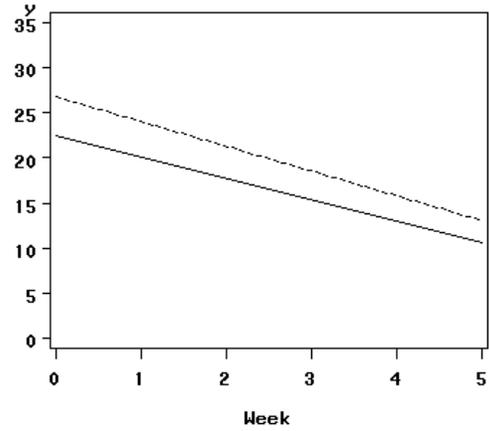
Patient and average regression lines  
Patient=101



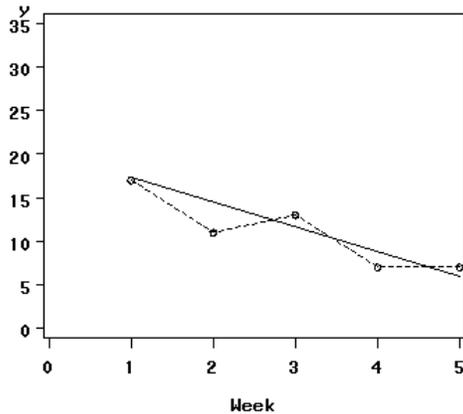
Regression line and observed trajectory  
Patient=103



Patient and average regression lines  
Patient=103



Regression line and observed trajectory  
Patient=114



Patient and average regression lines  
Patient=114

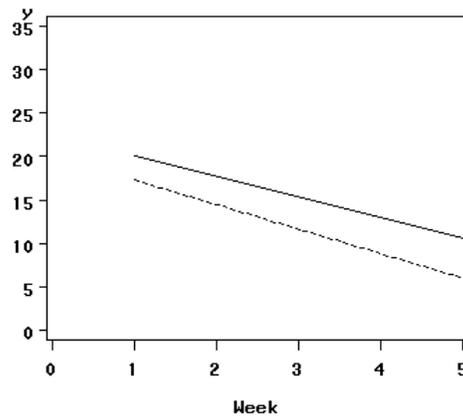


Figure 3.4: Average and unit-specific regression lines for patient with non-endogenous depression

From the top right graph in Figure 3.4 we find that patient number 101 had a higher initial HDRS score, but over time obtained a lower than average score. For patient 103, a higher than average predicted HDRS score is obtained at each time point, as illustrated in the second graph on the right. In contrast, patient 114 scored lower at each time point. Similar patterns for patients with endogenous depression are present in the graphs

in Figure 3.5. Patients 104, 106 and 108 were classified as having endogenous depression. In the case of patient 108, the predicted average regression line shows a consistently higher predicted HDRS score over time when these scores are compared to the predicted average regression line for the non-endogenous patients. The observed trajectories of patients 101, 114, 106 and 108 follow their predicted patient-specific regression line more closely than is the case for patients 103 and 104.

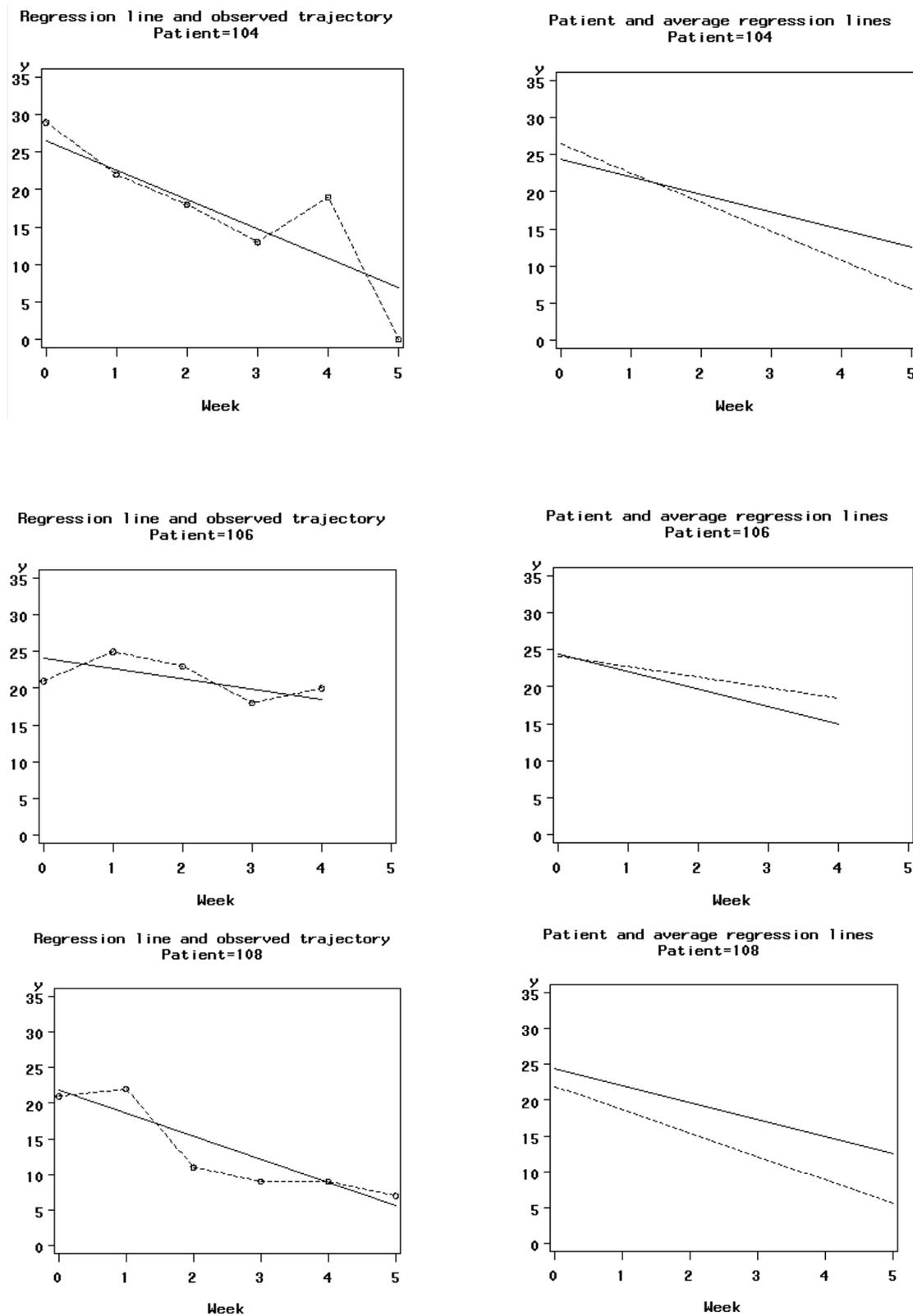


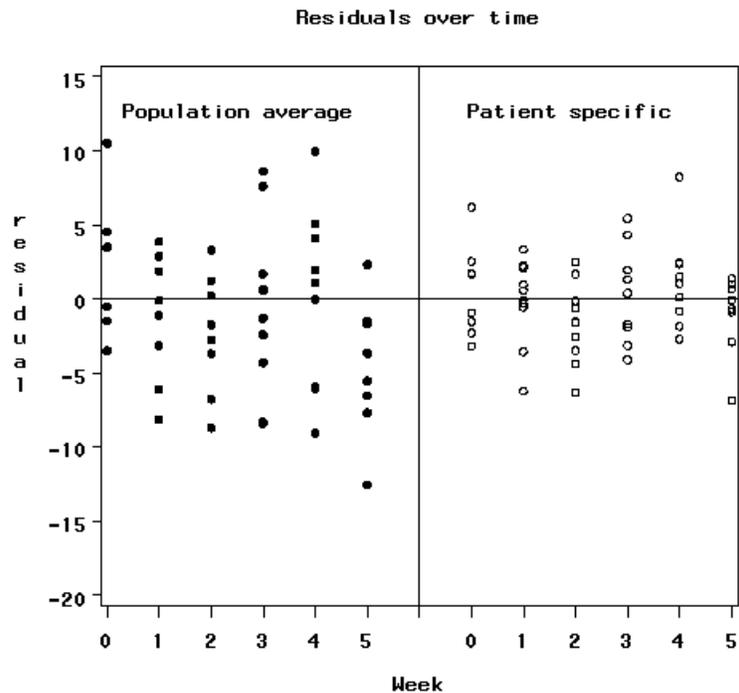
Figure 3.5: Average and unit-specific regression lines for patient with endogenous depression

Level-1 residuals can also be obtained, either for a typical or specific patient, by using the empirical Bayes estimates. The residuals for a typical patient are obtained as

$$\begin{aligned} \text{Average residual} &= \text{Observed HDRS score} - \hat{y} \\ &= \text{Observed HDRS score} \\ &\quad - [22.47626 - 2.36569(\text{Week}) + 1.98802(\text{ENDOG})] - 0.02706(\text{WxENDOG}) \end{aligned}$$

The residuals for a specific patient use the additional information given by the empirical Bayes residuals and have the form

$$\begin{aligned} \text{Patient-specific residual} &= \text{Observed HDRS score} - \hat{y} | \mathbf{u}_i \\ &= \text{Observed HDRS score} - \\ &\quad [22.47626 - 2.36569(\text{Week}) + 1.98802(\text{ENDOG}) - 0.02706(\text{WxENDOG}) + \hat{u}_{i0} + \hat{u}_{i1}(\text{Week})] \end{aligned}$$



**Figure 3.6: Comparison of population average and Bayes residuals for first 10 patients**

Inspection of these estimates can be useful in examining the distributional assumptions for the level-1 data, in this case at the measurement level. For the current example, residuals for a typical patient have a mean of  $-0.022$  and range between  $-15.3792$  and  $20.4995$ . The residuals for specific patients have a mean of  $0.000$  with a minimum value of  $-9.5942$  and a maximum value of  $12.6566$ . The range of the latter is much smaller. Empirical Bayes estimates are frequently referred to as "shrunk" estimates, as the empirical Bayes tend to pull the estimates closer to the sample average, thus "shrinking" them. The amount of shrinkage is a function of the precision of the intercept/slope estimates. The greater the precision for an estimate, the less shrinkage. This means that more shrinkage will occur for units where there is more uncertainty concerning the accuracy of the fixed intercept and slope regression estimates. Figure 3.6 shows a comparison of the two types of

residuals for our example. The Bayes residuals are much closer to the horizontal reference line at 0. Note the shrinking of the more extreme population average residuals in the left pane.

An alternative way to display the Bayes residuals is to plot 95% confidence intervals for patient intercepts and slopes, as shown in Figures 3.7 and 3.8. Information for the first 10 patients in the study were used to construct these graphs.

Confidence intervals were obtained as  $mean \pm 1.96(std. dev)$ . For the intercepts, the means are computed as  $\hat{\beta}_0 + \hat{u}_{0j}$ ,  $j = 1, 2, \dots, 10$ . Standard deviations are computed as the square roots of the variances of the empirical Bayes intercept residuals given in the **ba2** file shown earlier. For patient 1, for example,

$$\begin{aligned} \hat{u}_{01} &= 1.7950, & \text{var}(\hat{u}_{01}) &= 3.9520 \\ \hat{u}_{11} &= -2.0413, & \text{var}(\hat{u}_{11}) &= 0.45969. \end{aligned}$$

Therefore, the 95% confidence intervals for patient 1 are:

$$\begin{aligned} \text{intcept} &: (22.47626 + 1.7950) \pm 1.96 \times (3.9520)^{1/2} \\ &= (20.3749; 28.1677) \end{aligned}$$

$$\begin{aligned} \text{slope} &: (-2.36569 - 2.0413) \pm 1.96 \times (0.45969)^{1/2} \\ &= (-5.7359; -3.0781). \end{aligned}$$

In Figures 3.7 and 3.8, each mean is represented by a square. Figure 3.8 shows that the individual slopes for the first 10 patients are all negative.

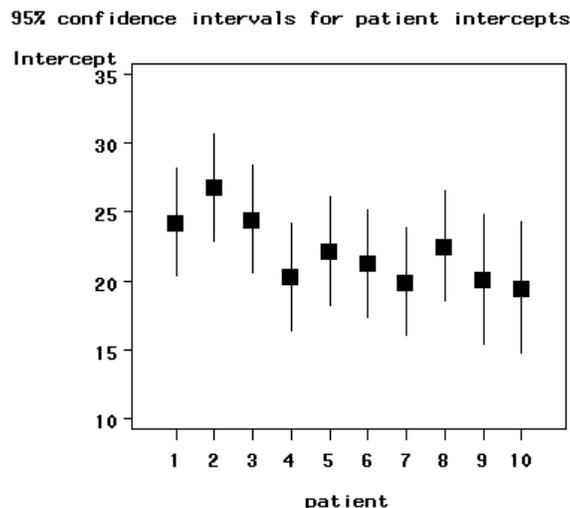


Figure 3.7: 95% confidence intervals for patient intercepts

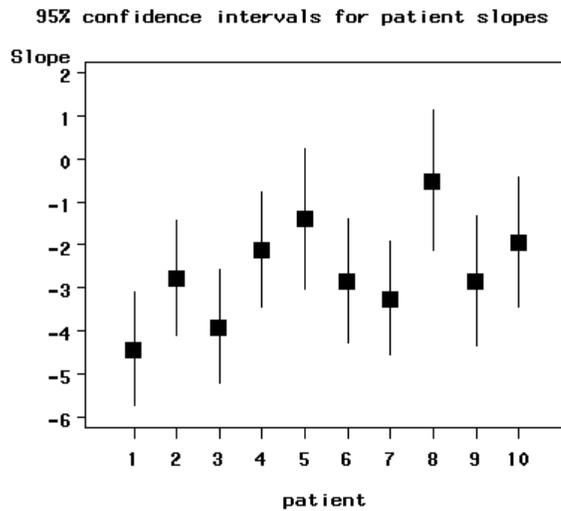
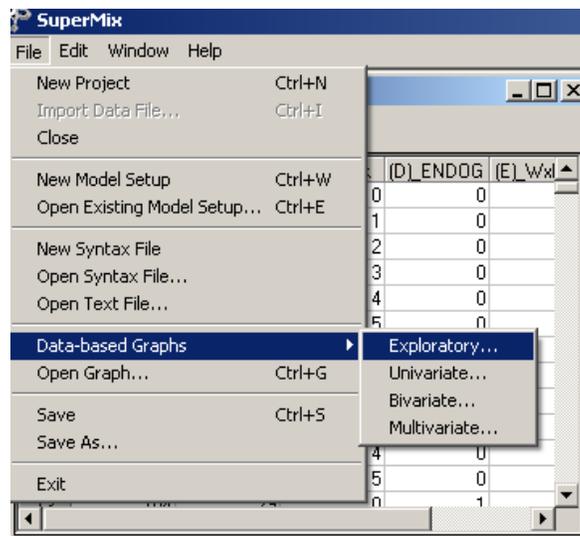


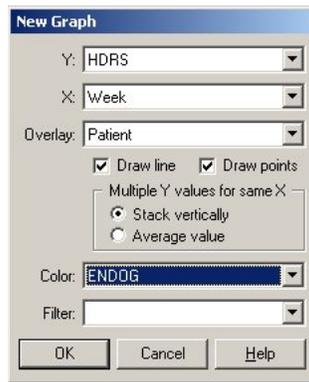
Figure 3.8: 95% confidence intervals for patient slopes

## Graphical displays

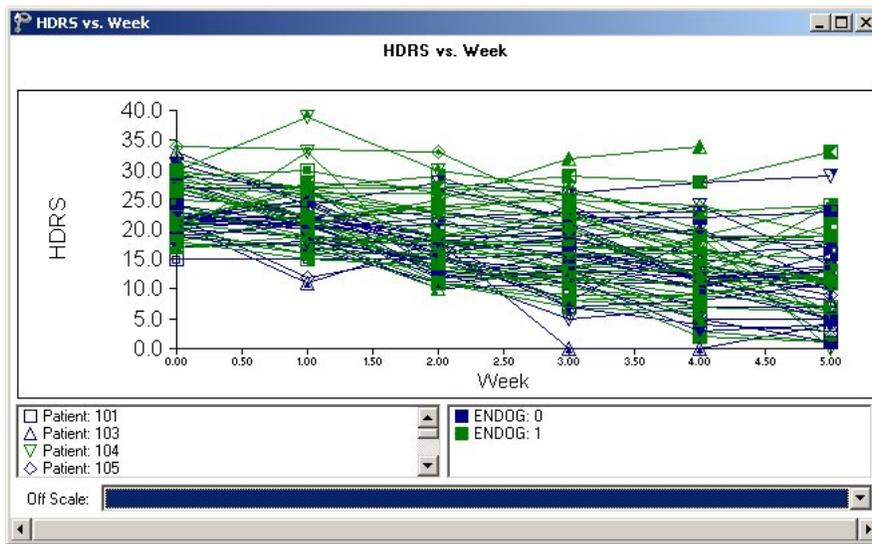
It is possible to obtain a great variety of graphical displays with SuperMix. To invoke the graphics procedure, open a SuperMix data file. To illustrate this, we use **reisby.ss3** in the **Continuous** subfolder. The next step is to select the **Data-based Graphs, Exploratory** option on the **File** menu as shown below



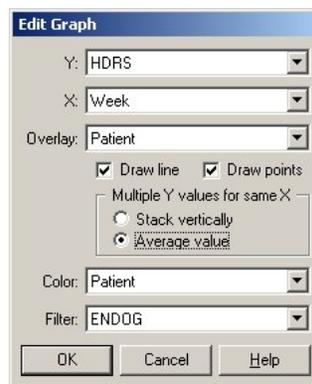
to activate the **New Graph** dialog box. Specify HDRS as the dependent (vertical axis) variable by selecting it from the **Y** drop-down list box and Week as the independent (horizontal axis) variable by selecting it from the **X** drop-down list box. A graph on the same axes-system is created for each patient by selecting the variable Patient from the **Overlay** drop-down list box. Furthermore, each graph is assigned a color by selecting ENDOG from the **Color** drop-down list box to produce the following **New Graph** dialog box.



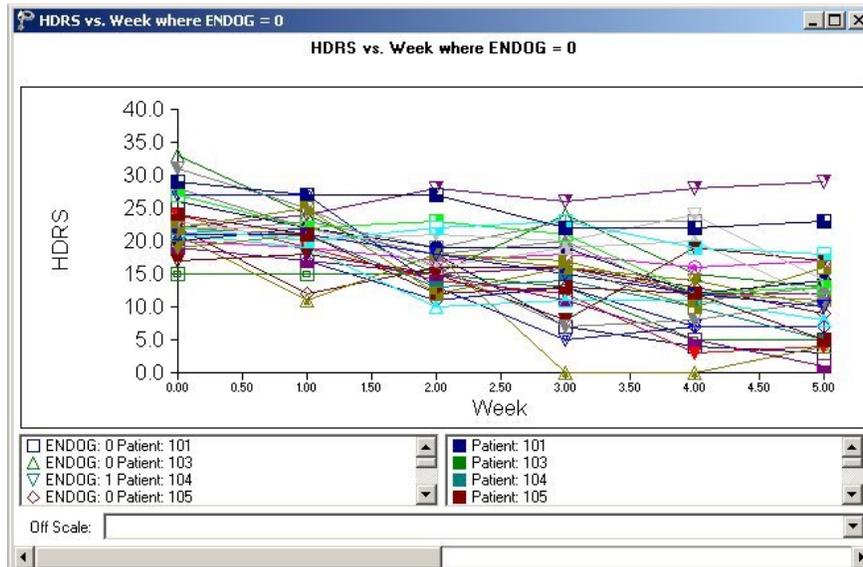
Click on the **OK** button to produce the following graph of the reaction trajectories over time for the 66 patients.



To modify the existing graphic displays, select the **Edit Graph** option from the **Settings** menu to open the **Edit Graph** dialog box. To obtain different graphs for the two categories of the covariate ENDOG, select it from the **Filter** drop-down list box to produce the following **Edit Graph** dialog box.



Click on the **OK** button to open the following graphics window.



At the bottom of the graphics window is a "slider" with left and right arrows. By clicking on the right arrow, one can obtain the next graphic shown below and by clicking on the left arrow, the graphic above.

