



Two-level random intercept and trend model

Contents

3.1.1	<i>The data</i>	1
3.1.2	<i>A 2-level random intercept model and trend model</i>	5

3.1.1 The data

To illustrate the application of the mixed-effects ordinal logistic regression model to longitudinal data, we examined data collected in the NIMH Schizophrenia Collaborative Study on treatment-related changes in overall severity. Specifically, Item 79 of the Inpatient Multidimensional Psychiatric Scale (IMPS; Lorr & Klett, 1966) was used. In this study, patients were randomly assigned to receive one of four medications: placebo, chlorpromazine, fluphenazine, or thioridazine. Since previous analyses (Longford, 1993, and Gibbons & Hedeker, 1994) revealed similar effects for the three anti-psychotic drug groups, they were combined in the present analysis. Finally, again based on previous analysis, a square root transformation of time was chosen to linearize the relationship of the IMPS79 scores over time.

Data for the first 10 observations are shown below in the form of a SuperMix spreadsheet file, named **schizx.ss3**.

 A screenshot of a SuperMix spreadsheet window titled 'SuperMix - [SCHIZX.ss3]'. The window has a menu bar with 'File', 'Edit', 'Window', and 'Help'. Below the menu bar is a text box containing '1103' and an 'Apply' button. The main area is a spreadsheet with 10 rows and 10 columns. The columns are labeled (A) ID, (B) IMPS79, (C) IMPS79D, (D) IMPS79O, (E) DRUG, (F) WEEK, (G) SQRTWEEK, and (H) WSQRTxDRUG. The data is as follows:

	(A) ID	(B) IMPS79	(C) IMPS79D	(D) IMPS79O	(E) DRUG	(F) WEEK	(G) SQRTWEEK	(H) WSQRTxDRUG
1	1103.00	5.50	1.00	4.00	1.00	0.00	0.00	0.00
2	1103.00	3.00	0.00	2.00	1.00	1.00	1.00	1.00
3	1103.00	-9.00	-9.00	-9.00	1.00	2.00	1.41	1.41
4	1103.00	2.50	0.00	2.00	1.00	3.00	1.73	1.73
5	1103.00	-9.00	-9.00	-9.00	1.00	4.00	2.00	2.00
6	1103.00	-9.00	-9.00	-9.00	1.00	5.00	2.24	2.24
7	1103.00	4.00	1.00	2.00	1.00	6.00	2.45	2.45
8	1104.00	6.00	1.00	4.00	1.00	0.00	0.00	0.00
9	1104.00	3.00	0.00	2.00	1.00	1.00	1.00	1.00
10	1104.00	-9.00	-9.00	-9.00	1.00	2.00	1.41	1.41

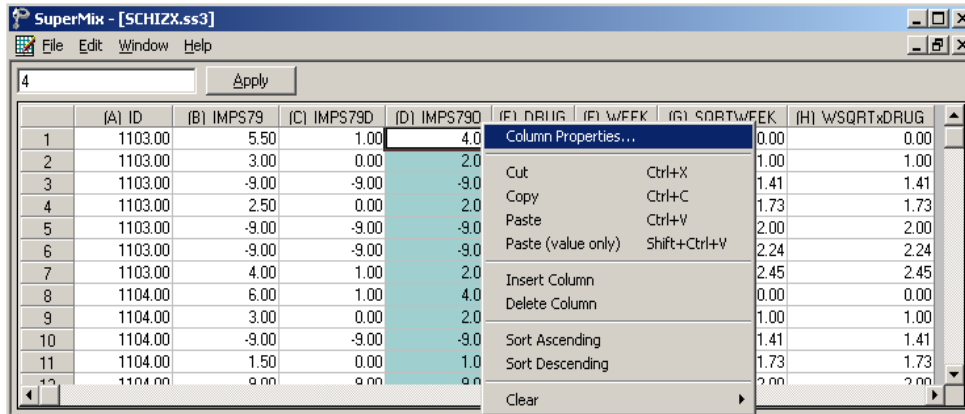
The variables of interest are:

- ID indicates the subject (437 patients in total).
- IMPS79 represents the original score on Item 79 of the Inpatient Multidimensional Psychiatric Scale. It was scored as: 1 = normal, or not at all ill; 2 = borderline mentally ill; 3 = mildly ill; 4 = moderately ill; 5 = markedly ill; 6 = severely ill; and 7 = among the most extremely ill.
- IMPS79D is a recoded version of the same scale, but in binary form, where scores up to, but excluding 3.5 were coded 0, and scores of 3.5 or higher were coded 1. The value "0" is associated with measurements classified as normal, borderline, mildly, or moderately mentally ill, while the value "1" was assigned to measurements corresponding to "markedly ill" through "most extremely ill."
- IMPS79O is also a recoded version of the same scale, but with the 7 original categories reduced to four: 1 = normal or borderline mentally ill, 2 = mildly or moderately ill, 3 = markedly ill, and 4 = severely or among the most extremely ill.
- DRUG indicates the treatment group, where 0 indicates the placebo patients, and 1 refers to the drug patients.
- WEEK represents the time during the course of the study when a specific measurement was made, and ranges between 0 and 6.
- SQRTWEEK is the square root of WEEK. This variable is generated within the SuperMix spreadsheet. For more information on data manipulation, please refer to Section 2.5.
- WSQRTxDRUG is the product of the treatment group and the square root of WEEK.

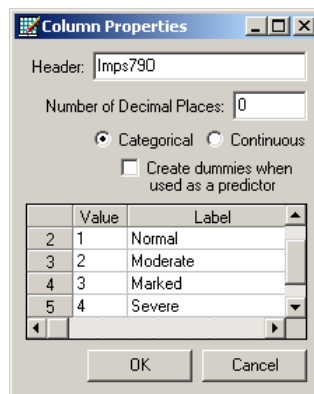
In this data file, each subject's data consist of seven lines, these being the repeated measurements on seven occasions. Notice that there are missing value codes (-9) for some subjects at specific time points. The data from these time points will not be used in the analysis, but data from these subjects at other time points where there are no missing data will be used in the analysis. Thus, for inclusion into the analysis, a subject's data (both the dependent variable and all model covariates being used in a particular analysis) at a specific time point must be complete. The number of repeated observations per subject then depends on the number of time points for which there are non-missing data for that subject. The specification of missing data codes will be illustrated in the model specification section to follow.

3.1.1.1 Defining column properties

Defining column properties for the ordinal data is recommended. We use the column of IMPS79O as an example. First, highlight the column of IMPS79O by clicking on its header. Then right click and select the **Column Properties** option as shown below to open the **Column Properties** dialog box.



The header of the **Column Properties** dialog box indicates the current variable name. Keep the default number of decimal places unchanged. Select the **Categorical** radio button to activate the grid field to enter the labels for each category as shown below.



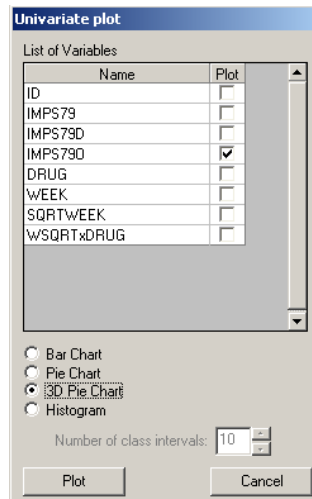
Click on the **OK** button and save the change to the data set by clicking on the **File, Save** option.

The outcome variable: univariate graphs

As a first step, we take a look at the ordinal variable IMPS790 which is the potential dependent variable in this study.

Pie chart

To generate a pie chart for IMPS790, first open the **schizx.ss3** SuperMix spreadsheet. Next, select the **File, Data-based Graphs, Univariate** option to load the **Univariate plot** dialog box. Select the variable IMPS790 and indicate that a **3D Pie Chart** is to be graphed as shown below.



Click the **Plot** button to display the following pie chart.

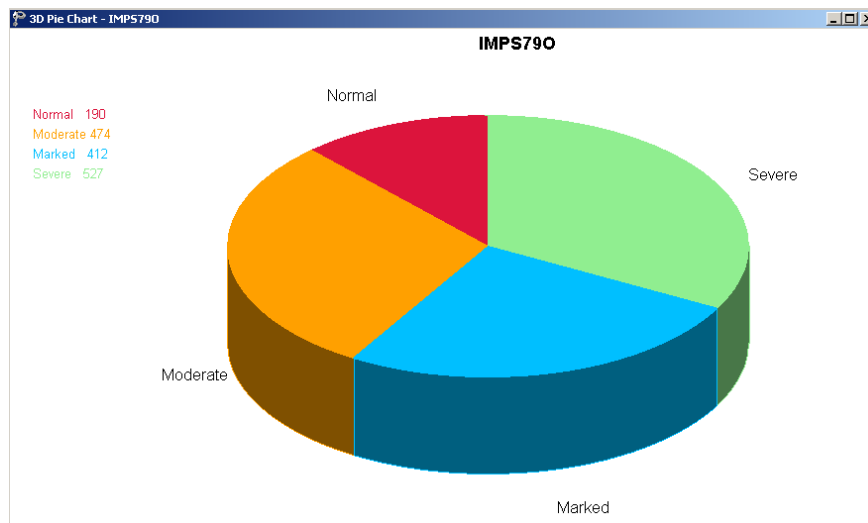


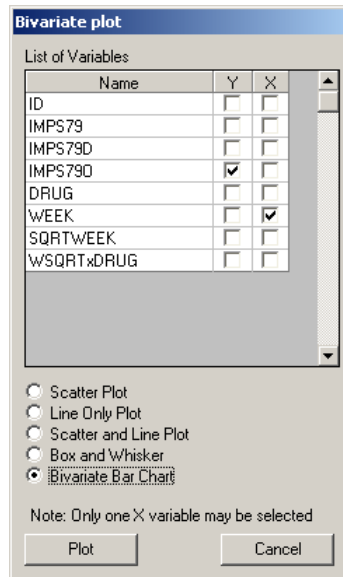
Figure 6.6: Pie chart of IMPS790 values

Note that most of the observations fall into the Severe illness category. Keep in mind that the pie chart takes all observations, regardless of the time of measurement, into account. As such, it is informative about the distribution of all observed values of the potential outcome, but does not provide any information on possible trends in illness level over time.

Relationships between variables: bivariate bar chart

It is hoped that the severity of the illness (IMPS790) will decrease over the treatment period. Before considering fitting a model to these data, we would like to explore the relationship between IMPS790 and WEEK using a bivariate bar chart.

Bivariate bar chart



A bivariate bar chart is accessed via the **Data-based Graphs, Bivariate** option on the **File** menu. The **Bivariate plot** dialog box is completed as below: select the outcome variable IMPS790 as the **Y**-variable of interest, and the predictor WEEK to be plotted on the **X**-axis. Check the **Bivariate Bar Chart** option, and click **Plot**.

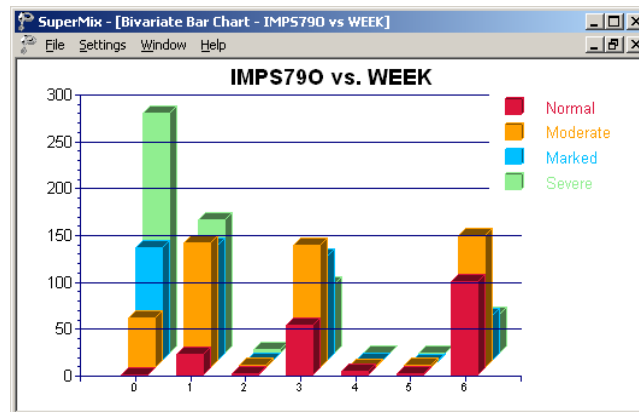


Figure 6.7: Bar chart of IMPS790 vs. WEEK

As shown above, most patients did not participate in the study at weeks 2, 4 and 5. At the beginning of the study (week 0), a large percentage of patients are markedly or severely ill. By the end of the study (week 6), most patients are reported as normal or moderate.

3.1.2 A 2-level random intercept model and trend model

In this section, we fit a model with random intercept and slope. To do this, the level-1 model is unchanged; only the level-2 model is modified.

3.1.2.1 The model

Level 1 model:

$$y_{ij} = \log\left(\frac{P_{ijc}}{1-P_{ijc}}\right) = \gamma_c - [b_{0i} + b_{1i}\text{DRUG}_i + b_{2i}\text{SQRTWEEK}_i + b_{3i}(\text{WSQRT}\times\text{DRUG})_i],$$
$$j = 1, \dots, n_i; c = 1, 2, \dots, C-1$$

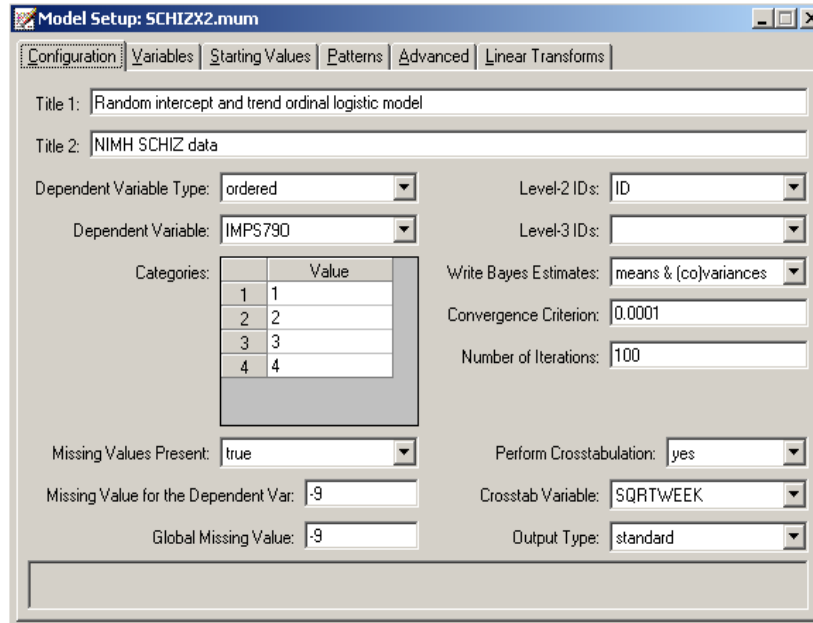
Level 2 model:

$$b_{0i} = \beta_0 + v_{0i}, \quad i = 1, \dots, N$$
$$b_{1i} = \beta_1$$
$$b_{2i} = \beta_2 + v_{2i}$$
$$b_{3i} = \beta_3$$

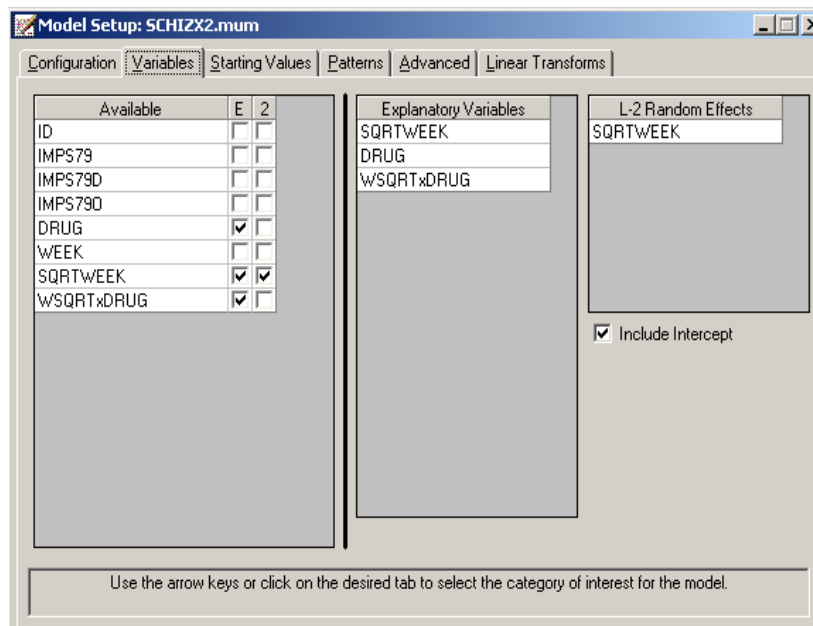
As shown above, the slope of the time variable b_{2i} is now estimated by a level-2 equation containing both a fixed and a random effect. It indicates that patients are now not only assumed to have different intercepts, but may also exhibit different responses to the treatment over time.

3.1.2.2 Setting up the analysis

Use the **File, Open Spreadsheet** option to re-open the previously used spreadsheet **schizx.ss3** from the **Examples\Ordinal** folder. Next, use the **File, Open Existing Model Setup** option to locate and open the syntax file **SCHIZX1.mum**. Click on **File, Save as** to save the model setup in a new file, such as **SCHIZX2.mum**. Next, we change the string in the **Title 1** text box on the **Configuration** screen (optional). Request a crosstabulation of the variable **SQRTWEEK** by the response variable **IMPS790** by selecting the **yes** option from the **Perform Crosstabulation** drop-down list box, followed by the selection of **SQRTWEEK** as the **Crosstab Variable**.



Proceed to the **Variables** tab, and check the **2** check box for SQRTWEEK to select it as a level-2 random variable as shown below.



Keep all the other settings unchanged. Save the changes to the file **SCHIZX2.mum** and click the **Analysis, Run** option to produce the output file **SCHIZX2.out**.

3.1.2.3 Discussion of results

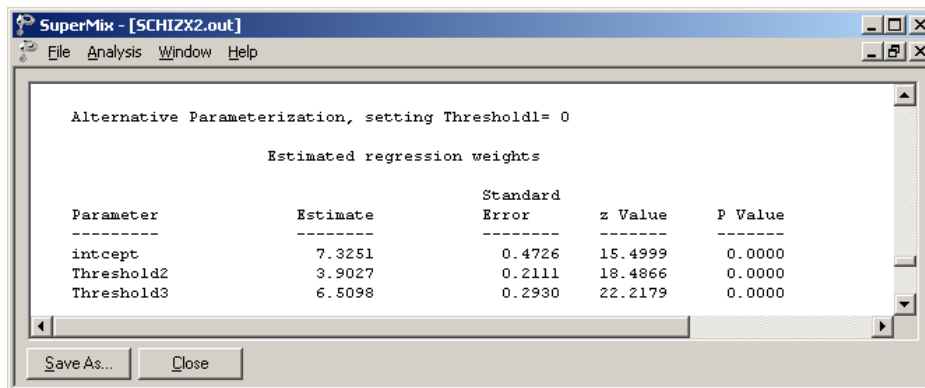
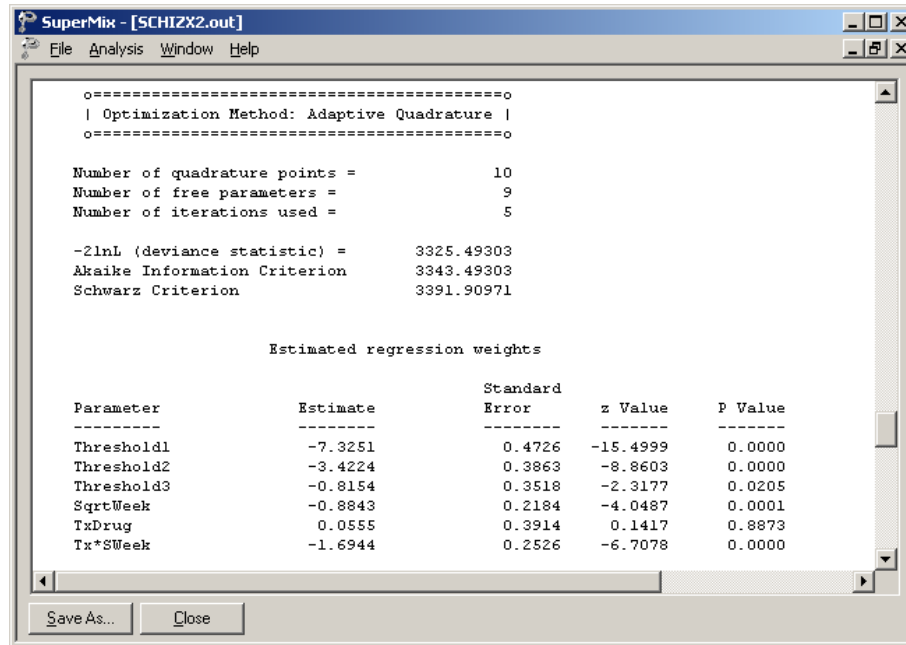
Crosstabulation

The following portion of the output is a crosstabulation of the seven distinct values of the variable SQRTWEEK by the four categories of the outcome variable IMPS790. We note that there are relatively few observations for the third, fifth and sixth weeks. For example, for week 5 (SQRTWEEK = 2.24), measurements on only 9 of the 437 patients are available. Looking down the columns (SQRTWEEK) we see the severity of symptoms (IMPS790) declining.

Y-Category	X-Category	Frequency
1	0.00	1
1	1.00	23
1	1.41	3
1	1.73	54
1	2.00	5
1	2.24	3
1	2.45	101
2	0.00	54
2	1.00	135
2	1.41	4
2	1.73	132
2	2.00	3
2	2.24	4
2	2.45	142
3	0.00	122
3	1.00	124
3	1.41	2
3	1.73	113
3	2.00	2
3	2.24	0
3	2.45	49
4	0.00	257
4	1.00	144
4	1.41	5
4	1.73	75
4	2.00	1
4	2.24	2
4	2.45	43
Total:		1603

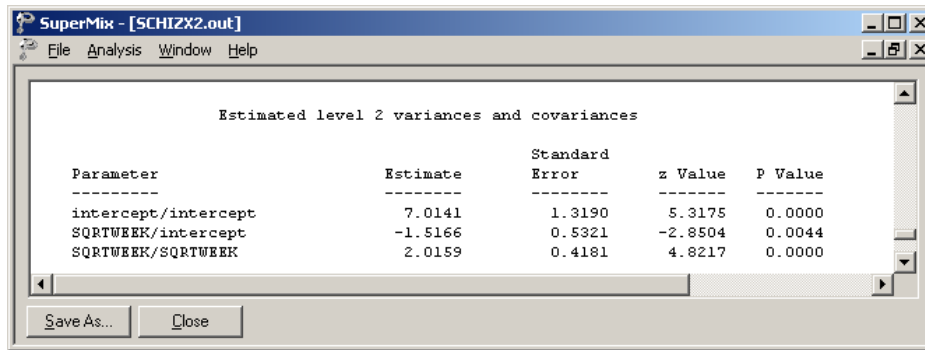
Fixed effect results

The final results after 23 iterations are listed below. While the values of the estimated coefficients differ from those in the random-intercept-only model, the overall picture remains very similar. The decline in severity over time noticed in the crosstabulation is captured by the significant fixed effect coefficient of -0.88295 for SQRTWEEK.



Random effects results

Note that the estimated coefficient for the random SqrtWeek slope is highly significant, indicating that patients not only start at different points but follow different paths during the treatment period.



3.1.2.4 Interpreting the output

Estimated outcomes for groups: unit-specific results

To evaluate the expected effect of the treatment group and the square root of time of treatment, while allowing for the interaction between treatment and the square root of time, we use the expression below:

$$\log\left(\frac{\hat{P}_{ijc}}{1-\hat{P}_{ijc}}\right) = \hat{\gamma}_c - \left[\hat{b}_{0i} + \hat{b}_{1i} \text{DRUG}_i + \hat{b}_{2i} \text{SQRTWEEK}_i + \hat{b}_{3i} (\text{WSQRT} \times \text{DRUG})_i \right]$$

so that

$$\hat{\eta}_{ijc} = \hat{\gamma}_c - 7.3793 + 0.0553 \times \text{DRUG}_i + 0.8841 \times \text{SQRTWEEK}_i + 1.6940 \times (\text{WSQRT} \times \text{DRUG})_i$$

As illustrated in the previous example, by substituting the values for DRUG, SQRTWEEK and WSQRTxDRUG, the results shown in Table 6.9 can be obtained.

Table 6.9: Estimated unit-specific results for random intercept & slope model

severity	Placebo patients (drug = 0)				Drug patients (drug = 1)			
	normal	moderate	marked	severe	normal	moderate	marked	severe
week 0	0.06%	2.96%	26.90%	70.08%	0.07%	3.13%	27.90%	68.91%
week 1	0.15%	6.87%	43.81%	49.17%	0.86%	29.42%	55.32%	14.40%
week 2	0.22%	9.61%	50.03%	40.15%	2.47%	51.90%	39.98%	5.81%
week 3	0.29%	12.32%	53.77%	33.62%	5.42%	68.72%	23.37%	2.49%
week 4	0.36%	15.09%	55.99%	28.55%	10.27%	74.85%	13.62%	1.26%
week 5	0.45%	17.94%	57.12%	24.49%	17.38%	73.94%	7.99%	0.69%
week 6	0.54%	20.84%	57.44%	21.17%	26.72%	68.08%	4.80%	0.40%

We can again represent the results from the above table graphically, as shown in Figures 6.10 and 6.11. The graphs tell us the same story as the previous model: patients from the treatment group showed more improvement over time than patients from the control group. While a

very small proportion of treatment patients were still diagnosed as being severely ill at the end of the treatment period (0.42% according to table 6.9), 20% of the control group were still classified as being severely ill by week 6.

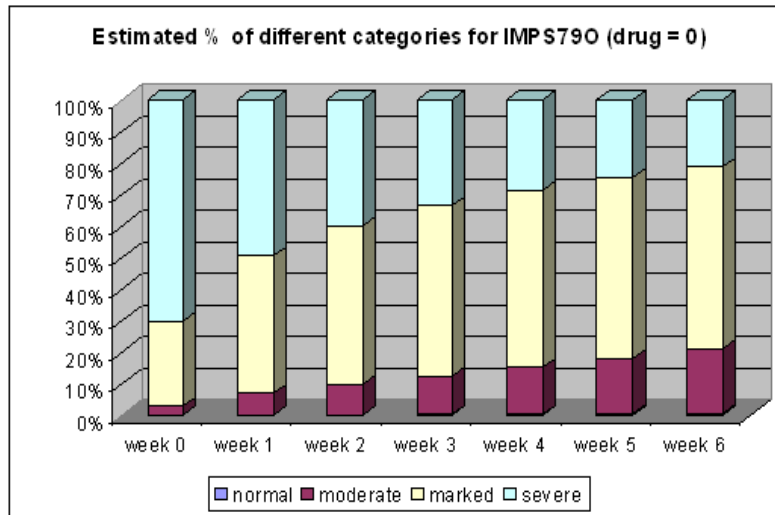


Figure 6.10: Estimated percentage of patients over time (treatment group)

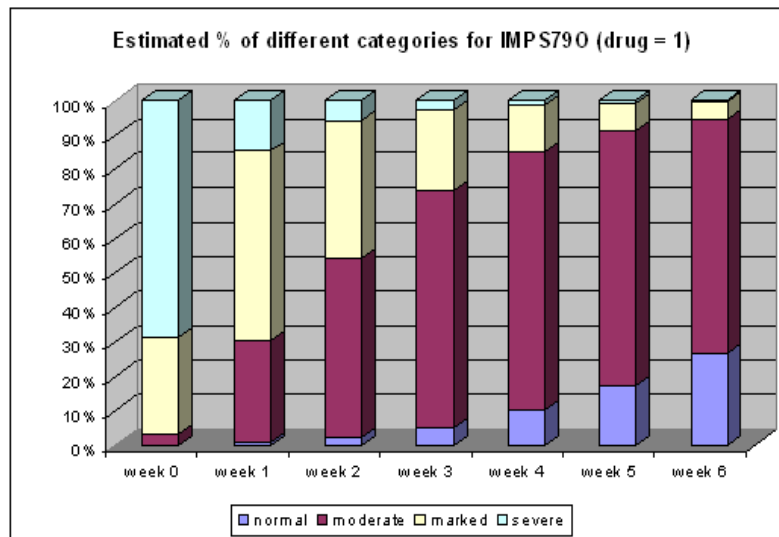


Figure 6.11: Estimated percentage of patients over time (treatment group)

Estimated time trend variance

When we consider the heterogeneity in responses across time, we notice that the estimated variance in the time trend is $\sigma_{v_1}^2 = (1.29774)^2 + (-0.57054)^2 = 2.0096$. The estimates for the time trends are -0.88295 for SQRTWEEK and -1.69416 for WSQRTxDR respectively. Thus the estimated trends for the placebo and drug groups are -0.88295 and

$-0.88295 - 1.69416 = -2.57711$. Thus the 95% confidence interval of the time trend for the placebo group is $-0.88295 \pm (1.96 \times \sqrt{2.0096}) = (-3.6615, 1.896)$. Similarly, the confidence interval for the drug group is $(-5.3556, 0.2014)$. Notice that both intervals are fairly large and include negative and positive slopes, which reflects the wide heterogeneity in trends. The estimated correlation value is -0.402 , which is moderately large. This indicates that the patients who are initially less severely ill improve at a smaller rate. The more severely ill patients improve at a greater rate.