A REGRESSION ANALYSIS OF MENTAL HEALTH DATA

by

ROBYN ATTEWELL

A Case Study submitted to the Australian National University towards a Master of Science Degree by coursework in the Department of Statistics, Faculty of Economics.

November 1982

I declare this material to be my own work except where stated otherwise.

R.G. Attewell

R. G. Attewell

Acknowledgements

I wish to thank Paul Duncan-Jones for making this problem available as a case study and for his encouragement and guidance in his capacity as my supervisor.

I also had some useful discussions with Professor Heathcote.

The C.S.C. Univac computer played an important role in this work and I acknowledge it's contribution to relieving the computational burden and the task of producing this manuscript.

Table of contents

1. Introduction to the problem

This study is concerned with investigating the associations between neurosis, a person's social contact, and exposure to adverse circumstances. Psychiatric illnesses which are classed as psychotic, (eg. schizophrenia), involve major functional disorder. Much research has been concerned with these illnesses since data are readily available from institutions and psychiatric consultations. Non-psychotic illnesses such as emotional disorders, anxiety and depression are harder to define and recognise and are much more widespread. It is this form of mental illness as identified in the general community that is the focus of this case study.

That there exists an association between lack of social and personal ties and mental disturbance is generally accepted. The direction of causality, that is, whether little social contact contributes to neurosis, or neurotic symptoms affect an individual's social interaction, or whether they are both influenced by a third factor(eg. personality), is still under investigation. Important and stressful events occurring in a person's life are also recognised as contributing to mental disturbance. But whether there is a significant interaction with social support is under current debate as well.

It has been shown in studies by Henderson et al (1980) and Henderson (1981) that deficiencies in social relationships and, more importantly, assessment of existing relationships as inadequate, are more strongly associated with subsequent symptoms under conditions of adversity. Brown and Harris(1978) made similar findings for the onset of depression in women in their study.However Miller and Ingham(1976) suggest that deficiencies in social bonds act independently of life events in the aetiology of neurotic illness. Brown et al (1975) identified four factors influencing the effects of stressful events in causing depression. These were, having an intimate confidant, loss of mother at an early age, having many young children at home, and lack of employment. However it was admitted that a sample of 220 women of whom only 10% were recent cases, is too small to determine whether significant interactions exist.

The investigation of interaction and comparisons between studies are made difficult by the problem of definition and measurement of the quantities involved. Though the data set of this study was gathered for purposes other than an intensive study of psychiatric epidemiology, it does have the advantage of being large. This offers the possibility of a more definitive statement about interaction than previous studies have provided. The main focus of the present investigation is to determine whether a model incorporating interaction between adversity and social support will account for materially more of the variance in a measure of neurotic illness than a model of independent effects, additive in the appropriate metric.

2. The Data and the Approach

In 1975 the Australian Bureau of Statistics conducted a health care survey for the Health Commission of NSW. People from approximately 3100 selected households in the Gosford,Wyong and Illawarra regions of NSW were asked quite detailed questions on all aspects of their health and use of local health services. The survey was the first large-scale household survey of morbidity and health services in Australia, and was undertaken primarily to assess the adequacy of health services in these areas.

We are grateful to the NSW Health Commission for making the data tapes available for secondary analysis. The data subset relevant to this study involved the responses to the mental health section of the questionnaire by 6067 adults (15 years of age and over).Patients in hospitals,convalescent homes and institutions were not included in the survey.

Although the overall sample design involved several stages and sampling techniques, simple random sampling was assumed in this analysis. Whereas the original analysis was concerned with estimation of prevalence of neurosis, our aim is to test hypotheses about the relationships between neurosis and other factors using techniques not appropriate to smaller sets of data.

Results from empirical studies by Kish and Frankel(1970) for similarly complex surveys have shown that the design effects computed for estimating standard errors using balanced repeated replication tend to be smaller for analytical statistics such as regression coefficients compared with aggregates and means.

Another detailed study by Landis(1982) for an American survey similar to this, has shown that effects found to be non-significant (in regression, anova and contingency table analysis) under the simplest assumptions are unlikely to be found significant if sampling weights and the design structure are taken into account.

In this study, given the sample size, marginally significant effects are of lesser scientific interest, and one would already be cautious of interpreting them, even if assured of the validity of the assumptions required for the significance tests.

2.1. Measuring Disturbance

The measure of neurosis available from the survey results was the score obtained on the 12 item version of the General Health Questionnaire (Goldberg 1972) , which is listed below.

Each question was answered with one of four graded responses. These are worded according to the question but follow the same general form. For example, the response set for question 1. is as follows-

(i) better than usual
(ii) same as usual
(iii) less than usual
(iv) much less than usual

General Health Questionnaire

1. Have you recently been able to concentrate on whatever you're doing?

2. Have you recently lost much sleep over worry?

3. Have you recently felt that you are playing a useful part in things?

4. Have you recently felt capable of making decisions about things?

5. Have you recently felt constantly under strain?

6. Have you recently felt that you couldn't overcome your difficulties?

7. Have you recently been able to enjoy your normal day-to-day activities?

8. Have you recently been able to face up to your problems?

9. Have you recently been feeling unhappy and depressed?

10. Have you recently been losing confidence in yourself?

11. Have you recently been thinking of yourself as a worthless person?

12. Have you recently been feeling reasonably happy, all things considered?

The responses were dichotomised to 0 for the 'normal' response and 1 for the 'pathological' response, and then summed to give a score ranging from 0 to 12. For example, a negative response to question 8 would contribute 1 unit to the overall score. This follows the normal scoring convention for this questionnaire. The total scores will subsequently be referred to as GHQ scores.

Assuming that neurotic disturbance can be measured on a single continuous axis ranging from hypothetical normality to severe disturbance, the score is interpreted as a quantitative estimate of the individual's degree of disturbance.The higher the score,the more neurotic the individual.In the original analysis for the Health Commission a score of 0 or 1 was taken to be normal.

Extensive validation studies have been carried out on the 60 and 30 item versions of this questionnaire, and justification of its form and content as an instrument for measuring non-psychotic disturbance is discussed at length in Goldberg(1972). It is shown that the 12 questions chosen for the smallest version have the best discriminatory power.

The GHQ shares shortcomings with other self-administered questionnaires in depending on the individual to understand the form of the questions and answer them accurately. By continually referring to a person's current state, chronically disturbed people may not score highly enough to be distinguished from the normal people. It is already popular, however, and its advantages include its ease of use, and objectivity since no interviewer assessment is required.

2.2. Social Support

The 9 questions of the survey selected for a social bonds measure are listed below. The possible responses were 0, 1-2, or 3+ .

Questions Pertaining to Social Support

1. How many people (friends,family or neighbours) would you find it easy to call on for PRACTICAL help in an emergency or crisis?

2. How many people would you find it easy to call on for EMOTIONAL help in an emergency or crisis?

3. How many times in the past week would you have talked to or corresponded with any of your RELATIVES not living with you?

4. How many times in the past week would you have talked to your NEIGHBOURS?

5. How many times in the past week would you have talked to or corresponded with your FRIENDS?

How many times have you attended or participated in the following activities in the past month?

6. SOCIAL or LICENCED CLUB (e.g. Leagues or R.S.L.)?

7. SPORTING or RECREATION group (e.g. team member or organiser)?

8. CHURCH SERVICE or MEETING?

9. any other COMMUNITY SERVICE, SPECIAL INTEREST GROUP or UNION MEETING (e.g. APEX, hobby class, school p. & c.)?

The answers to these questions (SS1 - SS9) convey information on the amount of interaction a person has with friends, relatives, neighbours, and the extent of participation in community activities. These are fairly crude measures, especially when compared with the detailed Interview Schedule for Social Interaction (Henderson, Byrne, Duncan-Jones 1981), which goes much further than 'counting heads' and includes personal assessments of the adequacy of friendships and social integration. However they appear adequate in the context of this large multi-purpose survey.

2.3. Adversity

Life events are defined as events which occur in a person's life which may cause stress. They usually involve significant changes in health or way of life. The survey included a list of 36 life events (e.g. marriage,divorce,death in the family,moving house,major physical illness, loss of job etc.) Each individual was required to indicate which of these they had experienced in the periods

(i) in the last 2 months
(ii) 3-6 months
(iii) 7-12 months
(iv) 12-24 months

Information of this kind is highly susceptible to recall error and since it is not each particular event that we are interested in, the information was summarised into 4 life event variables representing the total number of life events recalled by a person for each time period. These are subsequently referred to as LE1,LE2,LE3 and LE4.

With this construction each life event is given equal weight within each time period, yet it is plausible that some events may have more impact than others and some individuals may place more importance on particular events depending on their sex, age, social class or cultural background. The concensus of the current literature is that simple frequency counts of life events are inadequate measures of adversity and many methods have been proposed to best quantify their impact. However, no additional information was available in these survey results on the reactions of each individual to experiencing the events.

2.4. Method

The correlation structure of the variables indicated the general trend of relationships between these measures. Over the 6067 individuals, positive correlations were observed between GHQ score and all 4 life event variables (max 0.13 for LE1), and negative correlations for GHQ score and all 9 SS variables (-0.12 for SS5-friends), giving support to the basic hypotheses of associations between neurosis, the occurrence of stressful life events, and lack of social contact.

As this is a cross-sectional study, we cannot investigate the direction of causality. In the following regression analysis GHQ score will be modelled as the response variable, but it is recognised that neurotic symptoms may also influence a person's social behaviour, or cause events such as marriage breakup or loss of one's job.

A multiple regression of GHQ score on the 13 variables could explain only 7% of the variance and examination of the fit revealed that for large GHQ score the residuals were large and positive and became smaller with decreasing GHQ score. This may be attributed to the skewness of the distribution of the scores as illustrated by the frequencies in Table 2-1 and Plot 2.1.

Table 2-1. GHQ Score Frequencies

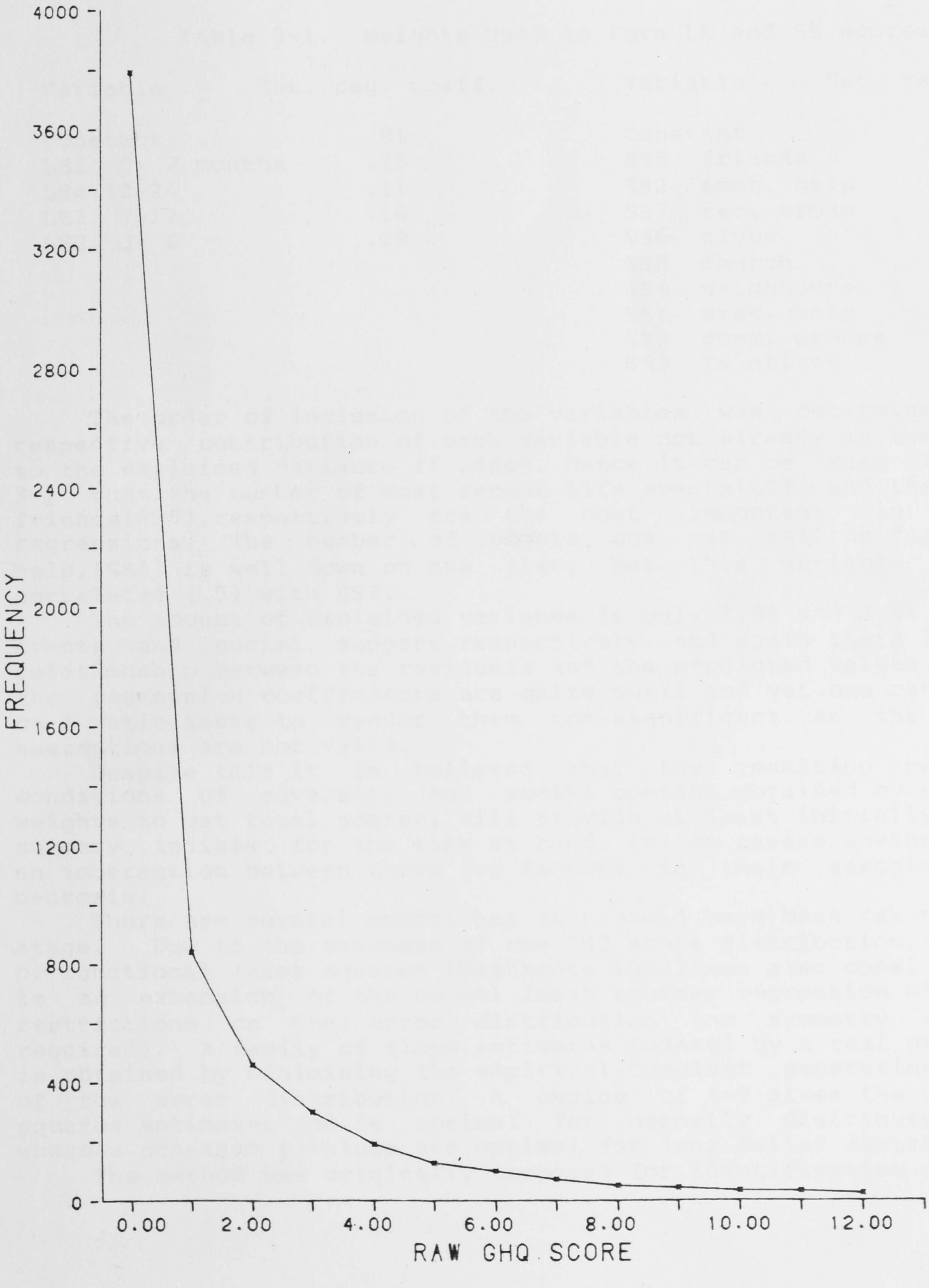
GHQ score	Frequency	% of Total
0	3791	62
1	841	14
2	460	7.6
3	300	4.9
4	192	3.2
5	128	2.1
6	100	1.7
7	72	1.2
8	51	.71
9	43	.71
10	33	.54
11	32	.53
12	24	.40
	6067	

The subsequent analysis tries to accommodate for this by
(i) selecting a threshold GHQ score to classify an individual as either normal or a case and using logistic regression techniques
(ii) dividing the GHQ score into several contiguous ordered categories and using a multivariate analogue to Generalised Linear Models as described by McCullagh(1980). Then parallel logistic regression lines are fitted with intercepts dependent on the category bounds.
(iii) modelling the counts in the 13 categories of GHQ score directly using exponential regression or using the (discrete) proportional hazards non-parametric approach of survival analysis.

In order to obtain a workable model for testing both general main effects and the interaction between social bonds and adversity, the information in the 4 LE and 9 SS variables was condensed into a single LE score and a single SS score for each individual. These scores were then divided into several contiguous categories and cross-classified with GHQ score to form 2-way tables for testing for main effects separately, and 3-way tables for testing between additive and interactive models. The next chapter explains in more detail the construction of the scores.

The analysis of each main effect using the dichotomous and multi-response models is treated in Chapter 4 and Chapter 5 contains the results of testing for interaction. In Chapter 6 the original variables are considered separately and the third approach is employed. A discussion and conclusion follow this.

PLOT 2.1 DISTRIBUTION OF GHQ SCORE



3. Data Reduction

3.1. Construction of Adversity and Social Support Scores

A single adversity score for each individual was obtained by using the regression coefficients from the multiple regression of the raw GHQ score on the 4 LE variables. The weights for a social support score based on the 9 SS variables were obtained on a similar but separate regression. Table 3-1 lists the estimated coefficients.

Table 3-1. Weights Used to Form LE and SS scores

Variable	Est. reg. coeff.	Variable	Est. reg. coeff.
constant	.81	constant	3.6
LE1 0- 2 months	.25	SS5 friends	-.27
LE4 12-24	.11	SS2 emot. help	-.20
LE3 7-12	.16	SS7 rec. group	-.16
LE2 3- 6	.09	SS6 clubs	-.14
		SS8 church	-.13
		SS4 neighbours	-.08
		SS1 prac. help	-.12
		SS9 comm. groups	-.03
		SS3 relatives	-.02

The order of inclusion of the variables was determined by the respective contribution of each variable not already in the equation, to the explained variance if added. Hence it can be seen from Table 3-1 that the number of most recent life events(LE1) and the number of friends(SS5),respectively are the most important in the two regressions. The number of people one can call on for practical help,(SS1) is well down on the list, but this variable is highly correlated (.5) with SS2.

The amount of explained variance is only 2.8% and 3.4% for life events and social support respectively, and again there is a marked relationship between the residuals and the predicted values. Some of the regression coefficients are quite small and yet one cannot depend on F ratio tests to render them non-significant as the normality assumptions are not valid.

Despite this it is believed that the resulting measures of conditions of adversity and social contact obtained by using these weights to get total scores, will provide at least initially adequate summary indices for the task at hand: ie. to assess whether there is an interaction between these two factors in their association with neurosis.

There are several approaches that could have been taken at this stage. Due to the skewness of the GHQ score distribution, the method of functional least squares (Heathcote 1982) was also considered. It is an extension of the normal least squares regression with minimum restrictions on the error distribution (no symmetry or moments required). A family of slope estimates indexed by a real parameter t, is obtained by minimising the empirical cumulant generating function of the error distribution. A choice of t=0 gives the usual least squares estimates and is optimal for normally distributed errors, whereas non-zero t values are optimal for long tailed distributions.

The method was originally proposed for identification of outliers with respect to least squares. Applied to this data set the procedure confirms the lack of normality but the convergence is slow and the resulting coefficients are extremely small and not readily interpretable. This is most probably due to the discrete nature of the data, especially the 3 value range of the SS variables, and justifies the contingency table approach discussed in subsequent chapters.

A drawback to the use of factor analysis at this stage, and selecting a single LE factor and a single SS factor is that some aspects of the data may be overlooked, which are important in accounting for the association with the dependent variable. Providing the two sets of variables are orthogonal the separate regression scores should prove suitable. The data supports this with the absence of significant correlations between the 4 LE variables and the 9 SS variables. Of these 36 correlations the two 'largest' are -.06 between LE2 and SS4 (neighbours), and 0.05 between LE4 and SS1 (emotional help). One may have expected LE1 to be the most likely to have an association since the social support questions were specifically about the previous week and month.

A more formal check of independence was made by calculating the canonical variates for these two sets of variables (ie. finding the two linear combinations of each set which maximises the correlation between the scores). The computations gave a life event canonical variate which could explain only 1% of the variance of the social support canonical variate (and vice versa).

Further evidence for orthogonality was given by the increase in R squared (regression ss/ GHQ ss) remaining at about 0.03 when the four LE variables were included in the regression, whether the nine SS variables were in the equation or not.

3.2. Distributions of these scores

The scores were divided into nine equal sized categories so that the minimum and maximum defined the lowest and highest bounds. Their distributions over the whole data set are indicated by the frequency table (3-2) and plots (3.1,3.2).

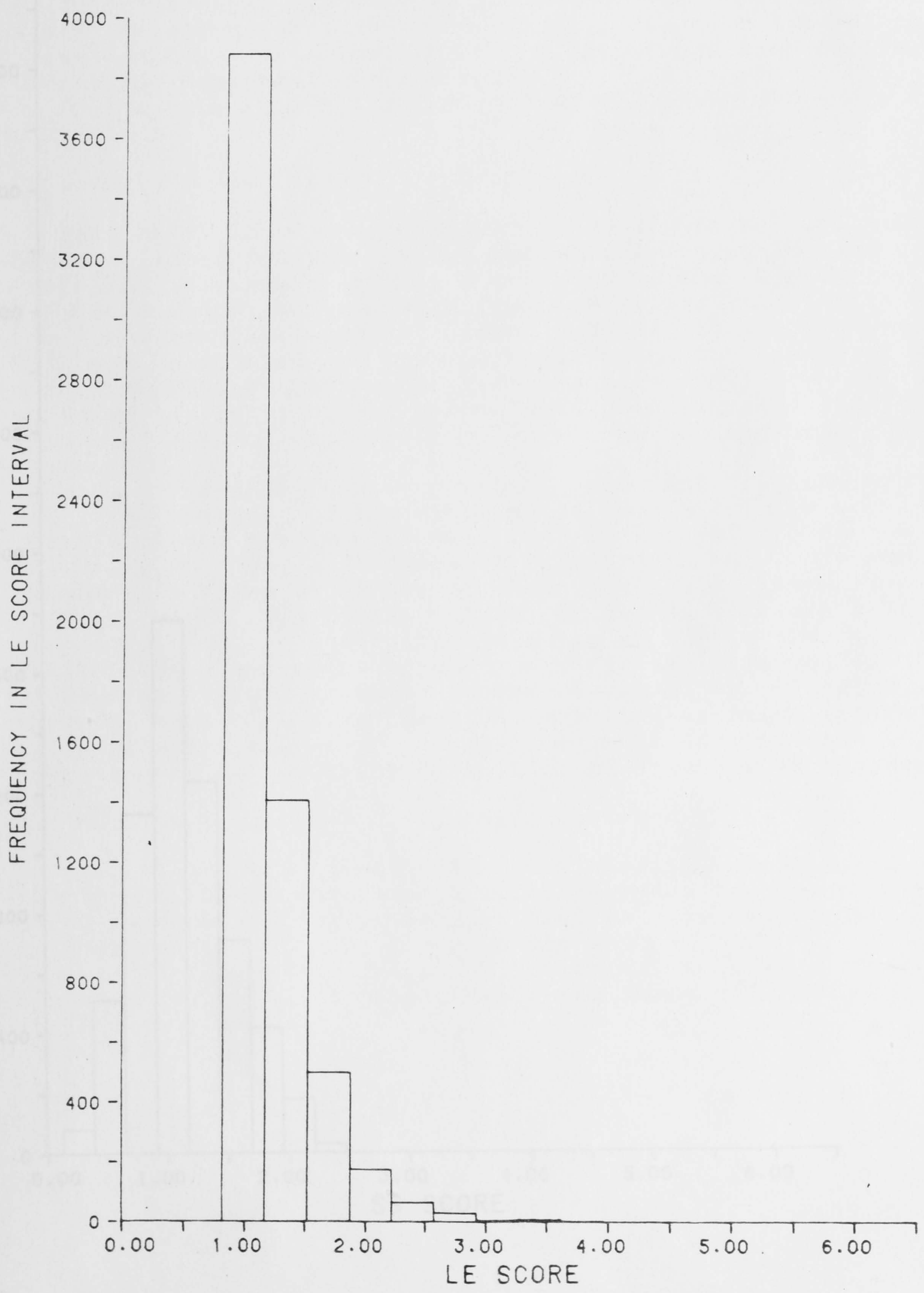
The skewed nature of the LE score reflects the skewness of both the GHQ score distribution and the distibution of the 4 LE variables. Most people recorded few life events occurring in each time period. In fact about 55% of the sample recorded no life events for 1 to 2 years and about 69% for each of the other time periods. The skewness parameters for the categorised forms of the LE and SS scores are 2.4 and .39 respectively, compared with 2.6 for GHQ score.

The SS score should be thought of more as a 'lack of contact' score since fewer friends and less social participation result in a higher numerical value of the score. Compared to the four ISSI scores in Henderson, Byrne and Duncan-Jones (1980), this SS score is most similar to the Availability for Social Interaction score, their other scores measuring adequacy and being more markedly skew.

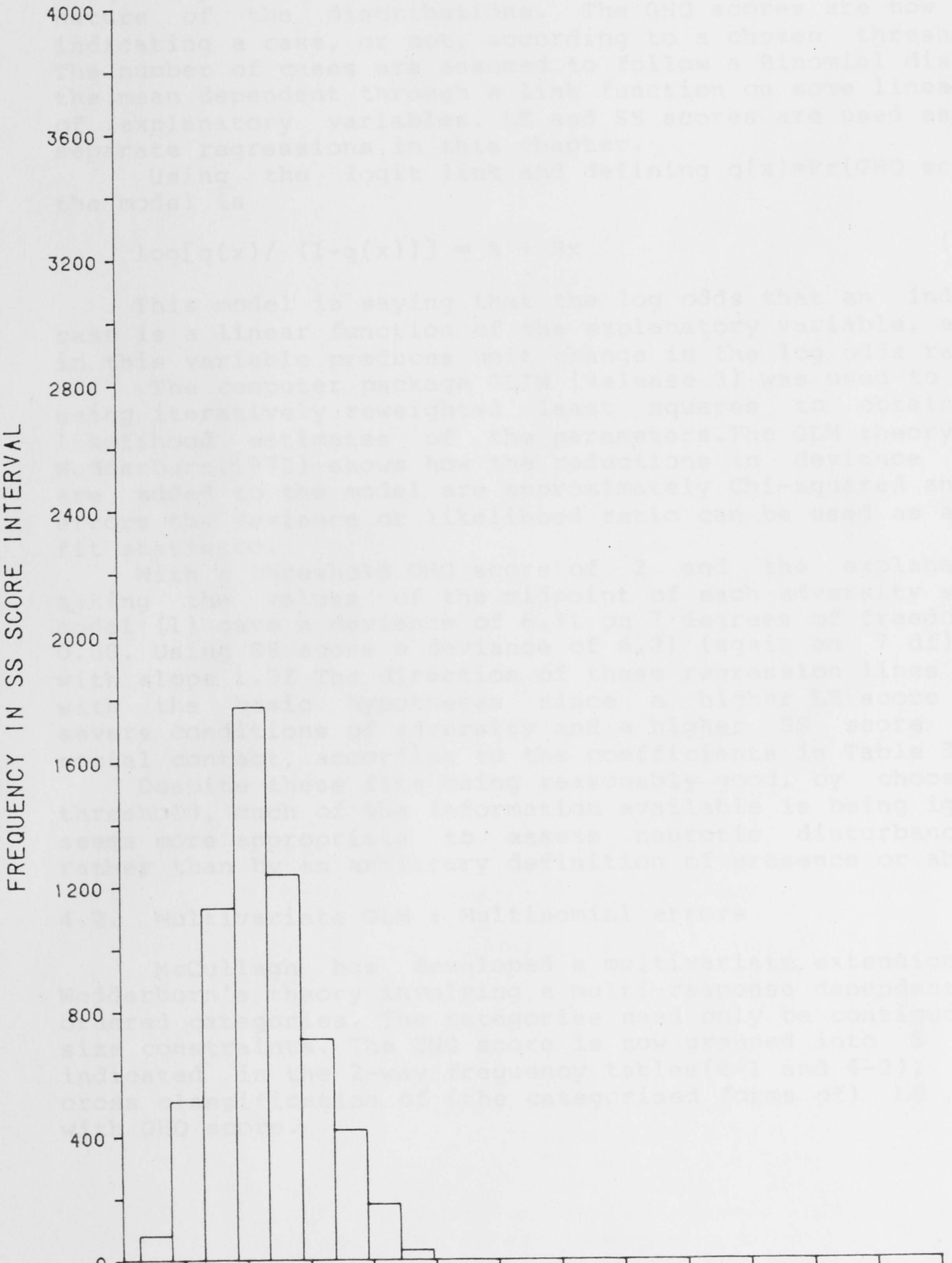
Table 3-2. LE and SS score Frequencies

LE Interval	Midpoint	Frequency	SS Interval	Midpoint	Frequency
0.81,1.16	0.99	3880	0.13,0.39	0.26	79
1.16,1.51	1.34	1406	0.39,0.65	0.52	511
1.51,1.86	1.69	498	0.65,0.91	0.78	1130
1.86,2.21	2.04	174	0.91,1.17	1.04	1772
2.21,2.56	2.39	64	1.17,1.43	1.30	1237
2.56,2.91	2.74	29	1.43,1.69	1.56	710
2.91,3.26	3.09	6	1.69,1.95	1.86	418
3.26,3.61	3.44	8	1.95,2.21	2.08	179
3.61,3.96	3.79	2	2.21,2.47	2.34	31
		6067			6067

PLOT 3.1 HISTOGRAM OF LIFE EVENT SCORES



PLOT 3.2 HISTOGRAM OF SOCIAL SUPPORT SCORES



4. Single Predictor Models

4.1. GLM : Binomial errors

We have seen that the normal linear regression model could not adequately describe the data and the problem lies with the skewed nature of the distributions. The GHQ scores are now interpreted as indicating a case, or not, according to a chosen threshold score(Θ). The number of cases are assumed to follow a Binomial distribution with the mean dependent through a link function on some linear combination of explanatory variables. LE and SS scores are used as predictors in separate regressions in this chapter.

Using the logit link and defining $q(x)=Pr(\text{GHQ score} > \Theta \mid X=x)$ the model is

$$\log[q(x)/\ (1-q(x))] = A + Bx \qquad (1)$$

This model is saying that the log odds that an individual is a case is a linear function of the explanatory variable, and unit change in this variable produces unit change in the log odds ratio.

The computer package GLIM (Release 3) was used to fit the model, using iteratively reweighted least squares to obtain the maximum likelihood estimates of the parameters.The GLM theory of Nelder and Wedderburn(1972) shows how the reductions in deviance as parameters are added to the model are approximately Chi-squared and for binomial errors the deviance or likelihood ratio can be used as a goodness of fit statistic.

With a threshold GHQ score of 2 and the explanatory variable taking the values of the midpoint of each adversity score interval, model (1) gave a deviance of 6.31 on 7 degrees of freedom with slope 0.88. Using SS score a deviance of 6.21 (again on 7 df) was obtained, with slope 1.01 The direction of these regression lines is consistent with the basic hypotheses since a higher LE score indicates more severe conditions of adversity and a higher SS score reflects less social contact, according to the coefficients in Table 3-1.

Despite these fits being reasonably good, by choosing a single threshold, much of the information available is being ignored. It also seems more appropriate to assess neurotic disturbance by degree, rather than by an arbitrary definition of presence or absence.

4.2. Multivariate GLM : Multinomial errors

McCullagh has developed a multivariate extension to Nelder and Wedderburn's theory involving a multi-response dependent variable with ordered categories. The categories need only be contiguous and have no size constraints. The GHQ score is now grouped into 5 categories as indicated in the 2-way frequency tables(4-1 and 4-2). These show the cross classification of (the categorised forms of) LE and SS score with GHQ score.

Table 4-1. LE score by GHQ score frequencies

LE interval mid-point	GHQ score 0	1	2	3..4	5..12	Total
.99	2648	475	243	267	247	3880
1.34	789	231	130	126	130	1406
1.69	237	88	53	56	64	498
2.04	81	29	20	18	26	174
2.39	25	10	9	14	6	64
2.74	10	5	3	5	6	29
3.09	0	1	0	4	1	6
3.44	1	2	1	2	2	8
3.79	0	0	1	0	1	2
Total	3791	841	460	492	483	6067

Table 4-2. SS score by GHQ score frequencies

SS interval mid-point	GHQ score 0	1	2	3..4	5..12	Total
.26	62	9	4	3	1	79
.52	365	72	30	26	18	511
.78	787	136	76	80	51	1130
1.04	1151	240	141	128	112	1772
1.30	721	190	95	121	110	1237
1.56	401	109	56	63	81	710
1.82	212	58	39	41	68	418
2.08	74	24	17	27	37	179
2.34	18	3	2	3	5	31
Total	3791	841	460	492	483	6067

Define $pj(x)=Pr(\text{GHQ score is in any category } 1..j \mid X=x) \quad 1<=j<5$

$=Pr(\text{GHQ score} <= \theta j \mid X=x)$

Then the Proportional Odds model is that

$$\log[pj(x)/(1-pj(x))] = Aj - Bx \qquad 1<=j<5 \qquad (2)$$

Model (2) is just a multivariate form of (1) with multinomial variation instead of binomial errors. McCullagh has written a computer package (called PLUM) which estimates the parameters by maximum liklehood in a manner analogous to GLIM. With the linear structure of model (2), 4 parallel logit regression lines are fitted with intercepts dependent on the cutoffs.

It is noted that these models can be fitted within the current framework of GLIM by using the $OWN directive to define a composite link function, and that this is one of the new features facilitated in the next version of GLIM. (Thompson, Baker 1981)

Significant reductions in deviance were observed in fitting model (2) with LE and SS score category midpoints as the explanatory variable. The deviances were 37.0 and 41.7 both on 31 df, compared with 207 and 208 on 32 df respectively for a model with B=0.

The relationship between models (1) and (2) is summarised in a table of parameter estimates (4-3). This shows the results of fitting model (1) four times with the threshold GHQ score at the four different cut-off points. Note that the intercepts decrease as the threshold increases. The sign difference between the estimates of A and Aj are due to the definitions of q(x) and pj(x).

Note the (opposite) trends in the estimated slopes for model (1) as threshold increases.(See also plots 4.1 (a) and (b)). The McCullagh model is assuming these are constant and the estimates of B for model (2) are closer to the estimates for model (1) at the lower thresholds since 76% of the observations fall into the first two GHQ score categories. The approximate 95% confidence intervals estimated for the slopes do not cover the range given by model (1), so the proportional odds model may not be the best summary of the data.

Table 4-3. Comparing Models (1) and (2)

		Model (1)					Model (2)			
Threshold		0	1	2	4	cutpoint	0	1	2	4
LE	A	-1.78	-2.34	-2.73	-3.42	Aj	1.68	2.26	2.85	3.66
	B	1.07	.964	.882	.792	B		.974		
DEVIANCE		14.6	7.77	6.11	7.32	DEV		37.0		
DF		7	7	7	7	DF		31		
SS	A	-1.38	-2.20	-2.87	-4.01	Aj	1.46	2.14	2.63	3.44
	B	.748	.874	1.01	1.27	B		.830		
DEVIANCE		8.12	6.34	6.24	2.79	DEV		41.7		
DF		7	7	7	7	DF		31		

A model called the proportional hazards model is obtained by replacing the log odds in model (2) by the complementary log log transformation.

$$\log[-\log(1-pj(x))] = Aj - Bx \qquad (3)$$

Whereas previously pj(x) was modelled as the logistic distribution function, model(3) is equivalent to

$$pj(x) = 1 - \exp(-\exp(Aj - Bx))$$

Now the ratio, log(qj(x1)/ log(qj(x2), instead of the odds ratio, is assumed constant over categories, and depends only on the difference between the covariate values.

When the SS score midpoint was used as the single explanatory variable, model (3) gave a deviance of 17.6 (cf 41.7 for the logit link). The deviance increased substantially (to 76.2) for LE score.

However by noting that the C-log log transformation is asymmetric and reversing the order of GHQ score categories in table 4-1, (effectively replacing p by q), a deviance of 35.3 was obtained. The differences in fit are explained by comparing the graphs of slope estimates for the bivariate response models with logit and cloglog links (at different thresholds). The improved fits of the proportional hazards model are reflected by the estimated slopes being more nearly constant over the different thresholds. See Plot 4.1 (a to f). The deviance for the corresponding McCullagh model is shown in brackets on each graph.

4.3. Continuation Odds Models

Another technique for modelling ordinal multivariate responses is to fit model (1) at the series of different thresholds but truncating the sample each time at the previous lowest (or highest) threshold. This is equivalent to conditioning on GHQ score being greater than (lower truncation), or less than (upper truncation) a particular value.

Models (4) and (5) describe two sets of continuation ratios.

$$\log [(P(\text{GHQ score} > j)/(P(\text{GHQ score} = j) = A + Bx \quad j=0,1,2,3,4 \qquad (4)$$

$$\log [(P(\text{GHQ score} < j)/(P(\text{GHQ score} = j) = A + Bx \quad j=4,3,2,1 \qquad (5)$$

Fienberg (1977) discusses these models and shows that the individual Chi-square statistics for each model can be added to obtain an overall goodness of fit statistic for the series.

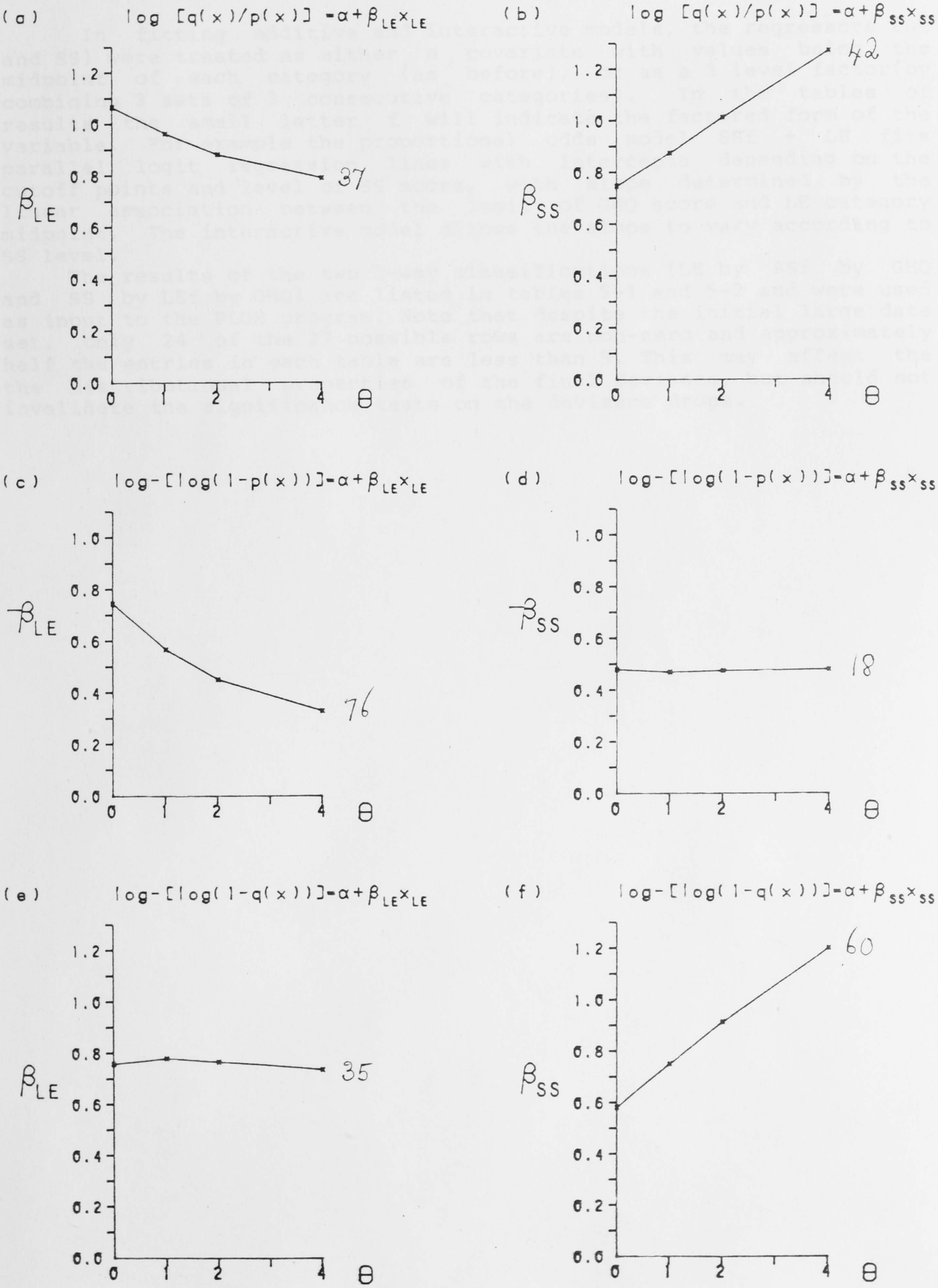
Such models have aims similar to the McCullagh models and the proportional hazards models, and in this study give more insight into the differences between the associations of LE and SS with GHQ score.

Model (4) fitted quite well at the lower thresholds, but once people with GHQ scores of 0 and 1 were excluded from the data, the linear relationship between log odds (GHQ score > threshold) and LE score was lost. It was maintained in the model (5) top truncated series.

Both models (4) and (5) proved suitable with SS substituted for x. The increasing trend in slope with increasing threshold, identified for Model (1), was again observed for the model (5) series, though the magnitude of the slope at each threshold was less for the reduced data sets as people with high scores were excluded.

PLOT 4.1

Slopes estimated for (logit and cloglog) models with 4 different threshold definitions of q(x)



5. Testing for Interaction

5.1. Logit link

In fitting additive and interactive models, the regressors (LE and SS) were treated as either a covariate with values being the midpoint of each category (as before), or as a 3 level factor(by combining 3 sets of 3 consecutive categories). In the tables of results the small letter f will indicate the factored form of the variable. For example the proportional odds model SSf + LE fits parallel logit regression lines with intercepts depending on the cutoff points and level of SS score, with slope determined by the linear association between the logit of GHQ score and LE category midpoint. The interactive model allows the slope to vary according to SS level.

The results of the two 3-way classifications (LE by SSf by GHQ and SS by LEf by GHQ) are listed in tables 5-1 and 5-2 and were used as input to the PLUM program. Note that despite the initial large data set, only 24 of the 27 possible rows are non-zero and approximately half the entries in each table are less than 5. This may affect the the distributional properties of the final deviance, but should not invalidate the significance tests on the deviance drops.

Table 5-1. LE score by SS factor by GHQ score Frequencies

		GHQ score					
LE	SS	0	1	2	3..4	5..12	Total
.99	1	828	127	55	59	25	1094
.99	2	1577	286	145	168	160	2336
.99	3	243	62	43	40	62	450
1.34	1	265	55	31	27	26	404
1.34	2	484	158	89	78	80	889
1.34	3	40	18	10	21	24	113
1.69	1	85	25	19	9	11	149
1.69	2	140	59	31	44	40	314
1.69	3	12	4	3	3	13	35
2.04	1	26	8	4	7	4	49
2.04	2	48	21	15	8	15	107
2.04	3	7	0	1	3	7	18
2.39	1	7	0	1	3	1	12
2.39	2	16	9	8	9	4	46
2.39	3	2	1	0	2	1	6
2.74	1	2	1	0	1	2	6
2.74	2	8	4	2	3	2	19
2.74	3	0	0	1	1	2	4
3.09	1	0	0	0	3	1	4
3.09	2	0	1	0	1	0	2
3.44	1	1	1	0	0	0	2
3.44	2	0	1	1	1	1	4
3.44	3	0	0	0	1	1	2
3.79	2	0	0	1	0	1	2

Table 5-2. SS score by LE factor by GHQ score Frequencies

SS	LE	GHQ score 0	1	2	3..4	5..12	Total
.26	1	59	8	4	3	1	75
.26	2	3	1	0	0	0	4
.52	1	355	69	29	21	18	492
.52	2	10	2	1	4	0	17
.52	3	0	1	0	1	0	2
.78	1	764	130	72	71	43	1080
.78	2	22	6	4	7	7	46
.78	3	1	0	0	2	1	4
1.04	1	1115	221	125	118	106	1685
1.04	2	36	17	15	9	5	82
1.04	3	0	2	1	1	1	5
1.30	1	698	177	88	114	101	1178
1.30	2	23	13	6	6	8	56
1.30	3	0	0	1	1	1	3
1.56	1	388	105	52	58	73	676
1.56	2	13	4	4	5	8	34
1.82	1	205	57	38	37	62	399
1.82	2	7	1	1	4	5	18
1.82	3	0	0	0	0	1	1
2.08	1	73	24	16	24	32	169
2.08	2	1	0	1	2	5	9
2.08	3	0	0	0	1	0	1
2.34	1	17	3	2	3	5	30
2.34	2	1	0	0	0	0	1

The results of fitting the successive proportional odds models (Table 5-3), do not support the hypotheses that a different slope is required for SS score according to the level of adversity , and vice versa.

Table 5-3.

MODEL	DEVIANCE	MODEL	DEVIANCE	DF
null	450.7	null	327.0	92
LE	270.1	SS	162.0	91
LE + SSf	121.3	SS + LEf	90.62	89
LE + SSf + LE.SSf	118.9	SS + LEf + SS.LEf	90.19	87

With the trends identified in Chapter 4 in mind, the single threshold models were again considered by combining the appropriate columns of tables 5-1 and 5-2 and fitting the additive and interative forms of model (1).

The additive model (SS+LEf) was sufficient at all 4 thresholds as was LE+SSf at the 2 lower thresholds. However there was some justification for adding the term LE.SSf at the higher thresholds. See Table 5-4.

Table 5-4.

MODEL	DEVIANCE $\theta=2$	DEVIANCE $\theta=4$	DF
null	238.4	174.8	23
LE	148.5	131.6	22
LE + SSf	31.72	24.95	20
LE + SSf +LE.SSf	22.24	18.99	18

For $\theta=2$ the last difference is significant at the 1% level and for $\theta=4$ it is just under the 5% significance point. Yet as mentioned in Chapter 2 design effects may affect these levels.

Table 5-5 lists the parameter estimates for the 'best' model in each case.

Table 5-5. Parameter Estimates & Deviances

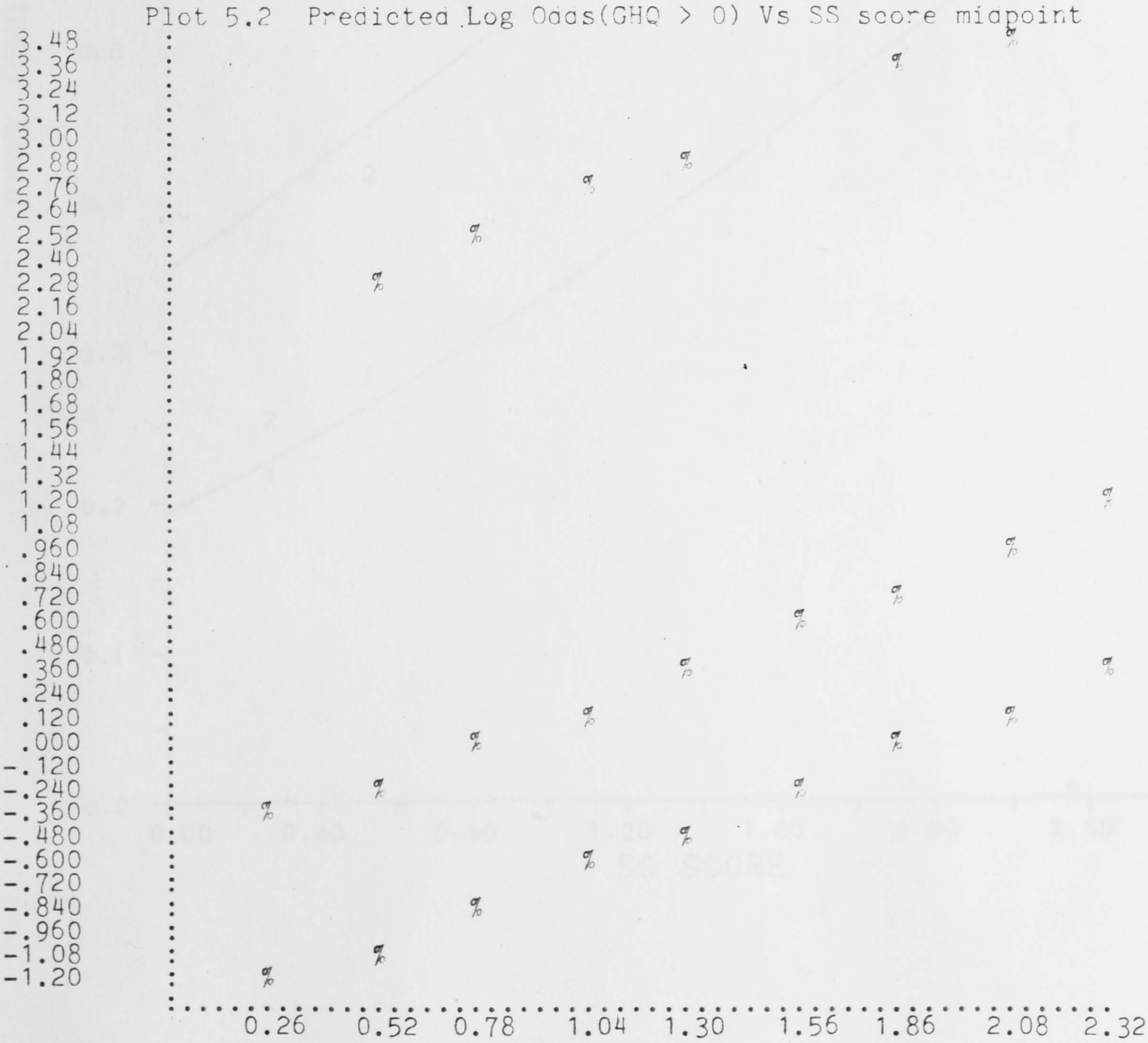
	GLIM					PLUM				
Threshold	0	1	2	4	cutpoint	0	1	2	4
A	-2.21	-2.83	-3.67	-4.77	Aj	2.10	2.80	3.30	4.12
SS(2)	.427	.470	1.18	1.58	SS(2)		.433		
SS(3)	1.01	1.18	1.07	1.91	SS(3)		1.12		
LE	1.11	1.01	.SS1 1.23	1.27	LE		1.01		
			.SS2 .721	.674					
			.SS3 1.45	1.11					
DEVIANCE	26.6	27.2	22.2	19.0			121.		
DF	20	20	18	18			89		
A	-1.43	-2.26	-2.93	-4.03	Aj	1.52	2.20	2.70	3.51
LE(2)	.820	.823	.766	.695	LE(2)		.806		
LE(3)	3.37	2.45	2.37	1.54	LE(3)		2.09		
SS	.754	.882	1.02	1.28	SS		.838		
DEVIANCE	14.9	17.8	20.6	17.8			90.6		
DF	20	20	20	20			89		

The following graphs help to illustrate the essential features of these results. Plot 5.1 shows the relationship between the log(odds for GHQ score > 0) and SS categories for different LE levels (denoted 1,2,3). Plot 5.2 indicates the form of the fitted lines for the additive model. Plot 5.3 shows the fit in terms of the actual probabilities rather than the log odds. There were only 31 individuals in the last SS category and only 1 with GHQ score 0. Even when this individual was removed, the fitted equations did not change significantly.

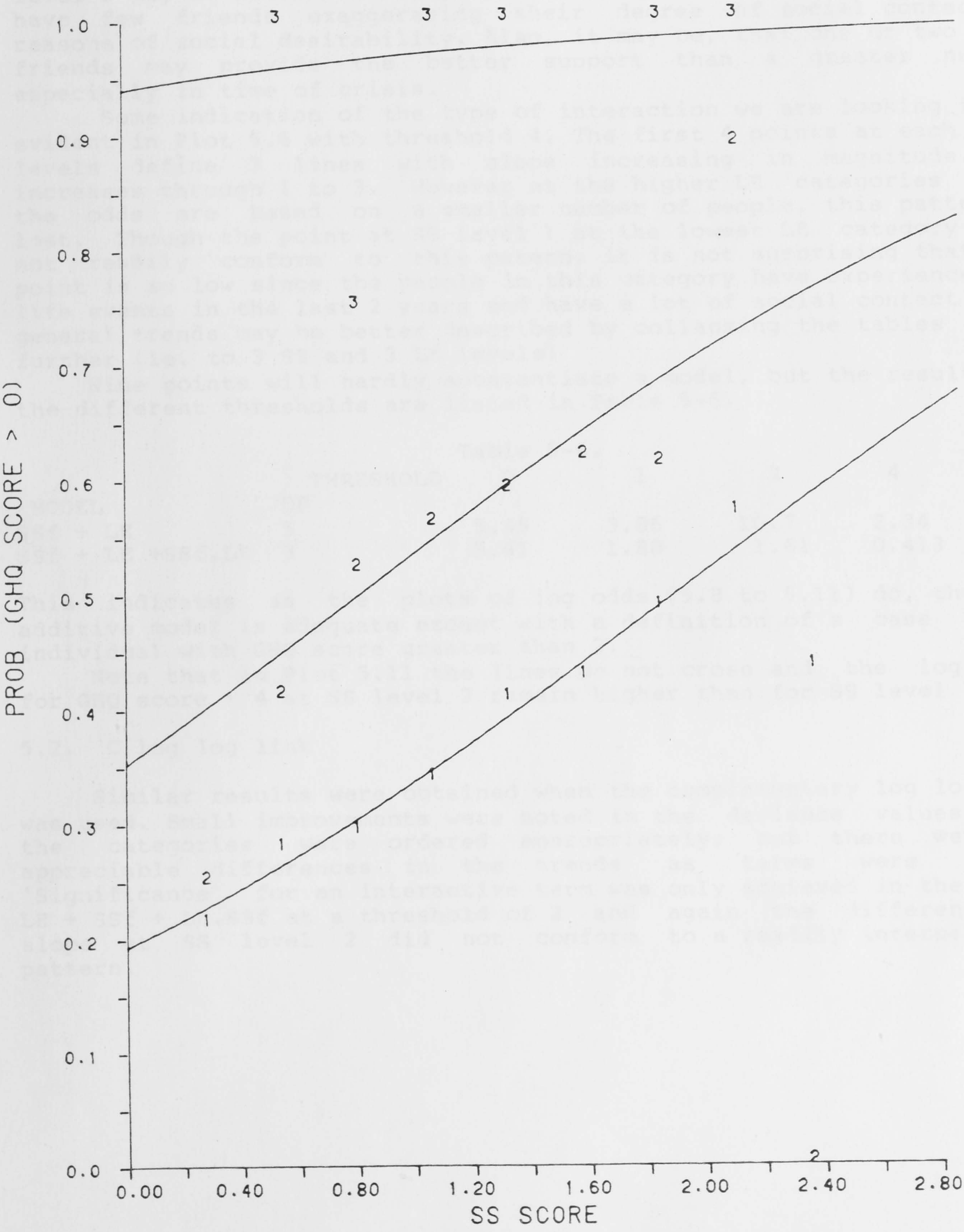
The intercept estimates in the table show that the difference between the fitted lines for LE levels 2 and 3, though large for threshold 0, decreases as the threshold increases.

Plot 5.1 Obs Log Odds (GHQ > 0) Vs SS score midpoint (1,2,3=LE$_f$ levels)

Plot 5.2 Predicted Log Odds(GHQ > 0) Vs SS score midpoint



PLOT 5.3 (1, 2, 3 = LE_f levels)



Plots (5.4,5.5) and (5.6,5.7) show the observed log odds and the fitted interactive models for $\theta=2$ and $\theta=4$. (Now LE is the horizontal axis and 1,2,3 refer to the SS levels). The last 3 LE categories contain only 6,8 and 2 individuals, but removal of the marked point (in 5.4) didn't influence the fit significantly. Their pattern changes considerably with threshold.

The crossing of the lines was not expected. The steep slope at SS level 1 may be attributed to people who scored high in the GHQ and have few friends exaggerating their degree of social contact for reasons of social desirability. Also, it may be, that one or two close friends may provide the better support than a greater number, especially in time of crisis.

Some indication of the type of interaction we are looking for is evident in Plot 5.6 with threshold 4. The first 4 points at each 3 SS levels define 3 lines with slope increasing in magnitude as SS increases through 1 to 3. However at the higher LE categories where the odds are based on a smaller number of people, this pattern is lost. Though the point at SS level 1 at the lowest LE category does not readily conform to this patern, it is not surprising that this point is so low since the people in this category have experienced no life events in the last 2 years and have a lot of social contact. The general trends may be better described by collapsing the tables still further (ie. to 3 SS and 3 LE levels)

Nine points will hardly substantiate a model, but the results at the different thresholds are listed in Table 5-6.

Table 5-6.

MODEL	DF	THRESHOLD 0	1	2	4
SSf + LE	5	5.45	3.86	10.7	2.34
SSf + LE +SSf.LE	3	5.41	1.80	1.61	0.413

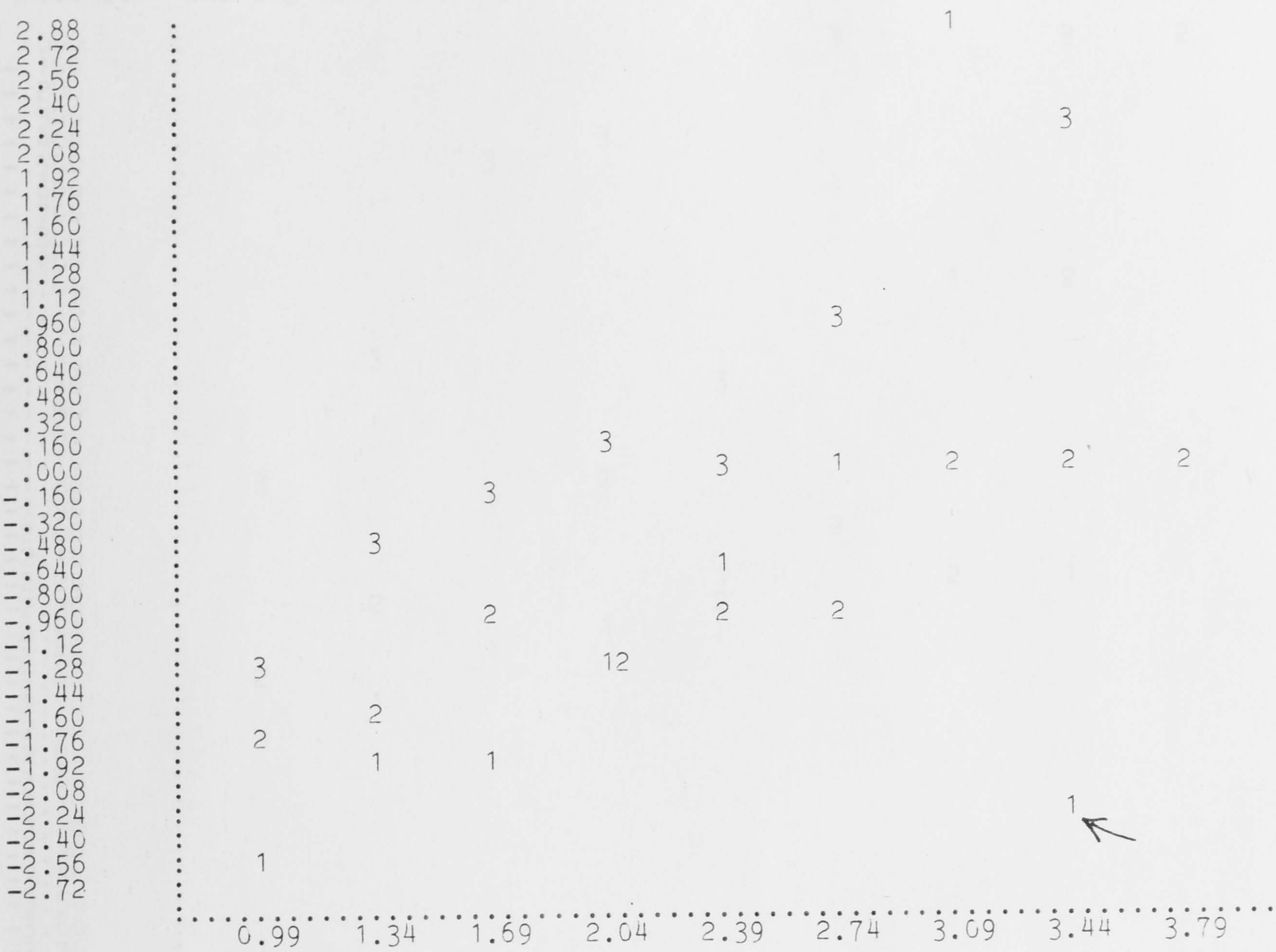
This indicates as the plots of log odds (5.8 to 5.11) do, that the additive model is adequate except with a definition of a case as an individual with GHQ score greater than 2.

Note that in Plot 5.11 the lines do not cross and the log odds for GHQ score > 4 at SS level 2 remain higher than for SS level 1.

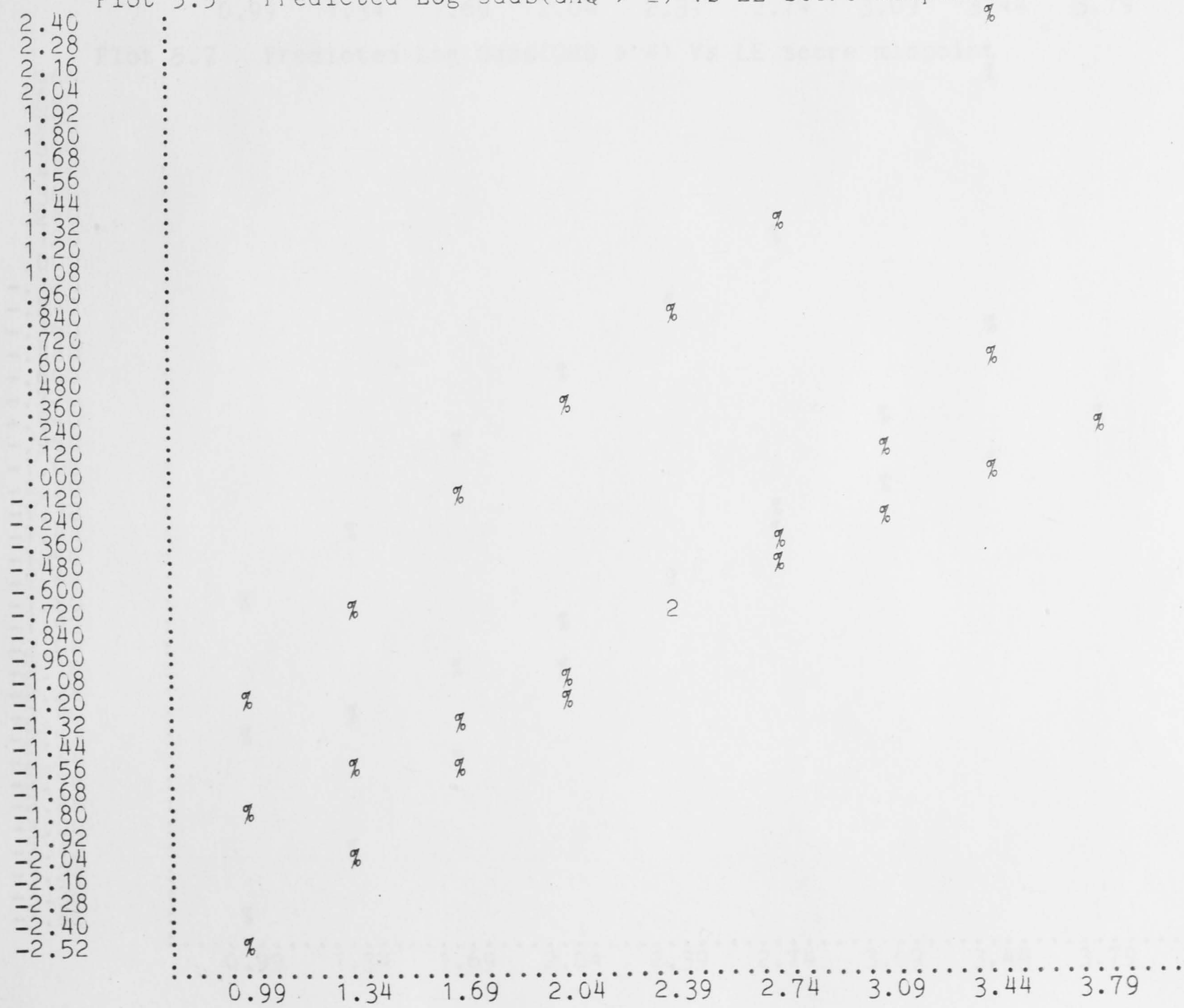
5.2. C log log link

Similar results were obtained when the complementary log log link was used. Small improvements were noted in the deviance values when the categories were ordered appropriately, but there were no appreciable differences in the trends as terms were added. 'Significance' for an interactive term was only achieved in the model LE + SSf + LE.SSf at a threshold of 2 and again the difference in slope at SS level 2 did not conform to a readily interpretable pattern.

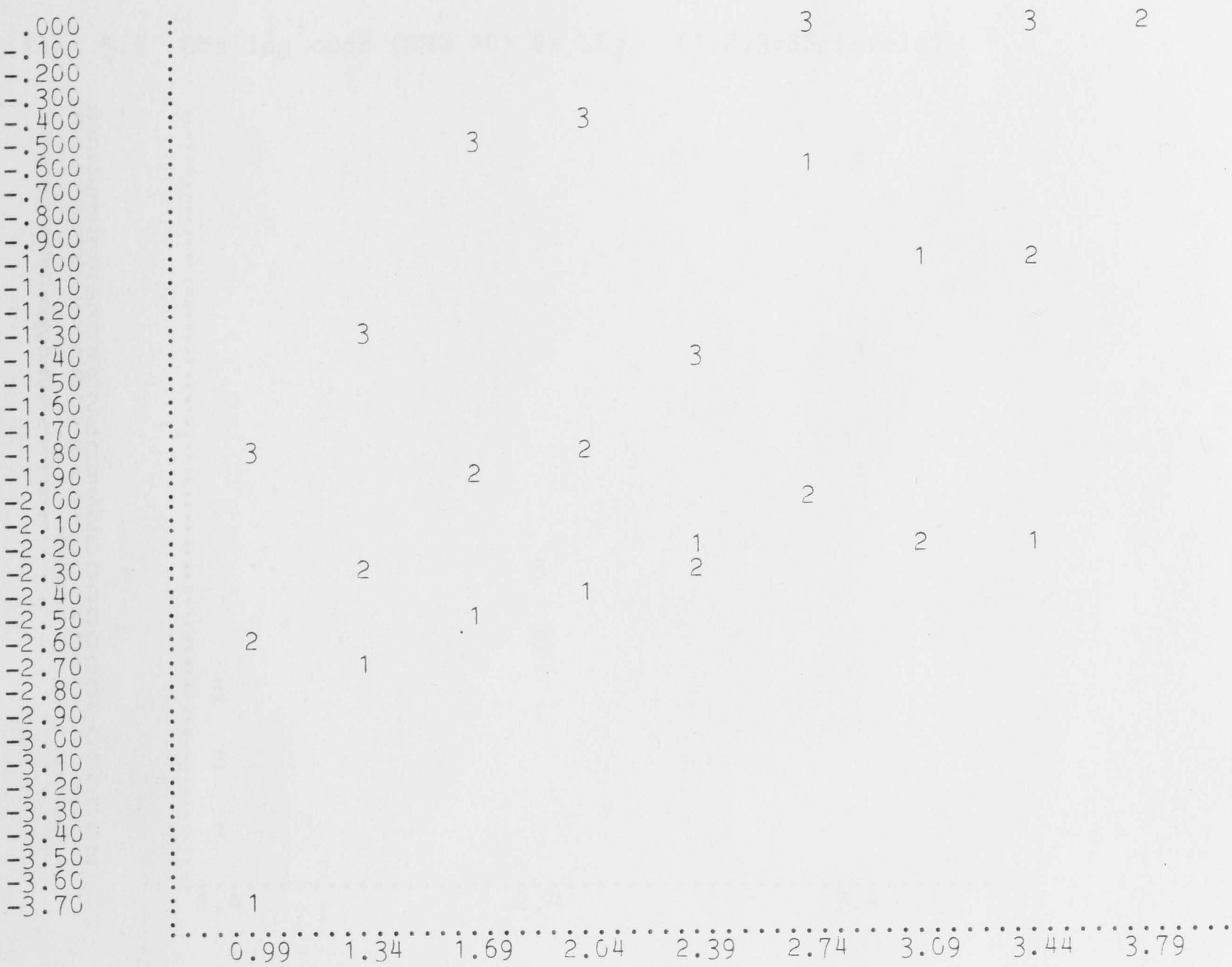
Plot 5.4 Obs Log Odds(GHQ > 2) Vs LE score midpoint (1,2,3=SS_f levels)



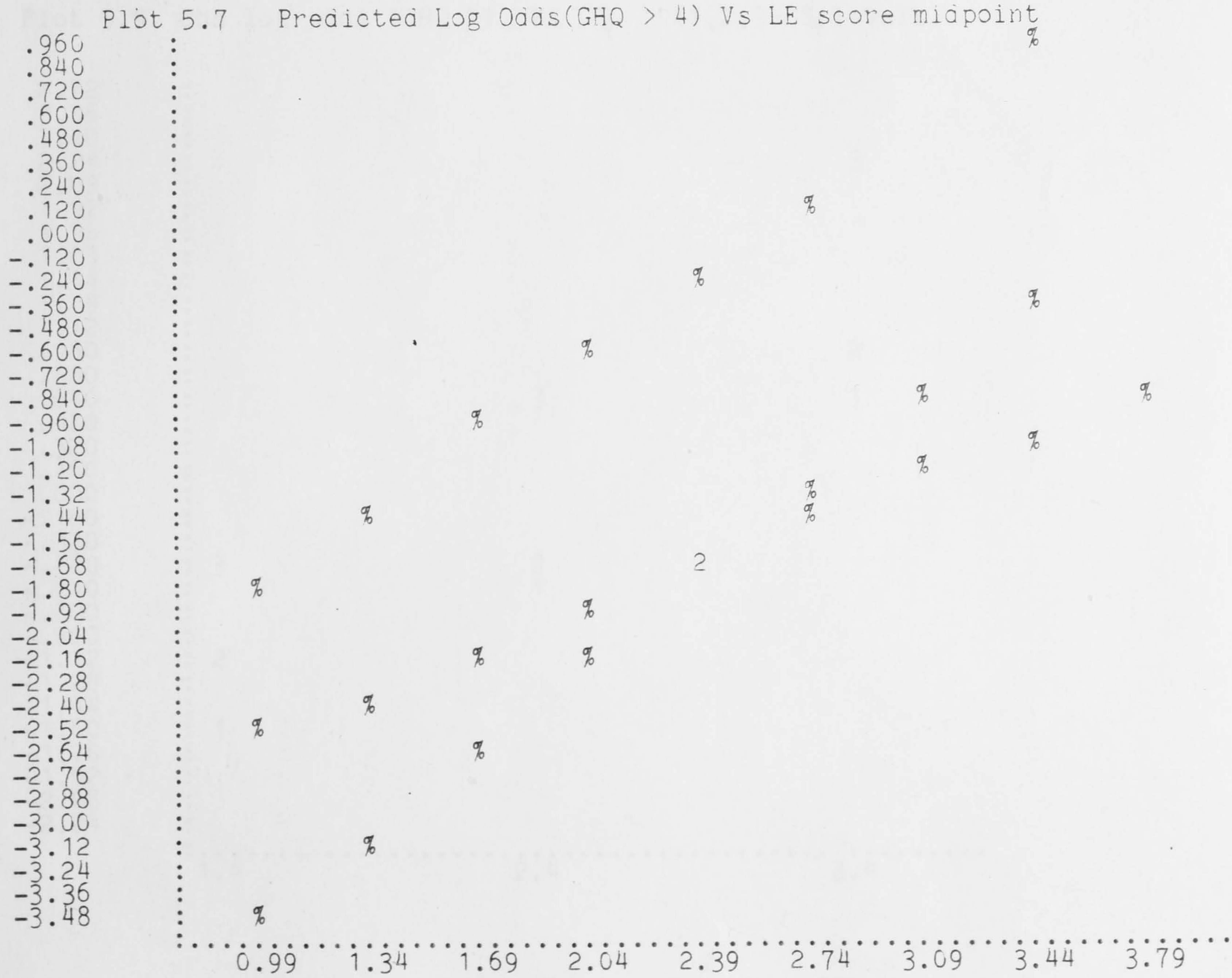
Plot 5.5 Predicted Log Odds(GHQ > 2) Vs LE score midpoint



Plot 5.6 Obs Log Odds(GHQ > 4) Vs LE score midpoint (1,2,3=SS_4 levels)



Plot 5.7 Predicted Log Odds(GHQ > 4) Vs LE score midpoint



Plot 5.8 Obs log odds (GHQ >0) Vs LE_f (1,2,3=SS_f levels)

```
 3.84    :
 3.68    :
 3.52    :                                                2
 3.36    :
 3.20    :
 3.04    :
 2.88    :
 2.72    :
 2.56    :
 2.40    :
 2.24    :                                                3
 2.08    :
 1.92    :
 1.76    :
 1.60    :
 1.44    :                                                1
 1.28    :
 1.12    :
 .960    :
 .800    :                        3
 .640    :
 .480    :
 .320    :                        2
 .160    :
 .000    : 3
-.160    :                        1
-.320    :
-.480    : 2
-.640    :
-.800    :
-.960    : 1
-1.12    :
         :..........................................................
           1.4                    2.4                     3.4
```

Plot 5.9 Obs log odds (GHQ >1) Vs LE_f (1,2,3=SS_f levels)

```
 2.72    :
 2.56    :
 2.40    :
 2.24    :                                                3
 2.08    :
 1.92    :
 1.76    :
 1.60    :
 1.44    :
 1.28    :
 1.12    :
 .960    :                                                2
 .800    :
 .640    :                        3                       1
 .480    :
 .320    :
 .160    :
 .000    :
-.160    :
-.320    :
-.480    : 3                      2
-.640    :                        1
-.800    :
-.960    :
-1.12    : 2
-1.28    :
-1.44    :
-1.60    : 1
-1.76    :
-1.92    :
-2.08    :
-2.24    :
         :..........................................................
           1.4                    2.4                     3.4
```

Plot 5.10 Obs log odds (GHQ >2) Vs LE$_f$ (1,2,3=SS$_f$ levels)

```
  2.40 :
  2.24 :                                  3
  2.08 :
  1.92 :
  1.76 :
  1.60 :
  1.44 :
  1.28 :
  1.12 :
  .960 :
  .800 :
  .640 :                                  1
  .480 :
  .320 :                3
  .160 :
  .000 :                                  2
 -.160 :
 -.320 :
 -.480 :
 -.640 :
 -.800 :
 -.960 :  3             1
 -1.12 :                2
 -1.28 :
 -1.44 :
 -1.60 :  2
 -1.76 :
 -1.92 :
 -2.08 :
 -2.24 :  1
 -2.40 :
 -2.56 :
       :.........................................
         1.4            2.4               3.4
```

Plot 5.11 Obs log odds (GHQ >4) Vs LE$_f$ (1,2,3=SS levels)

```
  .000 :                                  3
 -.120 :
 -.240 :
 -.360 :
 -.480 :
 -.600 :                3
 -.720 :
 -.840 :
 -.960 :
 -1.08 :                                  2
 -1.20 :
 -1.32 :
 -1.44 :                                  1
 -1.56 :  3
 -1.68 :
 -1.80 :
 -1.92 :                2
 -2.04 :
 -2.16 :                1
 -2.28 :
 -2.40 :  2
 -2.52 :
 -2.64 :
 -2.76 :
 -2.88 :
 -3.00 :
 -3.12 :
 -3.24 :  1
 -3.36 :
 -3.48 :
 -3.60 :
 -3.72 :
       :.........................................
         1.4            2.4               3.4
```

6. Models involving the original variables

The use of composite scores though justified for identifying general trends in the context of the sample as a whole, becomes tenuous as the individuals in each cell become more important when the data set is more finely cross-classified. A different method of categorising the LE and SS scores was also considered in which each of 6 categories contained about 1000 individuals and the covariate value used was the median score in each interval. Again the compounding effects hypothesis was not supported but a slope difference parameter for SS level 2 was found to be significant at the higher thresholds when LE was considered as the covariate and SS combined to a 3 level factor.

There are also problems with the time scale. A particularly distressing event or series of events may have occurred at a stage when the degree of social contact or number of confidents was quite different to that described by the SS score. Also, the General Health Questionnaire only measures a person's current mental state and there is no information on immediate reactions to events which happened a long time ago.

The actual parameters estimated for these models may not bear comparison with other sets of data since the construction of the scores was based on the initial regressions.

In order to be more specific, just the information on LE 1 was considered and the effect of experiencing one or more life events as opposed to none in the last two months, tested with some of the SS variables separately, using the techniques discussed above with thresholds and modelling the trend in scores more directly. Using the original SS variables eliminates the need for the somewhat arbitrary definition of low, medium and high levels of social support score and uses the levels 0, 1-2, and 3+ of the raw data.

6.1. More Logit Models

In particular, the number of persons available for practical help, emotional help, the number of friends spoken to in the last week, and the total number of community activities of the past month, were considered separately.

In each case the effects were significant, but additive with experiencing life events, in their influence on GHQ score. With each logit model the estimated coefficient for each of these aspects of Social Support increased in magnitude and the fit improved as the definition of a case became more severe.

When the actual number of life events experienced in the previous two months was taken into account there was again no significant compounding interaction, nor however, was there any evidence for the type of interaction found at the higher thresholds with the composite scores.

6.2. Exponential Regression

Borrowing techniques from survival analysis and substituting GHQ score for time, exponential regression models were fitted to the data cross-classified as displayed in Table 6-1, to see if the rate parameter could be related to the level of the particular Social Support variable and the presence or absence of life events in the last 2 months.

Table 6-1.
SS1 (practical help) by LE1 by GHQ score

SS1	LE1	GHQ score 0	1	2	3	4	5	6	7	8	9	10	11	12
0	0	56	17	13	9	6	6	1	3	2	1	1	1	1
0	1+	16	7	5	2	2	1	2	1	0	1	1	0	1
1-2	0	734	142	84	57	36	22	23	13	12	7	8	7	6
1-2	1+	218	67	49	31	26	21	13	9	7	6	2	6	6
3+	0	2005	374	177	124	76	44	44	25	12	15	11	9	4
3+	1+	762	234	132	77	46	34	17	21	18	13	10	9	6

SS2 (emotional help) by LE1 by GHQ score

SS2	LE1	GHQ score 0	1	2	3	4	5	6	7	8	9	10	11	12
0	0	210	43	27	17	11	7	9	9	3	3	4	2	3
0	1+	62	27	15	9	8	6	3	1	5	3	1	2	1
1-2	0	1067	245	124	88	51	38	30	19	16	12	9	10	6
1-2	1+	366	141	80	48	43	24	15	14	11	9	4	5	9
3+	0	1518	245	123	85	56	27	29	13	7	8	7	5	2
3+	1+	568	140	91	53	23	26	14	16	9	8	8	8	3

SS5 (friends) by LE1 by GHQ score

SS5	LE1	GHQ score 0	1	2	3	4	5	6	7	8	9	10	11	12
0	0	224	54	35	13	16	13	14	5	5	6	4	3	4
0	1+	62	26	14	14	7	8	0	3	6	2	3	1	3
1-2	0	436	95	43	41	25	18	15	11	10	5	4	5	4
1-2	1+	112	41	25	16	14	7	12	8	6	4	2	3	3
3+	0	2135	384	196	136	77	41	39	25	11	12	12	9	3
3+	1+	822	241	147	80	53	41	20	20	13	14	8	11	7

Letting y = GHQ score (+1) and x refer to the covariates, the model is

$$f(y) = \exp (Bx - y \exp (Bx))$$

This model can be fitted in GLIM using poisson errors and log(y) as an 'offset' (see Aitken 1980). The high frequencies for GHQ score 0 contributed to large deviances. Even if this column is removed some lines fit better than others as is seen in the table of deviances for fitting a model to each line (6-2). The worst is consistently the most sheltered group, (those with a lot of personal contact and no life events).

Table 6-2.

Deviances for fitting separate Exponential models to each line excluding GHQ score 0. (DF 11)

SS	LE1	Practical help	Emotional help	Friends
0	0	7.7	6.8	9.7
0	1+	4.6	7.0	16.
1-2	0	26.	54.	15.
1-2	1+	17.	30.	6.9
3+	0	137.	111.	170.
3+	0	57.	37.	67.

6.3. Proportional Hazards

The popular approach of survival analysis is to place less importance on finding the most appropriate parametric model and assume some underlying but unknown hazard function of the form $\lambda_0(t)$ exp (Bx) and base the estimation of B on a partial likelihood involving the ratio of hazards. The equivalent of the hazard function in this case is Pr(Y=y | Y >= y) and so the discrete analogue of this approach is equivalent to the continuation odds series model (4) of Chapter 4.

Again better results where achieved if people with GHQ score 0 were excluded and the deviances in Table 6-2 show the improvement over the exponential for the 3+ categories of Social Support.

Table 6-3.

Deviances for fitting separate prop. hazards models to each line excluding GHQ score 0 (DF 11)

SS	LE1	Practical help	Emotional help	Friends
0	0	5.8	15.	23.
0	1+	7.3	12.	24.
1-2	0	26.	35.	20.
1-2	1+	21.	31.	16.
3+	0	40.	20.	34.
3+	1+	44.	32.	39.

Fitting the covariate structure showed LE1, SS1, SS2 and SS5 to be significant with no evidence for interaction. Inspection of the coefficients revealed no significant difference between the 0 and 1-2 categories of SS1 and SS5 but it was evident for SS2, so having at least one confidant at an emotional level is better than none.

Reinvestigation of the separate logit models of 6.1 excluding people with GHQ score 0, also showed this effect. In addition the 1-2 and 3+ categories of SS2 were not significantly different. This may partly explain the aberrant interaction term found significant for the models with the factored form of the SS composite score in Chapter 5.

7. Discussion

Though evidence was found for the association between neurosis and the effect of life events and the effect of social support as separate factors, there was no conclusive evidence to support the hypothesis that conditions of adversity and lack of social contact together contribute more to neurosis than when taken account of separately.

Though there was some evidence for a different type of interaction, it occurred only in the models with high threshold, based on composite scores and with LE score as the regressor and SS score divided into a 3 level factor. This may be due partly to the smaller number of people with high GHQ scores leading to small expected values (with higher variance). So despite the 'advantage' of having a large sample, if the whole data set is used then the proportion of disturbed people is still small. Deciding on the best trade-off between using as much information as was available and using as much as was relevant in order to be precise, was a recurrent problem throughout the analysis. The results have shown the importance of distinguishing different levels of disturbance and especially excluding the most 'normal' group.

The size of the data set did allow us to investigate the trends that were significant, and compare the relative merits of different functional forms that could be used to describe them. Several methods of modelling the GHQ score as an ordered polychotomous response were considered. Of the contingency table methods, the threshold series approach was most useful.

When LE score was condensed to a 3 level factor, the McCullagh proportional odds model LEf+SS gave a reasonable summary of the data without the need for a strict case / non-case definition. The log odds for having a 'higher' GHQ score increases linearly with increasing SS category and a fixed amount for increasing LE level. However by assessing the other results as well it was found that this may be an over-simplification of the situation.

The results of fitting the LEf + SS proportional odds models at different thresholds (table 5-5) show that the distinction between LE levels decreases with a more severe definition of a case. The series of logit regressions presented in Chapter 4 (Table 4-3) show the different threshold GHQ scores which result in the steeper slopes, in modelling the log odds of being a case.

It is clear that the LE score and SS score of this analysis have different kinds of association at different levels of GHQ score. There was also some evidence (Chapter 6) of different effects of the Social Support variables. So it is not not surprising that the results of testing for interaction were slightly conflicting.

In conclusion it is proposed that the main effect of experiencing a number of life events over a period of 2 years, may be to move people from low to moderate levels of distress or anxiety and have little effect once people reach a certain level of disturbance. Accepting that people are more likely to seek clinical advice if they are severely disturbed, these findings are consistent with the results of Bebbington et al (1981), who showed a greater association between life events and mental disorders in cases in a community sample than with cases in a psychiatric out-patient sample.

On the other hand social support in the form of friends and community participation seems to protect people at all levels from becoming more distressed. The trend in estimated slopes for both composite and raw scores suggests that this effect is more important for more disturbed people, but at these levels neurotic symptoms could contribute more to isolation. This highlights the need for more thorough investigations into the direction of causality to interpret these different associations.

8. Bibliography

Aitken M.,Clayton D. (1980) The fitting of Exponential, Weibull and Extreme Value Distributions to Complex Censored Survival Data using GLIM. Applied Statistics, 29 No. 2, 156-163.

Bebbington P.E.,Tennant C.,Hurry J. (1981) Adversity and the Nature of Psychiatric Disorder in the Community. Journal of Affective Disorders, 3, 345-366.

Brown G.W. (1974) Meaning,Measurement and Stress of Life Events. 'Stressful Life Events: Their Nature and Effects'. ed. Dohrenwend & Dohrenwend. Wiley, 217-245.

Brown G.W., Bhrolchain M.N., Harris T.O. (1975) Social Class and Psychiatric Disturbance among Women in an Urban Population. Sociology, 9, 225-254.

Brown G.W., Harris T. (1978) 'Social Origins of Depression: A Study of Psychiatric Disorder in Women.' London Tavistock.

Fienberg S.E. (1977) 'The Analysis of Cross-Classified Categorical Data' MIT Press.

Goldberg D.P. (1972) 'The Detection of Psychiatric Illness by Questionnaire' Instit. of Psychiatry Maudsley Monographs No 21. London : Oxford University Press.

Heathcote C.R. (1982) Linear Regression by Functional Least Squares. Applied Probability Trust.

Henderson A.S., Byrne D.G., Duncan-Jones P., Adcock S. (1980) Social Relationships, Adversity and Neurosis : A Study of Associations in a General Population Sample. British Journal of Psychiatry, 136, 574-583.

Henderson A.S. (1981) Social Relationships, Adversity and Neurosis : An Analysis of Prospective Observations. British Journal of Psychiatry, 138, 391-398.

Henderson A.S., Byrne D.G., Duncan-Jones P., (1981) 'Neurosis and the Social Environment' Academic Press.

Kish L., Frankel M.R., (1970) Balanced Repeated Replications for Standard Errors. JASA 65, 1071-1094.

Landis R., (1982) Empirical Investigation of the Influence of Design Effects on Test Statistics Using Data from the 1st U.S. National Health and Nutrition Examination Survey.

McCullagh P., (1980) Regression Models for Ordinal Data. JRSS B 42, 109-142.

Miller P. McC., Ingham J.G. (1976) Friends, Confidants and Symptoms. Social Psychiatry. 11, 51-58.

Nelder J.A., Wedderburn R.W.M. (1972) Generalised Linear Models. JRSS A 135, Part 3, 370-384.

Thompson R.,Baker R.J. (1981) Composite Link Functions in Generalised Linear Models. Applied Statistics, 30 No. 2, 125-131.