An investigation into statistical methods for analysing ordered categorical data

John D. Gretton

A thesis submitted in fulfilment of the requirements for the degree of
Master of Philosophy

October 1998

School of Computing and Management Sciences

Sheffield Hallam University

Acknowledgements

I am sincerely grateful to my director of studies, Dr. Bekia Fosam, and my supervision team, Mike Grimsley, Eric Okell, Steve Wisher and Keith Jones, for their invaluable help and guidance throughout the course of the research. I was never found wanting for contact with people with bigger and different shaped brains than mine. Thanks a lot you guys.

I am indebted to my Examiners Dr Richard Gadsden and Professor Ian Dunsmore for their effort in inspecting this thesis (it can't have been enjoyable), and their attitude in allowing me to make amendments for re-examination without any hassle, when time and computing resources have been at a premium for me since starting full-time work.

Fond regards to the SSRC for all kinds of stuff, like contract work to give me some extra pennies, administration (and then some!) and generally just being friendly, helpful people.

Big ta's to Simon Alexander for mending my computer when it went down on me, and lending me the portable PC on which a fair proportion of this thesis was written. Same to Errol and Homed, too.

Much appreciation to Christine Straker for sorting me out loads of lecturing work to help my bank balance, communication skills and confidence, and to boost my CV. Cheers for being my referee, and the chats also.

Much respect is due to my good friends (Al, Paresh, Bob, Ad, Marky B., Jay, Si, Doggy Dave, Justin, Emma, Shelley & Dips) for listening to me moan about how hard I thought my research was, and helping to take my mind off it from time to time, as well as badgering me to get the corrections done for 10 months. Cheers Ears.

Many thankyous to Darren and Jo for a new place to live and the much sought after portable PC with which to correct the original incorrectness in this thesis. Same goes to Jonathan Gorst for his PC and office space for round 2 in the correction procedure.

Finally thanks to George Michael for some top songs, Wrighty (bon voyage, man), Dennis and Arsenal F.C. for giving me something to cheer about and having a great kit, and Cadbury's for their fantastic chocolate.

Summary

This investigation researches statistical methods for analysing ordered categorical data. Some standard descriptive and modelling procedures are described, and the data is analysed using a relatively new statistical package, CHAID, which is designed purely for categorical data analysis. The study is centered around the application of the proportional odds and continuation odds models, to data obtained from a survey of the opinions of South Yorkshire Police staff (SSRC (1994)). Morale within the South Yorkshire Police is the factor of interest, and is discussed in some detail. The two approaches of proportional odds and continuation odds models are discussed critically. Dummy variables and scored levels are employed for the treatment of ordinal variables. The effects of these two methods of coding ordinal data, on the results of the analyses, are also compared and discussed. Methods of assessing the goodness-of-fit of ordinal models are discussed, and a modification to the guidelines for using a recently presented technique (Lipsitz et al (1996)) is suggested and applied. The proportional odds model is successfully applied. The implications from the models produced are that job satisfaction, communication, public view of the police, promotion issues and length of service have an influence on the morale of an individual, in general.

Contents

Sections

Chapter 1: Introduction

1.1: The research problem - General

There is a distinct shortage of statistical methodologies that deal specifically with ordered categorical data. Methods that have been developed are not widely used to analyse ordinal data, more often techniques for analysing nominal or interval data are applied. Therefore, there is a need for greater understanding of how to treat ordinal data, and possibly greater accessibility of ordinal methods. There is uncertainty about the interpretation of some ordinal models, and ways to assess their goodness-of-fit.

This research is centered around the analysis of data with an ordinal response variable, and addresses the problems involved in analysing ordered categorical data. Ordinal data occurs when a categorical variable has an intrinsic ordering to its levels, so an underlying continuum is assumed. This type of data is very common in market research and medical studies, among other areas, thus the need for definitive methodologies is important.

1.2: Categorical and ordered categorical data

Categorical data arise frequently in many areas of research. A categorical variable is one where the measurement scale is a set of categories, e.g. political belief may be gauged as 'liberal', 'moderate' or 'conservative', or pain after an operation might have response categories of 'none', 'mild' or 'severe'.

A categorical variable whose levels have no natural or distinct ordering is called nominal. Examples of nominal information are religious affiliation (Catholic, Jewish, Protestant, other), mode of transport (car, bus, bicycle, foot, other), race, gender and marital status. For this type of variable, the ordering of the categories is irrelevant to any statistical analyses

Categorical variables which do have ordered levels are called ordinal. Examples of these could be social class (upper, middle, lower), attitude towards legalisation of abortion (strongly disapprove, disapprove, neither, approve, strongly approve) or diagnosis of multiple sclerosis (certain, probable, unlikely, definitely not). The categories of ordinal variables are clearly ordered, and in a lot of cases one could assume some underlying continuous scale. Whilst absolute distances between levels are unknown, one can conclude that someone categorised as 'mild' is in less pain than a person categorised as 'severe', although a quantitative measure of how much less pain the individual is in is realistically unobtainable. An interval variable is one which does have quantifiable distances between levels, e.g. income or age.

An ordinal data variable is one where there are distinct categories with a definite ordering. For example in medical research one might come across a pain response of none, mild or severe, or in market research response to a statement may be gauged by a likert scale variable with categories strongly agree, agree, neither agree nor disagree, disagree or strongly disagree. Both these examples assume an underlying continuous scale. The absolute distances between categories are not easily determinable, in that although no pain is better than mild or severe pain, and similarly mild pain is more favourable than severe pain, we cannot quantify precisely how much better. Similarly, whilst agreement or strong agreement with a statement may be desired, in the context of some research, one could not quantify how much better those responses are than strong disagreement, disagreement or neutrality. If this information were ascertainable, we would be able to turn the information into continuous or interval scale variables.

Despite the frequent and growing use of ordinal data, methods for analysing it are still a little sparse and uncommon. Most techniques used treat ordinal variables as nominal because they are categorical. Whilst the recognition of the categorical nature of the data is useful, the distinction between qualitative (nominal) and quantitative (ordinal) data is possibly more important, and ordinal variables should be treated more like interval variables in terms of descriptive measures and maybe modelling. Too frequently, ordinal data are split into binary variables representing, say, success and failure. There may be a genuine interest in the defined success and failure division, however, often there is a lack of thorough understanding of existing techniques to analyse ordinal independent variables, or lack of accessibility which leads to the reduction of the response to binary, and often less useful analysis.

The way a characteristic is measured determines the form of data generated and hence determines plausible methods of analysis. For instance, a variable 'education' can be nominal if measured by types of education such as public school or private school, or ordinal when measured in terms of infant, junior, secondary, fifth form, sixth form, university and postgraduate, and interval when measured by number of years in education 0, 1, 2, ...etc.

Nominal variables are qualitative - distinct levels differ in quality not in quantity. Interval variables are quantitative - distinct levels have differing amounts of the characteristic in question. The position of ordinal variables in terms of quantitative/qualitative classification is often ambiguous. Frequently ordinal data are analysed as qualitative, because they are categorical like nominal variables, but in many respects ordinal variables are more like interval variables, as they possess important quantitative features, in that each level has a smaller or greater magnitude of the characteristic than another level.

1.3: Rationale for the proportional odds and continuation odds models

Much evolution has taken place for methods of analysing a continuous response or a binary response, however, techniques for analysing an ordinal response are in their infancy, relatively. Ordinal regression models in general, are not widely used, and scarcely covered in any undergraduate statistical study, whereas literature for, say, multiple regression, logistic regression and analysis of variance is widely available.

The binary logistic regression model analyses a dichotomous response, representing the 'success' and 'failure' of a defined event. If ordinal analyses are inappropriate or unfeasible, it is common to split an ordinal response into two groups of interest and analyse the dichotomised variable using logistic regression (Carroll (1993)). If an ordered categorical response variable, response to a statement say, has classes agree strongly, agree, neither agree nor disagree, disagree and disagree strongly, then a dichotomy of interest may be to combine those who agree (agree or agree strongly) versus those who do not agree (neither agree nor disagree, disagree or disagree strongly). The binary logistic model compares the log odds of an individual agreeing with the statement against those for an individual not agreeing, given specific covariate characteristics. The binary logistic model accommodates ordinal information in this context, but does not utilise the ordinality of a variable. This model can be fitted simply using many standard packages such as GLIM, SAS and SPSS. The goodness-of-fit of the model can be tested by a measure of deviance using GLIM (Lindsey (1989)), as well as goodness-of-fit tests proposed by Hosmer and Lemeshow (1989).

The proportional odds and continuation odds models are specifically designed for an ordinal response variable. The proportional odds and continuation odds models permit single sweeping statements about the effect of independent variables on an ordinal response. The methodology shows that the proportional and continuation odds models are effectively a method of combining or simultaneously fitting several logistic regression models, so the concept perhaps is not revolutionary. The models operate by using a single log odds ratio, that represents several log odds ratios pertaining to binary splits of the response. If using a single 'global' log odds ratio is statistically feasible, then the implications of a proportional odds or continuation odds model may enable a single decision or interpretation, rather than many. For example, taking the proportional odds model, if an ordinal response is 'pain after an operation', diagnosed as none, moderate and severe, the most desirable response (especially for the patient) is none, and the next most desirable response is moderate. If we can determine the odds of a patient experiencing no pain versus moderate or severe, and the odds of none or moderate pain versus severe, with respect to influencing factors, and conclude that these odds are equivalent, then it is the odds of a more desirable level of pain that are examined. Subsequently we may be able to use a single model from which to draw implications about influencing factors rather than two models. Similarly for the continuation odds model, the odds of no pain versus moderate or severe, and the odds of moderate pain versus severe are simultaneously estimated for explanatory characteristics. The proportional odds and continuation odds models are also more parsimonious than a model without the assumption of global odds, logically, as the models produce a single parameter per covariate, rather than parameters pertaining to the possible adjacent dichotomies. The proportional odds model may be fitted using SAS very simply (Carroll (1993)), and instruction on fitting the model using GLIM is given by Hutchison (1985). The goodness-of-fit of the proportional odds model can be determined by statistics proposed recently by Lipsitz et al (1996). The continuation odds model may be fitted using SAS procedure LOGISTIC (Carroll (1993), Berridge and Whitehead (1991)), involving some manipulation of data, or using SAS procedure PHREG to fit the model as a proportional hazards model (Iyer (1985)). Iyer (1985) also gives direction on how to fit the model using GLIM. The goodness-of-fit of the continuation odds model may also be tested by statistics outlined by Lipsitz et al (1996).

Ordinal logistic regression is equivalent to simultaneously fitting binary logistic models to all possible adjacent dichotomies of the response variable, adhering to the set ordering of the response categories, and therefore only dichotomises an ordinal dependent variable between adjacent levels. This models may be fitted simply using most standard statistical packages, e.g. SAS Proc CATMOD, and the goodness-of-fit assessed by maximum likelihood deviance analysis produced within the SAS procedure.

The stereotype model is also designed for an ordinal response, though more suitable for a measure that is perhaps a sum of qualitative indicators (Greenland (1994)). The form of the model follows the ordinary polytomous regression model, using scores for the levels of the response variable. The stereotype model may be fitted via constrained polytomous regression using a standard statistical package such as SAS (Proc CATMOD). The goodness-of-fit of the model may be tested using maximum likelihood deviance statistics.

Using standard parametric methods, i.e. regression on scores for the levels of an ordinal response variable, depends to a large extent on the distribution of the data.

Multiple regression requires that explanatory covariates are treated as known or fixed, with the response (and therefore error terms of a model, also) being normally distributed. For ordinal or categorical response variables, this is not likely to be the case. The approaches and principles that guide, say, linear regression analysis can be used to guide categorical and ordinal data modelling, but the distributional considerations are vital to the success and robustness of any technique, and therefore parametric methods for analysing ordered categorical data are not explored in this research.

If the assumptions of proportional odds and continuations are satisfied, the resultant models are simple to interpret and relatively parsimonious, which is the motivation for fitting a model of this type over a different, often less efficient way of analysing an ordinal response. Therefore, the use of these more sophisticated models is exploited and evaluated in more detail than other methods discussed.

Advantages to using Ordinal methods over standard nominal include the following (Agresti (1984)) :-

Ordinal methods have greater power for detecting important alternatives to null hypotheses such as independence.

Ordinal data description is based on measures that are similar to those used in ordinary regression and analysis of variance for continuous variables, i.e. correlations, slopes, means.

Ordinal analyses can use a greater variety of models, most of which are more parsimonious and have simpler interpretations than the standard models for nominal variables.

Interesting ordinal models can be applied in settings where the standard nominal models are trivial or else have too many parameters to be tested for goodness of fit.

1.4: The research problem - Application

In order to examine and evaluate any techniques available for analysing ordinal data, the methods need to be applied to an appropriate situation. The data used in this research emanates from a survey of the South Yorkshire Police, designed to evaluate the opinions of the staff on a number aspects of their work and factors affecting it in some way (SSRC (1994)). A factor of interest in the survey is the morale of South Yorkshire Police staff, measured on a five point scale from very high to very low, with a central neutral category. Being ordinal in nature, a discrete version of a one-dimensional continuum, with distances between categories unknown, the variable morale is suitable for the application of the more sophisticated ordinal models - the proportional odds and continuation odds models (Chapter 4).

The generation of appropriate explanatory variables is based on theoretical grounds, in terms of factors that may feasibly be related to the concept of morale sociologically (Viteles (1954)), Hollway (1991)), as well as statistically. The data and variables used are discussed further in sections 3.1.1 to 3.1.5.

The relationship between morale and explanatory factors is examined descriptively using CHAID (Chi-squared Automatic Interaction Detection), a relatively uncommon technique, which helps to parsimoniously describe large data sets (Kass (1980)). CHAID segments the data into specific subsets according to the 'best' predictor variables for describing the behaviour of the response. The method can be used as a precursor for more sophisticated analyses, to identify pertinent factors, or as a purely descriptive tool. The methodology and concept of CHAID is discussed in section 2.2.3, and the technique applied to the South Yorkshire Police data set in section 3.2.

1.5: Organisation of the thesis

Chapter 2 describes the methodologies used in this investigation. Basic exploratory analysis of contingency tables, odds ratios and some measures of association for ordinal variables are discussed in the early sections of the chapter. Models for ordinal variables are then introduced, the more straightforward loglinear and logit modelling procedures are presented, including the binary logistic model. The proportional odds and continuation odds models, designed specifically for an ordinal dependent variable, are then described in some detail. The chapter finishes with a discussion on criteria for assessing the fit of the models described.

Chapter 3 introduces the data from the South Yorkshire Police survey, 1994. The variable of interest, morale, is discussed theoretically and statistically. The potential explanatory variables are discussed, and exploratory data analysis is reported, including the use of the statistical package CHAID, designed specifically for categorical data analysis.

Chapter 4 reports the results of fitting the proportional odds and continuation odds models to the South Yorkshire Police dataset. The implications of models fitted are discussed, and the chapter finishes with a discussion, comparing critically the approaches of the two models to analysing an ordinal response variable.

Chapter 5 draws the investigation to a close, with conclusions to the research. Original work contained within the study is highlighted, some general discussion points are raised, and some ideas for further research in the area are proposed.

Chapter 2: Methodology

2.1: Relevant Developments of the Methodology

This section reviews some relevant literature on methods developed for the analysis of ordered categorical data.

The proportional odds model was introduced by McCullagh (1980). The concept was utilising the ordinal nature of a response variable without the need to assign scores to its levels. The motivation for using this technique is to model the log odds of a 'more favourable' response, thus using a global odds ratio. Many papers have applied the proportional odds model (sometimes referred to as the McCullagh model), including Hutchison (1985), Hastie et al (1989) and Ashby and West (1989), who all give adequate description of the theory of the model, and guidance for diagnostic checking, though interpretation of the implications of the model is not always clear. Hutchison (1985) describes a way of fitting the proportional odds model in GLIM, including testing the proportional odds assumption. Carroll (1993) gives easy to follow description of the methodology, and describes in detail the SAS code to fit the proportional odds model using Proc Logistic.

Cox and Chaung (1984) compare the proportional odds model with the continuation odds model and a base logit model, although the results and conclusions are not clear or easy to follow. Cox and Chaung (1984) do, however, give code for fitting the proportional odds and continuation odds models using the programming languages Fortran and BMDP3R. Other ways to fit the continuation odds model are given by Iyer (1985) and Berridge and Whitehead (1991), both describe the theory of the method quite well. Iyer (1985) gives instruction on fitting the model using GLIM, whilst Berridge and Whitehead (1991) fit the model using SAS, with some clever data manipulation, utilised in this study, and Carroll (1993) also gives clear instruction on the method by Berridge and Whitehead. Iyer (1985) also comments that the continuation odds model is a discrete version of Cox (1972) proportional hazards model for survival data. An alternative to the proportional odds and continuation odds models is given by the late John Anderson (1984) in the form of the stereotype model, Greenland (1994) describes the stereotype model fully, along with the continuation odds and proportional odds model. The stereotype model is in essence a polytomous regression model with an order constraint, imposed by assigning scores to the levels of the dependent variable, therefore representing a drawback of the method.

Analysis of data from contingency tables using logit and loglinear models for ordinal data is discussed by Agresti (1984,1990), Haberman (1974) and Fienberg (1980) among others. The technique CHAID (CHi-squared Automatic Interaction Detection) is introduced by Kass (1980), the method addresses the problem of parsimoniously analysing large data sets. The statistical package SPSS contains a module for CHAID which is explored, the SPSS CHAID user manual also gives some technical insight into the technique.

Testing the goodness-of-fit of ordinal logistic models is an area where there has been relatively little progress. Goodness-of-fit statistics for binary response models are given by Tsiatis (1980) and Hosmer and Lemeshow (1980 and 1989), based on residuals for aggregated data in a particular partition of the covariate space. The test described by Tsiatis (1980) is used less often as he does not give instruction on the partitioning whereas Hosmer and Lemeshow (1980 and 1989) do. Lipsitz et al (1996) extend the Hosmer and Lemeshow test for ordinal data, and this technique for assessing goodness-of-fit is applied to the data in this study. A modification, or extra guideline for using the test given by Lipsitz et al (1996), when discrete/categorical explanatory variables are present, is given in this study.

On the topic of scoring the categories of variables, Agresti (1984, 1990) gives some discussion on this matter, Thomas and Kiwanga (1993) mention different approaches but do not go into much detail and Koch et al (1977) discuss the assignment of integer scores and its merits. None of these give a definitive guide on scoring strategies, but offer the alternative ways of assigning scores to the levels of ordinal variables with discussion. The SPSS/CHAID manual (SPSS (1993)) gives some guidance and instruction on calibrating scores using maximum likelihood estimation, within the statistical package CHAID.

2.2: Exploratory analysis for categorical and ordinal data

2.2.1: Contingency tables

If X and Y denote two categorical variables, with I and J number of levels respectively, then when an individual is classed on both variables there are IJ possible classifications. The responses (X, Y) of individuals have probabilities π_{ij} that they fall in a cell in row i and column j of cross-classification or contingency table (Pearson (1904)).

The probability distribution $\{\pi_{ij}\}$ is the joint distribution of X and Y, and the marginal distributions are the row and column totals obtained by summing the joint probabilities, denoted by $\{\pi_{i+}\}$ for the row variable X and $\{\pi_{+j}\}$ for the column variable Y.

In many cases of contingency tables, one variable is a response or dependent variable (Y, say) and one is an explanatory or independent variable, X. When X is fixed or controlled, Y has a probability distribution for fixed levels of X, rather than defined as a joint distribution for X and Y. Given that an individual is classified in row i of X, then $\pi_{j|i}$ is the probability of classification in column j of Y. The probabilities $\{\pi_{1|i},....,\pi_{J|i}\}$ are the conditional distribution of Y at level i of X. Note that interesting cases when X and Y are both responses may also occur.

Many studies are centered around the conditional distribution of Y at various levels of explanatory variables. For an ordinal response variable it is best to use the cumulative distribution function (cdf), as this keeps the adjacency between levels, and therefore preserves the ordering of the variable. The conditional cdf

$$F_{j|i} = \sum_{b\leq j}\pi_{b|i}\ , j = 1, ..., J$$

is the probability of classification in one of the first j columns, given classification in row i.

2.2.2: Independence between variables

For two variables X and Y, the joint and conditional distributions are related. Using the conditional distribution of Y given X, it is related to the joint distribution of X and Y by:-

$$\pi_{j|i} = \pi_{ij}/\pi_{i+} \qquad \text{for all i and j.}$$

The variables are statistically independent if all joint probabilities are equal to the product of their marginal probabilities, ie $\pi_{ij} = \pi_{i+}\pi_{+j}$. When X and Y are independent :-

$$\pi_{j|i} = \pi_{ij}/\pi_{i+} = (\pi_{i+}\pi_{+j})/\pi_{i+} = \pi_{+j}$$

which means that two variables are independent when the probability of column response j is the same in each row.

Table 2.1 illustrates joint, marginal and conditional distributions for a 2×2 contingency table.

Table 2.1: Notation for joint, conditional and marginal probabilities

	Column 1	Column 2	Total
Row 1	π_{11} ($\pi_{1\|1}$)	π_{12} ($\pi_{2\|1}$)	π_{1+} (1.0)
Row 2	π_{21} ($\pi_{1\|2}$)	π_{22} ($\pi_{2\|2}$)	π_{2+} (1.0)
Total	π_{+1}	π_{+2}	1.0

For sample distributions replace π with p, e.g. $\{p_{ij}\}$ denotes the sample joint distribution in a contingency table, cell frequencies are denoted by $\{n_{ij}\}$ with $n = \Sigma_i\Sigma_j n_{ij}$

being the total sample size, therefore $p_{ij} = n_{ij}/n$. Given row i, the proportion of subjects responding in column j is :-

$$p_{j|i} = p_{ij}/p_{i+} = n_{ij}/n_{i+}$$

where $n_{i+} = np_{i+} = \Sigma_j n_{ij}$.

2.2.3: CHAID

The technique CHAID (CHi-square Automatic Interaction Detection) partitions data into mutually exclusive, and exhaustive, subsets that most adequately describe the behaviour of the response variable (Kass (1980)). Results from CHAID can be useful to aid model building. Often, small groups of explanatory variables are identified and selected from many , and then, say, these variables may be used in subsequent analyses. The technique may also simply be used as an end in itself, in terms of descriptive analysis of a given set of data.

For a categorical or ordinal dependent variable with $j \geq 2$ categories, and a number of categorical or ordinal predictor variables with $k \geq 2$ the CHAID procedure follows an algorithm :-

1. For each predictor in turn, cross-tabulate the categories of the predictor with the categories of the dependent variable (to address the subproblem of optimal categorisation of the predictor variables examined in steps 2 and 3 below).
2. Find the pair of categories of the predictor (only bearing in mind allowable pairs depending on the type and nature of the predictor variable, e.g. monotonic, polytomous etc.) whose $2 \times j$ sub-table is least statistically significantly different. If the significance does not exceed a critical value, then the categories are merged to and the step repeated, using the newly formed compound category.
3. For each compound category consisting of three or more original categories, find the most significant binary split (again constrained by the type of predictor) into

which the merger can be rearranged. If the significance exceeds a critical value, the split is implemented and step 2 repeated.

4. Examine the statistical significance of the relationship between each optimally categorised predictor and the dependent variable, and take the most significant predictor. If this significance exceeds a critical value, then subdivide the data according to the categories of the chosen predictor.
5. For each partition of the data not yet analysed, repeat step 1. This step may be modified by excluding partitions created with a small number of observations.

The following description of the technique refers more to the methodology and use of CHAID within the statistical package SPSS (the technical aspects are obviously in accordance with the ideas proposed by Kass (1980)). The partitioned subsets are referred to as nodes. The analysis can be tailored to a certain 'depth' if required. Depth 0 is the parent node, i.e. the full sample, Depth 1 is the first split of the data on the variable with the strongest statistical association with the response, there will be only 1 variable at depth 1. At depth 2 there could be as many different significant variables as there are levels of the first predictor variable, so depth does not imply number of variables identified (SPSS (1993)).

The predictor variables are all specified as a 'type' before the analysis begins, in order not to break any logical 'rules', most specifically when merging categories. For example, a nominal variable is specified as 'free' as the ordering of the categories is unimportant and it is feasible to merge levels that aren't adjacent to each other. An ordinal predictor may be classed as 'monotonic' or 'float', depending on how missing values are treated or defined within the dataset. CHAID treats missing values as an extra category of each of the variables in the analysis, so if it is feasible that this missing category could be collapsed with any non-adjacent level of the ordinal variable then this variable is assigned as float, however, if the missing category should only be merged with the last category of the variable then it is strictly monotonic. To avoid this problem, treatment of missing values must be sorted out in the dataset in SPSS before the CHAID analysis is performed.

There are constraints and options that CHAID uses when merging and splitting the data on categories of variables. There are two subgroup size constraints. The first is the 'before merge subgroup size', whereby if a subgroup contains fewer observations than the specified value, then it is not analysed further, i.e. not split on another predictor variable and therefore becomes a segment node or completed path. The default is 100, this value is used for the analysis of the South Yorkshire Police data (Chapter 3). The second is the 'after merge subgroup size' which constrains CHAID from splitting the data into a subgroup of less than the specified value. The default is 50, this again will be used in analyses performed later.

CHAID's merge level controls the merging of categories of predictor variables. It takes values between 0 and 1 and is a level of difficulty for combining categories, where the higher the value the more difficult it is for categories to be merged. It is effectively a significance level for the probability that two levels show the same pattern of observations in terms of proportions contained in response levels, below which the categories are deemed to be dissimilar enough to remain distinct. The default of 0.05 is used in this analysis.

Eligibility level is essentially the chosen significance level for accepting a predictor variable's association with the response as statistically significant. The eligibility level takes values between 0 and 1, for the following analyses it is set to 0.05.

CHAID can perform two different types of analyses. It uses either the Nominal or Ordinal method, referring to the nature of the response variable. If the response variable is nominal, then CHAID will produce output in terms of proportions of observations contained in response categories for the subgroups created, whereas the ordinal method gives results pertaining mean response scores.

The nominal method assumes cell counts for a two way table, say, between variables A and B with levels 1 to I and 1 to J respectively, occur from a saturated loglinear model:-

$$\ln (F_{ij}/(1-W_{ij})) = \lambda + \lambda_{(A)i} + \lambda_{(B)j} + \lambda_{(AB)ij}$$

Where F_{ij} denotes expected cell counts and W_{ij} is the average sampling weight.

The nominal method tests for independence by testing whether the parameter $\lambda_{ij}^{AB}=0$.

An ordinal dependent variable may not necessarily be analysed by the ordinal method in CHAID, although it is probably beneficial to do so as the way the package calculates probabilities takes into account the ordinal nature of the response to give more powerful inference. Within the ordinal method, when testing for independence between variables, CHAID utilises category scores and therefore uses an unsaturated model, the Y association model (Magidson (1992)) :-

$$\ln (F_{ij}/(1-W_{ij})) = \lambda + \lambda_{(A)i} + \lambda_{(B)j} + x_i(y_j - \bar{y})$$

Where y_j is the category score for the jth level of B, x_i is an unknown coefficient for the y_j's and $\bar{y}$ is the mean score for the response variable.

CHAID tests for independence by testing whether $x_1 = x_2 = \ldots = x_I$.

The ordinal method of calculating probabilities ignores non-relevant sources of non-independence, i.e. it concentrates on the Y association involving the ordinality of the response, therefore uses fewer degrees of freedom making the test more powerful.

CHAID can also estimate scores for levels of an ordinal dependent variable if they are unknown. The package uses maximum likelihood calibration to estimate response level scores that are most likely to be associated with a particular explanatory variable. As a single predictor variable is used as a calibration instrument, the estimated scores can vary, possibly dramatically, between different covariates, therefore is a degree of arbitrariness using this method to assign scores to the dependent variable, as there is with all scoring systems (see section 2.3.1 for discussion on scoring). As an example of

CHAID's score estimation, the South Yorkshire Police data, described in Chapter 3, is used. Respondent's own morale (omor) is most strongly associated statistically with job satisfaction (jobsat) (shown in section 3.1.5), therefore if we use this covariate to calibrate scores for the response, the following results are given :-

Table 2.2: CHAID scores for omor calibrated using jobsat

omor	v. high	high	neither	low	v. low
est. score	0	23.2	56.19	83.3	100

The end category scores are constrained to be 0 and 100, and the order, i.e. ascending, can be reversed so the scale is 100 to 0, with the same inter-category distances. The category scores above are not too much of a departure from equidistant scores. However, if we use a different predictor to calibrate the scores, say, promotions given to those who earn them (promearn), which is also strongly statistically associated with the response, the following scores are obtained :-

Table 2.3: CHAID scores for omor calibrated using promearn

omor	v. high	high	neither	low	v. low
est. score	0	10.16	24.06	40.36	100

for which the distances between first four categories are not too dissimilar, but between levels of morale low and very low, there is a distance greater than that between the level very high morale at the opposite end of the scale.

Estimating scores does not affect the analysis procedure, but obviously the choice of calibration variable may affect any substantive conclusions, if one is using mean scores, as given when using the ordinal method. Therefore care should be taken when interpreting the results. One can also assign scores within CHAID, or only estimate some, rather than all, scores. Score estimation within CHAID is a very useful tool for descriptive purposes, and more so when a modelling procedure requires assignment of scores to the levels of a variable, and no obvious choice exists. The package only estimates scores for dependent variables. If necessary, one could temporarily use an explanatory variable as the response purely for the purpose of estimating scores for its categories, using, say, the real response variable as the calibration instrument. This produces scores for the explanatory variable that are most likely to be associated with the response, this process is employed and discussed in the application of the proportional odds model in Chapter 4. The application of CHAID to the South Yorkshire Police data is given in section 3.2.

2.2.4: Odds and Odds ratios

Using the 2×2 table 2.1, within row 1, the 'odds' of a response in column 1 as opposed to column 2 is defined as :-

$$\Omega_1 = \pi_{1|1}/\pi_{2|1}$$

and similarly within row 2, the corresponding odds are :-

$$\Omega_2 = \pi_{1|2}/\pi_{2|2}$$

Each Ω_i is non-negative, and greater than 1.0 if response 1 is more likely than response 2, eg if $\Omega_1 = 4.0$, then response 1 is 4 times as likely as response 2, within the first row. The ratio of these odds, Ω_1 and Ω_2, is :-

$$\theta = \Omega_1/\Omega_2 = \frac{(\pi_{11}/\pi_{12})}{(\pi_{21}/\pi_{22})} = \frac{\pi_{11}\pi_{22}}{\pi_{12}\pi_{21}} \qquad (2.1)$$

Called, logically, the odds ratio.

Independence between the row and column variables, X and Y, is equivalent to $\theta = 1$. When $1 < \theta < \infty$, subjects in row 1 are more likely to make response 1 than are subjects in row 2, eg if $\theta = 4.0$, the odds of the first response are 4 times higher in row 1 than in row 2. When $0 < \theta < 1$, the first response is less likely in row 1 than in row 2. If a cell has zero probability, $\theta = 0$ or ∞.

For sample frequencies $\{n_{ij}\}$, the sample odds ratio is given by :-

$$\theta = \frac{n_{11}n_{22}}{n_{12}n_{21}}. \qquad (2.2)$$

The sample odds ratio does not change when cell frequencies within a row are multiplied by a constant, or similarly when cell frequencies within a column are multiplied by a constant.

The odds ratio is invariant to changes in orientation of the table, i.e. rows become columns and vice versa. Two different values for θ represent the same level of association, in opposite directions, when one is the inverse of the other, eg if $\theta = 0.25$, the odds of response 1 are 0.25 as high in row 1 as in row 2, and/or equivalently the same odds are 4 times as high in row 2 as in row 1 (as $1/0.25 = 4$).

The log odds ratio, $\log(\theta)$, is sometimes used, especially in logit models where the parameters actually are the log odds ratios, so that the values of parameters are not constrained, ie they can take any real value rather than just positive values. Independence corresponds to $\log(\theta)= 0$, and the log odds ratio is symmetric about this value. Therefore the property described above that two values for θ, where one is the inverse of the other, represent the same level of association, now becomes two values of $\log(\theta)$ the same except for sign, i.e. $\log(4)= 1.39$ and $\log(0.25)= -1.39$.

2.2.5: Odds ratios and ordinal data

For an ordinal variable, a response variable Y, say, (note that the identification of dependent and independent variables is unnecessary for odds ratios) with categories 1, ..., k, (k>2), in order to calculate an odds ratio one would have to collapse the data to binary in some way. If we consider first an original table with a binary explanatory covariate, X :-

Table 2.4: Probability distribution for a binary covariate and an ordinal response with k levels

	Y=1	Y=2	...	Y=k
X=1	p_{11}	p_{12}	...	p_{1k}
X=2	p_{21}	p_{22}	...	p_{2k}

Two possible ways of collapsing this table to k-1 2×2 sub-tables are, firstly :-

Categorisation of the response into all possible divisions of 'success' and 'failure' or 'favourable' and 'unfavourable', assuming the categories are ordered 'best' (Y=1, say) to 'worst' (Y=k) in some sense, or similarly vice versa :-

Table 2.4.1: Category 1 vs Categories 2 to k

	Y=1	Y=2 to k
X=1	p_{11}	$p_{12}+ \ldots +p_{1k}$
X=2	p_{21}	$p_{22}+ \ldots +p_{2k}$

Table 2.4.2: Categories 1 and 2 vs Categories 3 to k

	Y=1 + Y=2	Y= 3 to k
X=1	$p_{11}+p_{12}$	$p_{13}+ \ldots +p_{1k}$
X=2	$p_{21}+p_{22}$	$p_{23}+ \ldots +p_{2k}$

.

.

Table 2.4.k-1: Categories 1 to k-1 vs Category k

	Y=1 to k-1	Y= k
X=1	$p_{11}+ \dots +p_{1k-1}$	p_{1k}
X=2	$p_{21}+ \dots +p_{2k-1}$	p_{2k}

When collapsing the data, it is important to consider the logic of collapsing certain categories. For example, if you have a variable with categories ordered as very high, high, low and very low, combining the categories very high, high and low, or high, low and very low, can make interpretations of the merged category difficult, as the levels have contrary interpretations.

The odds ratios for the k-1 tables can be calculated, to give an idea of the differences in effect on the different dichotomies of the response. Sub-divisions in this manner form the basis of logistic regression and proportional odds modelling procedures, for ordered categorical data with more than two categories. For the latter in particular, from this approach, insight may be gained into whether the odds ratios across the k-1 divisions are approximately constant or similar, with a view to using a global odds ratio to describe the odds of 'success' vs 'failure' for the covariate.

Another way to collapse the response, assuming again that levels are ordered 'best' to 'worst' or vice versa, is to make the divisions according to membership of the 'most favourable' category available:-

Table 2.5.1: Category 1 vs Categories 2 to k

	Y=1	Y=2 to k
X=1	p_{11}	$p_{12}+ \dots +p_{1k}$
X=2	p_{21}	$p_{22}+ \dots +p_{2k}$

Table 2.5.2: Category 2 vs Categories 3 to k

	Y=2	Y= 3 to k
X=1	p_{12}	$p_{13}+ \dots +p_{1k}$
X=2	p_{22}	$p_{23}+ \dots +p_{2k}$

Table 2.5.k-1: Category k-1 vs Category k

	Y=k-1	Y= k
X=1	p_{1k-1}	p_{1k}
X=2	p_{2k-1}	p_{2k}

As mentioned for the previous collapsing of the data, though only applying to the right hand side of the dichotomy, it is important to make sure the collapsing of categories does not make interpretation difficult, i.e. that no levels with contrasting meanings are combined.

The collapsing of the response in this way allows comparison of the odds that given an individual has responded in category j or worse, they have responded in the most favourable of these categories available, j. The interpretations of these odds ratios are different from those pertaining to 'success' and 'failure'. Sub-dividing the response in this manner forms the basis for the continuation odds modelling procedure, which seeks to describe the k-1 tables above with a single global odds ratio.

2.2.6: Measures of association for ordinal data

Concordance and discordance are measures similar to that of Pearson correlation. When the ordering of a pair of individuals on each of two ordinal variables, X and Y, is observed, the pair can be classified as concordant if the individual ranking higher on X also ranks higher on Y. The pair is discordant if the subject ranking higher on X ranks lower on Y, and the pair is tied if they both have the same classification on X and/or Y. To illustrate, the following example uses data from the South Yorkshire Police survey described in Chapter 3 (SSRC (1994)) :-

Table 2.6: Cross classification of Own Morale by Communication with More Senior Officers/Managers

Communication	Respondent's		Own	Morale	
	Very High (VH)	High (H)	Neither (N)	Low (L)	Very Low (VL)
Very Good (VG)	51	142	71	33	7
Good (G)	58	302	244	120	37
Neither (N)	35	169	218	139	41
Bad (B)	5	54	65	103	38
Very Bad (VB)	1	5	14	13	27

Consider two individuals, one classified in the cell (VG, VH) and the other (G, H). This pair is concordant as the first subject is ranked higher than the second on both scales. Each of the 51 subjects in cell (VG, VH) form concordant pairs when matched with each of the 302 classified (G, H), so there are $51\times302=15402$ concordant pairs from those two cells. The 51 individuals classified (VG, VH) also form concordant pairs with each of the other (244 + 120 + 37 + 169 + 218 + 139 + 41 + 54 + 65 + 103 + 38 + 5 + 14 + 13 + 27) individuals they are ranked higher than on both variables. Similarly, the 142 subjects in cell (VG, H) are part of a concordant pair when matched with the (244 + 120 + 37 + 218 + 139 + 41 + 65 + 103 + 38 + 14 + 13 + 27) individuals they are ranked higher than on both variables.

The total number of concordant pairs, denoted by C, equals :-

$$
\begin{aligned}
C &= 51(302+244+120+37+169+218+139+41+54+65+103+38+5 \\
&\quad +14+13+27) \\
&\quad +142(244+120+37+218+139+41+65+103+38+14+13+27) \\
&\quad +71(120+37+139+41+103+38+13+27) +33(37+41+38+27) \\
&\quad +58(169+218+139+41+54+65+103+38+5+14+13+27) \\
&\quad +302(218+139+41+65+103+38+14+13+27) \\
&\quad +244(139+41+103+38+13+27) \\
&\quad +120(41+38+27) \ldots +65(13+27) +103(27) \\
&= 736{,}012
\end{aligned}
$$

The number of discordant pairs of observations, D, is :-

$$D = 142(58+35+5+1) + 71(58+302+35+169+5+54+1+5) + \ldots + 103(1+5+14) + 38(1+5+14+13) = 334,075.$$

Therefore in this example, C>D suggests that lower morale has a tendency to occur with the feeling that communication is bad, and higher morale to occur with good communication.

A measure of association that uses the above statistics is gamma, γ, defined as the difference between the probabilities of concordance and discordance (Goodman and Kruskal (1954)). For the sample case :-

$$\gamma^{hat} = (C - D) / (C + D).$$

As for a correlation coefficient, the range of gamma is $-1 \le \gamma \le 1$, and as $\gamma \rightarrow |1|$, the stronger the association between the two variables. Independence between variables implies that $\gamma=0$, but the inverse is not necessarily true, as some non-linear association, eg a U-shaped joint distribution , may not be detected by gamma :-

Table 2.7: U-shaped joint distribution of two variables X and Y

	y_1	y_2	y_3
x_1	0.2	0	0.2
x_2	0.2	0	0.2
x_3	0	0.2	0

Here both C and D = 0.08, therefore $\gamma^{hat} = 0$, but it seems there is some form of association between the variables as the distribution of proportions in the cells have a distinct pattern.

For the morale example above it was found that C = 736,012 and D =334,075. Of the concordant and discordant pairs, 68.78% are concordant and 31.22% discordant, therefore the difference in proportions gives $\gamma^{hat} = 0.376$, indicating a moderately strong tendency for morale to be higher when communication with more senior managers/officers is deemed better. Measures of association for ordinal variables are discussed fully in Agresti (1984).

2.3: Standard models modified for ordinal variables

The following sub-sections briefly outline how some standard categorical modelling procedures can be adapted to accommodate ordinal information. The adaptation of standard loglinear and logit models for ordinal variables hinges on the use of scores for the levels of explanatory variables, rather than the utilisation of ordinality in a dependent variable. These models are examples of Generalised Linear Models (GLM). In brief, GLMs are a class of models first developed by Nelder and Wedderburn (1972). GLMs are models basically specified by three components - A random component which identifies the probability distribution of the response variable; a systematic component which specifies the form of the model, in terms of the linear function of the explanatory variables; a link function which describes the relationship between the systematic component and the expected value (mean) of the random component. Full details of GLM's are contained in McCullagh and Nelder (1989).

2.3.1: Scoring the levels of ordinal variables

One of the main objectives of this research is to examine models and methods that do not require scoring of the levels of an ordinal variable. However some of the methods and examples described within this thesis require the assignment of scores, so that the ordinality of a variable is utilised in some way, rather than lost to nominality.

An ordinal variable is quantitative in the sense that each level on its scale can be compared in terms of whether it corresponds to a greater or smaller magnitude of a certain characteristic than another level. In reality, it is almost impossible to measure the ‘distance’ between categories of an ordinal variable, and therefore assigning scores is often arbitrary.

Sometimes a score may be an actual numerical response, eg the number of cancerous lungs (0, 1, 2), or the midpoint of an interval, if the variable is a grouping of an underlying continuous variable, eg age (<16, 16-25, 26-39, 40+) or salary (<£6k, £6k-£12k, £13k-£20k, £20k+).

Where no obvious choice of scores exists, integer scores are often used. Assuming the levels of an ordinal variable are equally spaced leads to easy interpretation of statistics or models fitted for that variable (Koch et al (1977)).

Alternatively, if it is not appropriate to assume equal spacing, and there are suspicions or further information about inter-category distances, one could assign a variety of ‘reasonable’ sets of scores, to see if, or how much, substantive conclusions depend on the choice of scores. One may settle for a set of scores that gives the most desirable results, though care must be taken when interpreting and/or reporting results in such cases.

Another approach is to use distributional scores. In some cases it may be assumed that there is an underlying continuous measurement scale for which a particular distribution, with distribution function F, is suitable, eg a normal or uniform distribution. Scores for the categories of the variable could be functions of the ranks. For example, scores may be estimated from the data by evaluating $F^{-1}(r_j/(N+1))$, where r_j is the midrank score for category j, for j=1, ..., k, and N is the total number of observations. Many statisticians have voiced concerns over the use of such scoring methods and prefer preassigned scores. For further discussion see Thomas and Kiwanga (1993).

Scores for categories can also be estimated from the data to make them optimal in some sense, these are called optimal or calibrated scores. CHAID in SPSS (see section 2.2.3) can estimate scores for an ordinal dependent variable. Using maximum likelihood estimation the package calibrates scores from a particular explanatory covariate, so that those scores are most likely to be associated with that covariate. A drawback to this method is that the scores vary with the choice of calibration instrument, i.e. explanatory variable, so substantive conclusions will therefore probably be dependent on the scores obtained.

Most methods of assigning scores to the categories of an ordinal variable and/or their interpretation are subjective. Methods of estimating scores are dependent on the data used to calibrate them, and therefore not in accordance with a preconceived suspicion. Agresti (1984) gives some discussion on scoring. In this investigation preassigned integer scores and CHAID estimated scores are used, the motivation for which is to compare the results obtained by the different scoring methods.

2.3.2: Loglinear modelling

Loglinear analysis models cell frequencies or probabilities from contingency tables, therefore there is no dependent variable as such. A loglinear model shows how the factors affect the distribution of observations within the cells of a table, and how the factors associate with each other.

Earlier, in section 2.2.2, it was seen that if two variables, X and Y, with levels i and j, are independent, then $\pi_{ij} = \pi_{i+}\pi_{+j}$ for all i and j. Equivalently for expected cell frequencies $\{m_{ij}=n\pi_{ij}\}$, if X and Y are independent then $m_{ij}=n\pi_{i+}\pi_{+j}$ for all i and j. Therefore on a logarithmic scale, independence corresponds to :-

$$\log m_{ij} = \log n + \log \pi_{i+} + \log \pi_{+j}$$

Referred to alternatively as :-

$$\log m_{ij} = \mu + \lambda_{(X)i} + \lambda_{(Y)j} \qquad (2.3)$$

where μ is the overall mean of the log cell frequencies, and the λ parameters are the effects of the variables X and Y, on the log cell frequencies adjusting for the overall mean.

This is called the loglinear model for independence. In standard loglinear modelling, the next more complex model is the saturated model (saturated means there are as many parameters in the model as cells) incorporating an interaction between the two variables :-

$$\log m_{ij} = \mu + \lambda_{(X)i} + \lambda_{(Y)j} + \lambda_{(XY)ij} \qquad (2.4)$$

Which is the most general model for two variables. It provides a perfect fit to the data and has no degrees of freedom, i.e. it has a parameter for every cell, therefore it is not really useful, and shows nothing new.

However, if one or both of the variables are ordinal, an unsaturated specialised loglinear model can be formed. Firstly suppose that X and Y are both ordinal with known category scores u_i and v_j respectively, then a simple model that utilises the ordinality in the variables and accounts for an association between them is :-

$$\log m_{ij} = \mu + \lambda_{(X)i} + \lambda_{(Y)j} + \beta(u_i-\bar{u})(v_j-\bar{v}) \qquad (2.5)$$

where $\bar{u}$ and $\bar{v}$ are the means of the scores u_i and v_j.

This model is called the uniform association model. Note that this model only requires 1 more parameter than the independence model as opposed to ij extra parameters of the saturated standard model, and does not require extra parameters if the number of levels of X and Y increases. This increase in efficiency is the biggest advantage of such a model, further to the employment of the ordinality in the variables (Haberman (1974)). The fit of the model can be assessed using a chi-square (χ^2) statistic, therefore this model will have more degrees of freedom (df) than the corresponding standard loglinear model, i.e. the ordinal loglinear model is more parsimonious. β describes the association between X and Y, therefore if $\beta=0$ then the variables are independent. The term $\beta(u_i-\bar{u})(v_j-\bar{v})$ reflects a deviation of log m_{ij} from the independence model. If $\beta>0$ then more observations are expected to have (large X, large Y) or (small X, small Y) values, than if X and Y were independent, and if $\beta<0$ one would expect more (large X, small Y) or (small X, large Y) values.

If only one of the variables, say Y, is ordinal with known category scores v_j, a similar model to that above is given by the row effects model:-

$$\log m_{ij} = \mu + \lambda_{(X)i} + \lambda_{(Y)j} + \tau_i(v_j-\bar{v}) \qquad (2.6)$$

The $\{\tau_i\}$ are row effects (hence the name of the model, although it can also be called the column effects model if it is the column variable that is ordinal). Within a particular row, i, the deviation of log m_{ij} from independence is a linear function of the ordinal variable. If $\tau_i=0$, X and Y are deemed to be independent. If $\tau_i>0$, then in row i the probability of classification above $\bar{v}$ on Y is higher than would be expected if X and Y were independent. If $\tau_i<0$ then observations in row i are more likely to be classified at the lower end of the scale of Y.

These concepts in loglinear modelling can easily be applied to higher dimensions of variables, for instance, consider another variable Z with k categories. If all 3 variables are nominal, a loglinear model more complex than the independence model but unsaturated, describing the association between the variables and the distribution of observations would be :-

$$\log m_{ij} = \mu + \lambda_{(X)i} + \lambda_{(Y)j} + \lambda_{(Z)k} + \lambda_{(XY)ij} + \lambda_{(XZ)ik} + \lambda_{(YZ)jk} \qquad (2.7)$$

which includes all 2 way interactions between the variables, i.e. the pairwise partial associations, but excludes a 3 factor interaction, and thus is not saturated. The parameters can be interpreted as for the previous loglinear models given for two dimensions.

If all these variables are ordinal with X and Y having category scores u_i and v_j as for model (2.5) and Z with category scores w_k, a model utilising this information, equivalent to (2.7) but more parsimonious can be given by :-

$$\log m_{ij} = \mu + \lambda_{(X)i} + \lambda_{(Y)j} + \lambda_{(Z)k} + \beta_{(XY)}(u_i-\bar{u})(v_j-\bar{v}) + \beta_{(XZ)}(u_i-\bar{u})(w_k-\bar{w})$$
$$+ \beta_{(YZ)}(v_j-\bar{v})(w_k-\bar{w}) \qquad (2.8)$$

which only has 3 more parameters than the independence model compared to (ij+ik+jk) more parameters than the independence model for (2.7).

If, say, X is nominal while Y and Z are ordinal, the row effects model (2.6) for 2 dimensions can be extended to give :-

$$\log m_{ij} = \mu + \lambda_{(X)i} + \lambda_{(Y)j} + \lambda_{(Z)k} + \tau_{(XY)i}(v_j-\bar{v}) + \tau_{(XZ)i}(w_k-\bar{w})$$
$$+ \beta_{(YZ)}(v_j-\bar{v})(w_k-\bar{w}) \qquad (2.9)$$

where the X-Y and X-Z association terms have the same form as in the row effects model, and the Y-Z association term has the same form as in the uniform association model (2.5). This model, again, is more parsimonious than a model treating all variables as nominal.

The ordinality of variables can be taken into account by these types of models in different ways, also, for instance, log-multiplicative models are a form of ordinal loglinear models which estimate the scores u_i and v_j as parameters. For two ordinal variables X and Y, if the ordered scores u_i and v_j from before are treated as unknown parameters μ_i and ν_j, the two dimensional log-multiplicative model is given by :-

$$\log m_{ij} = \mu + \lambda_{(X)i} + \lambda_{(Y)j} + \beta\mu_i\nu_j \qquad (2.10)$$

which simplifies to the loglinear model for independence if $\beta = 0$. The score parameters μ_i and ν_j are estimated from the data to give the model best fit, and therefore probably should not be used for any other purpose than in the model itself.

2.3.3: Logit modelling

Logit models can be equated to loglinear models, but take a slightly different form, logit models describe the effects of a set of explanatory variables on a response variable, but do not describe associations between explanatory variables. Logit models with respect to ordinal explanatory variables are considered here, whilst in section 2.3.4, the binary logistic model is described in the context of a basis for the more sophisticated proportional and continuation odds models.

Consider 3 categorical variables X, Y and Z with levels i, j and k respectively, where Z is a dichotomous response variable. $\{\pi_{ijk}\}$ and $\{m_{ijk}\}$ denote cell probabilities and frequencies. The conditional probability of response k at levels i and j of X and Y is $\pi_{k(ij)} = \pi_{ijk}/\pi_{ij+}$. The logit for Z is the log odds of an event (a response) :-

$$\log [\, \pi_{2(ij)}/(1-\pi_{2(ij)})]$$

$$= \log (\pi_{ij2}/\pi_{ij1})$$

$$= \log (m_{ij2}/m_{ij1}).$$

First suppose X and Y are nominal, the logit of Z could be modelled by :-

$$\log (m_{ij2}/m_{ij1}) = \alpha + \tau_{(X)i} + \tau_{(Y)j} + \tau_{(XY)ij} \qquad (2.11)$$

Where α is the log odds that Z=2 in this case, if the τ parameters are zero. $\{\tau_{(X)i}\}$ pertains to the partial association between X and Z, and $\{\tau_{(Y)j}\}$ pertains to the partial association between Y and Z. The $\tau_{(XY)ij}$ terms are accounting for the joint effects of X and Y on Z. If all $\tau_{(X)i}=0$, then Z is conditionally independent of X, given Y, and the same applies to Y and Z when all $\tau_{(Y)j}=0$.

If X and Y are ordinal, with, as before, known category scores u_i and v_j, in order to take into account this quantitative information, the logit of Z can be modelled by :-

$$\log (m_{ij2}/m_{ij1}) = \alpha + \beta_{(X)}(u_i-\bar{u}) + \beta_{(Y)}(v_j-\bar{v}) + \beta_{(XY)}(u_i-\bar{u})(v_j-\bar{v}) \qquad (2.12)$$

Where $\beta_{(X)}$ and $\beta_{(Y)}$ represent local log odds ratios for the partial X - Z, and Y - Z associations, and $\beta_{(XY)}$ represents the joint effects of X and Y on Z therefore if $\beta>0$ the log odds that Z=2, i.e. the logit of Z, increases. If all β parameters are zero, this means that Z is independent of X and Y.

If X is nominal and Y is ordinal with scores v_i, a combination of the above logit models can be applied :-

$$\log (m_{ij2}/m_{ij1}) = \alpha + \tau_{(X)i} + \beta_{(Y)}(v_j-\bar{v}) + \tau_{(XY)i}(v_j-\bar{v}) \qquad (2.13)$$

The interpretation of parameters is the same as for the corresponding parameters in the two previous models (2.11) and (2.12).

Logit models (2.11), (2.12) and (2.13) can be fitted for higher dimensions of variables similarly to loglinear models.

The loglinear and logit models above accommodate ordinal information mainly with respect to explanatory covariates, the proportional odds and continuation odds models, discussed later, are designed for an ordinal response.

2.3.4: The Binary Logistic Model

The binary logistic model is a fairly typical logit model, and in the description below does not account for ordinality in any of the variables. The model is discussed in more detail, mainly because it is like a building block for the more sophisticated proportional and continuation odds models, which account for an ordinal response variable. The proportional and continuation odds models degenerate to the binary logistic model when the response variable has only two categories. Also when these more sophisticated models fail to be appropriate, the binary logistic offers a method of modelling the response in at least some context of interest.

Let Y be a binary response variable where Y=1 for success and 0 for failure (possibly a dichotomised ordinal response). If there are m explanatory covariates, $x_1, x_2, \ldots, x_m$, thought to influence success, then an individual, i, with specific covariates, $\mathbf{x}_i$, has probability of success $\pi(\mathbf{x}_i)$. The response variable Y_i follows a Bernoulli distribution.

The logistic regression model is given by :-

$$\log\{\pi(\mathbf{x}_i) / [1-\pi(\mathbf{x}_i)]\} = \beta_0 + \boldsymbol{\beta}'\mathbf{x}_i \tag{2.14}$$

where $\boldsymbol{\beta}' = (\beta_1, \beta_2, \ldots, \beta_m)$, and $\mathbf{x}_i = (x_{1i}, x_{2i}, \ldots, x_{mi})$.

If we let $\eta_i = \beta_0 + \boldsymbol{\beta}'\mathbf{x}_i$ and rearrange the model formula above, the probability of success can be obtained :-

$$\pi(\mathbf{x}_i) = \exp(\eta_i) / [1+\exp(\eta_i)] \tag{2.15}$$

The probability of success, $\pi(\mathbf{x}_i)$ is modelled via the logit transformation. This transforms $\pi(\mathbf{x}_i)$ from the range (0, 1) to the range $(-\infty, +\infty)$, so the parameters of the model β_0 and $\boldsymbol{\beta}$ are not restricted and can take any real value.

The binary logistic model is modelling the log odds of a successful response for an individual with specific covariates $\mathbf{x}_i$. The $\boldsymbol{\beta}$ parameters show how changes in the values of covariates affect the probability of success.

To compare 2 individuals a and b, with covariates $\mathbf{x}_a$ and $\mathbf{x}_b$, the difference in log odds of success is :-

$$\log [\pi(x_a)/\{1-\pi(x_a)\}] - \log [\pi(x_b)/\{1-\pi(x_b)\}] = \boldsymbol{\beta}'(\mathbf{x}_a-\mathbf{x}_b)$$

rearranging :-

$$\log \frac{[\pi(x_a)/\{1-\pi(x_a)\}]}{[\pi(x_b)/\{1-\pi(x_b)\}]} = \boldsymbol{\beta}'(\mathbf{x}_a-\mathbf{x}_b) \qquad (2.16)$$

Therefore $\boldsymbol{\beta}'(\mathbf{x}_a-\mathbf{x}_b)$ represents the log odds ratio of success for an individual with covariates x_a compared to an individual with covariates x_b.

2.3.4.1: Parameter estimation for the binary logistic model

Consider a random sample of N individuals, all with a set of covariates, $\mathbf{x}_i$, and a response of either success, $Y_i=1$, with probability $\pi(\mathbf{x}_i)$, or failure, $Y_i=0$, with probability $1-\pi(\mathbf{x}_i)$. Y_i is therefore a Bernoulli random variable, $B\{\pi(\mathbf{x}_i)\}$, with probability of success related to β_0 and $\boldsymbol{\beta}$.

The likelihood is then proportional to a product of N such random variables :-

$$L \propto \prod_{i=1}^{N} \pi(\mathbf{x}_i)^{Y_i} [1 - \pi(\mathbf{x}_i)]^{1-Y_i}$$

The parameters β_0 and $\boldsymbol{\beta}$ are estimated by calculating partial derivatives and equating to zero. In most cases these equations can only be solved numerically using Newton-Raphson type iterative techniques.

Alternatively, Least Squares could be used to estimate the parameters. Iteratively reweighted least squares has similarities to Newton-Raphson. The statistical package SAS uses iteratively weighted least squares in the LOGISTIC procedure.

The methods of estimation are asymptotically the same, and therefore converge to identical parameter estimates. However, often in the real world, with finite sample sizes, parameter estimates from the different estimation techniques will be very similar, but not exactly the same (Agresti (1984)).

2.3.5: Interactions between explanatory covariates

The following section briefly describes interaction terms, which are a little more complex to interpret than main effect parameters in any given model. If we have a categorical response variable Y (ordinal or binary) and two explanatory variables A and B, which could be of any scale type, but for illustration's sake, say, they are categorical, if the association between the response variable and A is the same within each level of B (or vice versa), then there is said to be no interaction between A and B. In general, the absence of interaction is characterised by a model which contains no second or higher order terms involving two or more variables.

If interaction is present, then the association between the response and A differs, or depends in some way on the level of B (or vice versa). The implication of this is that any conclusions regarding the odds of a response for an individual with characteristic A should be made with respect to a specific level of B, i.e. the effect of A depends on the specific level of B.

This concept also applies to interactions between more than two variables, for example, if a third explanatory variable, C, was involved in an interaction with both A and B, then the association between the response and C differs, or depends in some way on the levels of A and B, i.e. the odds of an event for an individual with characteristic C should be estimated with respect to a specific level of A and a specific level of B.

Determining whether an interaction is present, i.e. significant, is fairly straightforward. One must first decide whether an interaction between two or more variables is plausible, and consequently if an interaction is logically possible, then a term is added to the model and its significance can be determined by a Wald Chi-square statistic, which measures the probability that the term is actually equal to zero, and therefore makes no significant improvement to the fit of the model, as for main effects parameters (Wald (1943)).

The interaction parameter estimate on its own cannot be interpreted meaningfully, as the value of an interaction term is an adjusting term, whereby the main effects of two variables, say A and B again, at specific levels or values are combined and then adjusted by the value of the respective interaction term to determine the effect of a certain level A at a certain level of B. Discussion of interactions in logistic regression models can be found in Hosmer and Lemeshow (1989).

2.4: Models for an ordinal dependent variable

2.4.1: The Ordinal Logistic Regression Model

For an ordinal response variable, Y, with categories 1, ..., K (K>2), ordinal logistic regression creates all possible dichotomies of the variable, without violating the adjacency of any pair of levels. For example, if k=3, then dichotomies for calculating the odds of category 1 versus 2 and 3, and categories 1 and 2 versus 3 are created, for given covariates. However, the odds of category 2 versus 1 and 3 are not estimated as the variable is assumed to follow a continuum. This process is described further in section 2.2.5 (illustration contained in tables 2.4.1 to 2.4.k-1).

The odds that an individual with specific covariates, ($\mathbf{x}_i$), responds in category j or less of Y, i.e. a dichotomy of the variable representing, say, success, are given by :-

$$P(Y_i \leq j)/[1-P(Y_i \leq j)] = \gamma_j(\mathbf{x}_i)/\{1-\gamma_j(\mathbf{x}_i)\} \qquad (2.17)$$

where

$$P(Y_i \leq j) = \gamma_j(\mathbf{x}_i) = \pi_1(\mathbf{x}_i) + \pi_2(\mathbf{x}_i) + ... + \pi_j(\mathbf{x}_i). \qquad (2.18)$$

Ordinal logistic regression produces results equivalent to simultaneously fitting K-1 binary logistic models. The ordinal logistic model takes the form :-

$$\log [\gamma_j(\mathbf{x}_i)/\{1-\gamma_j(\mathbf{x}_i)\}] = \alpha_j + \boldsymbol{\beta}_j'\mathbf{x}_i \qquad 1 \leq j \leq K-1 \qquad (2.19)$$

This model, effectively, gives no more information than fitting the k-1 binary logistic models separately. A couple of advantages of using ordinal logistic regression over binary include the fact that it is unnecessary to create dichotomies 'by hand' as the analysis is automatically performed on the binary cutpoints, and a little time may be saved by running a program once, rather than several (k-1) times. As ordinal logistic regression offers no real technical improvement over methods for analysing nominal data, its use is not pursued for the purpose of this research.

2.4.2: The Stereotype Model

If a dependent variable, Y, has k>2 categories, and each observation has a set of covariates, $\mathbf{x}_i$, the most flexible logistic regression model is the polytomous model, given by :-

$$\log [\pi_j(\mathbf{x}_i)/\{\pi_0(\mathbf{x}_i)\}] = \alpha_j + \boldsymbol{\beta}_j'\mathbf{x}_i \qquad 1 \leq j \leq K-1 \qquad (2.20)$$

Where $\pi_j(\mathbf{x}_i)$ is the probability of a response in category j of Y given the set of covariates, with $\pi_0(\mathbf{x}_i)$ being the probability of a response in level zero. The odds of each of the response levels above zero are compared to that category in terms of logit functions. For example, if k=3 with the levels coded 0,1,2, then the polytomous model computes the logit functions :-

$$g_1(\mathbf{x}_i) = \log [\pi_1(\mathbf{x}_i)/\{\pi_0(\mathbf{x}_i)\}]$$

$$g_2(\mathbf{x}_i) = \log [\pi_2(\mathbf{x}_i)/\{\pi_0(\mathbf{x}_i)\}]$$

The logit for comparing Y=2 to Y=1 can be obtained as the difference between $g_1(\mathbf{x}_i)$ and $g_2(\mathbf{x}_i)$ (Hosmer and Lemeshow (1989)).

If Y is ordinal in nature, the coefficients $\boldsymbol{\beta}_j$ from the polytomous model (2.20) may be replaced by :-

$$\boldsymbol{\beta}_j = \boldsymbol{\beta} s_j \qquad 1 \le j \le K\text{-}1$$

where the parameter s_j represents the score attached to outcome y_j (Anderson (1984)). From this modification of the polytomous model, to allow the utilisation of the ordering of the response categories, the stereotype model is defined (Greenland (1994)) :-

$$\log [\pi_j(\mathbf{x}_i)/\{\pi_0(\mathbf{x}_i)\}] = \alpha_j + \boldsymbol{\beta} s_j{}'\mathbf{x}_i \qquad 1 \le j \le K\text{-}1 \qquad (2.21)$$

The parameters $\boldsymbol{\beta} s_j$ represent the log odds ratio for $Y=y_j$ versus $Y=y_0$ per unit increase in $\mathbf{x}_i$. As the scores s_j are multiplicative on the logit scale, modest score spacing represents large odds ratio changes. The scores may be assigned on external grounds or estimated from the data (see section 2.3.1 for discussion on scoring techniques). The need for assigning scores to the categories of the dependent variable represents a drawback of the stereotype model. Techniques which utilise the ordinality in a response variable, without the need to score levels, are of most interest in this research, therefore the stereotype model and its application are not examined further.

2.4.3: The Proportional Odds Model

The Proportional Odds Model was first introduced by McCullagh (1980), and thus is sometimes called the McCullagh model. It is considered an extension of the Generalised Linear Model (presented by Nelder and Wedderburn (1972) (details can also be found in McCullagh and Nelder(1989)).

For an ordinal variable, Y, with response categories k=1,2,...,K, where we assume that the scale of the variable is such that there is a most favourable or desirable response down to a least favourable or desirable response, for example, if a variable "reaction to treatment of kidney failure" has possible responses none, moderate and good, the most desirable category is obviously good, or similarly for a variable "pain after operation" with response categories none, mild and severe, the most desirable outcome is none.

The possible responses for an ordinal variable are deemed to represent a categorisation of an underlying continuous variable, i.e. the K categories form contiguous intervals on the continuous scale with cut points or divisions between categories denoted by α_1, α_2, ..., α_{K-1}.

When considering the odds of an event with an ordinal variable, it is not appropriate to simply dichotomise the categories as response 1/not 1 and response 2/not 2 etc.. The adjacency of the levels must be respected, so that when looking at the odds of a response in category 2 we do not combine level 1 with 3 to k. Therefore when looking at odds it makes more sense to divide the ordinal variable cumulatively, ie level 1/levels 2 to k, levels 1 & 2/levels 3 to k, ..., levels 1 to k-1/k or similarly, as can be seen for the continuation odds model, level 1/levels 2 to k, level 2/levels 3 to k, ..., level k-1/level k.

The proportional odds model models the log odds of a 'more desirable' response regardless of which category an individual might have responded in, ie it simultaneously models the log odds of :-

Category 1 versus Categories 2 to K

Categories 1 and 2 versus Categories 3 to K

.

.

Categories 1 to K-1 versus Category K

Therefore, the model uses a global odds ratio (a single parameter that represents the effects at the k-1 adjacent dichotomies of the response, i.e. it represents k-1 local odds ratios) to reflect the odds of an individual giving a more favourable response given their explanatory covariate characteristics.

To illustrate this concept further consider a situation where there is only one covariate coded as binary (with levels 1 and 2). If the response is binary, too, (with levels, for instance, of desirable (d) and undesirable (u)), then we have a 2×2 table. A standard measure of association for such a table is the odds ratio :-

$$\theta = p_{1d}(1-p_{2d}) / p_{2d}(1-p_{1d})$$

If the response has more than 2 categories and is ordinal then the data are contained in a 2×n table (where n is the number of response categories). Denote the response levels as C_1, C_2, ..., C_n, where C_1 is the most favourable and C_n is the least favourable, the data can be reconstituted into any of n-1 2×2 tables with response categories C_1 to C_{j-1} vs C_j to C_n (where j takes any value between 2 and n inclusive), and for each of the n-1 tables an odds ratio can be calculated as above. The proportional odds assumption is that all of these n-1 odds ratios are equivalent, ie not statistically significantly different from each other.

If we say that n=3 for simplicity's sake, the original 2×3 table would look like table 2.8:-

Table 2.8: Probability distribution for a binary covariate and an ordinal response with 3 levels

	C_1	C_2	C_3
1	p_{11}	p_{12}	p_{13}
2	p_{21}	p_{22}	p_{23}

Reconstituting the data as described previously, table 2.8 would break down into tables 2.8a and 2.8b below

Table 2.8a: As table 2.8, except levels 2 and 3 of the response are collapsed

	C_1	$C_2 + C_3$
1	p_{11}	$p_{12}+p_{13}$
2	p_{21}	$p_{22}+p_{23}$

Table 2.8a has odds ratio $\theta_{2.8a} = p_{11}(p_{22}+p_{23})/p_{21}(p_{12}+p_{13})$

Table 2.8b: As table 2.8, except levels 1 and 2 of the response are collapsed

	$C_1 + C_2$	C_3
1	$p_{11}+p_{12}$	p_{13}
2	$p_{21}+p_{22}$	p_{23}

Table 2.8b has odds ratio $\theta_{2.8b} = (p_{11}+p_{12})p_{23}/(p_{21}+p_{22})p_{13}$

Under the Proportional Odds Assumption, $\theta_{2.8a} = \theta_{2.8b}$.

The odds that an individual with specific covariates, ($\mathbf{x}_i$), responds in category j or less of Y are given by :-

$$P(Y_i \le j)/[1-P(Y_i \le j)] = \gamma_j(\mathbf{x}_i)/\{1-\gamma_j(\mathbf{x}_i)\}$$

where

$$P(Y_i \le j) = \gamma_j(\mathbf{x}_i) = \pi_1(\mathbf{x}_i) + \pi_2(\mathbf{x}_i) + ... + \pi_j(\mathbf{x}_i).$$

The proportional odds model takes the form :-

$$\log [\gamma_j(\mathbf{x}_i)/\{1-\gamma_j(\mathbf{x}_i)\}] = \alpha_j + \boldsymbol{\beta}'\mathbf{x}_i \qquad 1 \le j \le K-1 \qquad (2.22)$$

Which shows that it is the cumulative probabilities ($\gamma_j(x_i)$) that are modelled, not the individual response probabilities ($\pi_j(\mathbf{x}_i)$), using the logit transformation. The cumulative probabilities can be calculated as follows :-

Say $\eta_{(j)i} = \alpha + \boldsymbol{\beta}'\mathbf{x}_i$ and rearrange the model formula above :-

$$\gamma_j(\mathbf{x}_i) = \exp[\eta_{(j)i}]/\{1 + \exp[\eta_{(j)i}]\} \qquad (2.23)$$

As $\gamma_j(\mathbf{x}_i) = \pi_1(\mathbf{x}_i) + \pi_2(\mathbf{x}_i) + ... + \pi_j(\mathbf{x}_i)$, the individual probability of a response in category j, $\pi_j(x_i)$ can be found by rearranging the previous expression :-

$$\pi_j(\mathbf{x}_i) = \gamma_j(\mathbf{x}_i) - \gamma_{j-1}(\mathbf{x}_i). \qquad (2.24)$$

The Proportional Odds Model is in effect fitting k-1 binary logistic models simultaneously and using an averaged or global odds ratio, ie it is modelling :-

$$\log [\gamma_1(\mathbf{x}_i)/\{1-\gamma_1(\mathbf{x}_i)\}] = \alpha_1 + \boldsymbol{\beta}'\mathbf{x}_i$$

$$\log [\gamma_2(\mathbf{x}_i)/\{1-\gamma_2(\mathbf{x}_i)\}] = \alpha_2 + \boldsymbol{\beta}'\mathbf{x}_i$$

$$\log[\gamma_{k-1}(\mathbf{x}_i)/\{1-\gamma_{k-1}(\mathbf{x}_i)\}] = \alpha_{k-1} + \beta'\mathbf{x}_i$$

at the same time. In separate binary logistic models, the slope parameter, β, would not be general, but specific to each of the k-1 models, so we would have β_1, β_2, ..., β_{k-1}, and it is from these values that a mean is calculated to form the global odds ratio parameter in the proportional odds. This information is very useful in that one could actually fit these separate models in order to check the local odds ratios with the global one, ie compare magnitude and direction of the covariate parameters, doing this would be especially relevant when the proportional odds assumption fails or there is suspicion that the odds of a 'more favourable' response are not constant over the levels of the dependent variable.

To compare 2 individuals, a and b, with covariates $\mathbf{x}_a$ and $\mathbf{x}_b$, the difference in log odds of the event of a response $Y_i \leq j$ is :-

$$\log[\gamma_j(\mathbf{x}_a)/\{1-\gamma_j(\mathbf{x}_a)\}] - \log[\gamma_j(\mathbf{x}_b)/\{1-\gamma_j(\mathbf{x}_b)\}] = \beta'(\mathbf{x}_a-\mathbf{x}_b)$$

rearranging :-

$$\log \frac{[\gamma_j(\mathbf{x}_a)/\{1-\gamma_j(\mathbf{x}_a)\}]}{[\gamma_j(\mathbf{x}_b)/\{1-\gamma_j(\mathbf{x}_b)\}]} = \beta'(\mathbf{x}_a-\mathbf{x}_b), \qquad 1 \leq j \leq k-1 \qquad (2.25)$$

so the log odds ratio of the event $Yi \leq j$ is the same for all j=1, ..., k-1, that is the difference in log odds is independent of the response category involved. This is the proportional odds assumption.

2.4.3.1: The Proportional Odds Assumption

If the assumption that a global odds ratio parameter adequately represents the k-1 underlying local odds ratios is not reasonable, then the proportional odds model (2.22) given earlier would be modified to :-

$$\log [\gamma_j(x_i)/\{1-\gamma_j(x_i)\}] = \alpha_j + \boldsymbol{\beta}_j \mathbf{x}_i \qquad 1 \le j \le k-1 \qquad (2.26)$$

where $\boldsymbol{\beta}_j = (\beta_{j1}, \beta_{j2}, ..., \beta_{jm})$ when there are m covariates in the model. Thus the model is saying that covariate effects are specific to each response level, whereas for the proportional odds model $\beta_{1z} = \beta_{2z} = ... = \beta_{k\text{-}1z} = \beta_z$ for $1 \le z \le m$.

This new model has m(k-1) β and (k-1) α parameters, ie (m+1)(k-1) parameters to be estimated compared to m+k-1 parameters under the proportional odds assumption. The increased efficiency of the proportional odds model is clear to see, therefore if the concept of proportional odds is feasible, the model is more parsimonious than a model such as (2.26)

In order to test the proportional odds assumption one could apply a score statistic. The statistical package SAS carries out this test automatically within the LOGISTIC procedure, the following is an outline description of how this is done:-

First consider the multivariate response model, where the number of response levels is strictly greater than 2, with m covariates included in the model :-

$$g(\Pr(Y \le i|\mathbf{x})) = (1,\mathbf{x}')\, \gamma_i$$

for i=1, ..., k, where Y is the response. $\gamma_i = (\gamma_{i0}, \gamma_{i1}, ..., \gamma_{im})'$ is a vector of parameters to be estimated, consisting of an intercept γ_{i0} and m slope parameters. So the parameter vector for the full model here is :-

$$\gamma = (\gamma_1', ..., \gamma_k').$$

Under the proportional odds assumption these parameters are equivalent :-

$$\gamma_{1z} = \gamma_{2z} = \ ... = \gamma_{kz} \qquad \text{for all } z = 1, ..., m.$$

Let $\alpha_1^{hat}, \ldots, \alpha_k^{hat}$ and $\beta_1^{hat}, \ldots, \beta_m^{hat}$ be the maximum likelihood estimates under the proportional odds assumption. Then for all i :-

$$\gamma_i^{hat} = (\alpha_i^{hat}, \beta_1^{hat}, \ldots, \beta_m^{hat})'.$$

The chi-squared score statistic is evaluated at

$$\gamma_0 = (\gamma_1^{hat}, \ldots, \gamma_k^{hat})'$$

and has an asymptotic chi-squared distribution with m(k-1) degrees of freedom (SAS institute (1989).

Score statistics in general are only approximate and have relatively low statistical power, and thus a significance level of 10%, rather than the standard 5%, is advisable (Carroll (1993)). More can be found about score statistics in Rao (1973).

If the proportional odds assumption is accepted this does not imply that the model is a good fit to the data. What is implied is that it is appropriate to use a single odds ratio per explanatory variable, to represent the odds of events described by the k-1 possible cumulative dichotomies of the response. Ways of evaluating the adequacy of the proportional odds model are discussed in section 2.5.

If the proportional odds assumption fails, it is a good idea to fit and examine the k-1 binary logistic models that correspond to the proportional odds model, to see why and/or where the assumption is violated. Comparing the parameter estimates for the covariates, for each dichotomy of the response, for consistency of magnitude and direction can be most insightful. Rather than discarding the proportional odds, the problem may be a single covariate or sub-division of the response. Depending on the focus of the investigation or feasibility, one might remedy this problem by omitting a variable, collapsing levels of a variable, if an 'offending' covariate is categorical (if appropriate), or if an explanatory variable is ordinal, a different scoring method may improve the model. A similar option would be to collapse levels of the response variable if this were appropriate or feasible.

2.4.3.2: Parameter estimation for the proportional odds model

Techniques to estimate parameters in regression type models are fairly standard with many theoretical texts giving details, such as Collett (1991), Agresti (1984, 1990) and Hosmer and Lemeshow (1989), therefore this section will be a short overview.

In a random sample of N subjects where the dependent variable is ordinal with k>2 levels, each individual has a response j, $1 \leq j \leq k$, with probability $\pi_j(\mathbf{x}_i)$, where $\mathbf{x}_i$ is a vector of covariates. $\pi_j(\mathbf{x}_i)$ is related to α_j and β through :-

$$P(Y_i \leq j) = \gamma_j(\mathbf{x}_i) = \pi_1(\mathbf{x}_i) + \pi_2(\mathbf{x}_i) + \ldots + \pi_j(\mathbf{x}_i),$$

$$\log\left[\gamma_j(\mathbf{x}_i)/\{1-\gamma_j(\mathbf{x}_i)\}\right] = \alpha_j + \beta'\mathbf{x}_i \qquad 1 \leq j \leq k-1,$$

and

$$\pi_j(\mathbf{x}_i) = \gamma_j(\mathbf{x}_i) - \gamma_{j-1}(\mathbf{x}_i).$$

The likelihood is then proportional to a product of N probabilities :-

$$L \quad \propto \quad \prod_{i=1}^{N} \prod_{j=1}^{k} \pi_j(\mathbf{x}_i)^{Y(j)i}$$

where $Y(j)i = 1$ if the i^{th} individual responds in category j

$Y(j)i = 0$ otherwise

The parameters α_j and β are estimated by calculating partial derivatives and equating to zero. As described for the binary logistic model, in most cases these equations can only be solved numerically, using Newton-Raphson type iterative techniques.

Alternatively, as stated on page 35, least squares estimation could be used.

2.4.4: The Continuation Odds Model

The Continuation Odds model is another extension of the generalised linear model and the concept is quite similar to that of the proportional odds model. The following describes what the model does.

Let Y be an ordinal response variable with k levels ordered from 'best' to 'worst' or 'most favourable' to 'least favourable', as described for the proportional odds model. Again, it is assumed that these k categories form contiguous intervals on an underlying latent continuous scale, with the divisions between levels denoted by $\alpha_1, \alpha_2, ..., \alpha_{K-1}$.

Within the notion of the continuation odds model, the event of a response is conditional. Only the actual response category or worse levels of the dependent variable are available, so the odds considered are that an individual responds in category j, given that he/she could only have responded in categories j or worse. Therefore, the dichotomies of the response which enable those odds to be examined are level 1/levels 2 to k, level 2/levels 3 to k, ..., level k-1/level k.

Therefore, the continuation odds model models the log odds of the 'most favourable' response available, regardless of which category an individual may have responded in, i.e. it simultaneously models the log odds of :-

Category [1 / 1 to k] vs categories [2 to k / 1 to k]

Category [2 / 2 to k] vs categories [3 to k / 2 to k]

.

.

Category [k-1 / k-1 to k] vs categories [k / k-1 to k]

Therefore the model uses a global odds ratio to reflect the odds of an individual giving the most desirable response available, given their explanatory covariate characteristics.

Using the same approach as for the proportional odds model to illustrate this concept further, consider a situation where there is only one covariate coded as binary (with levels 1 and 2). If the response has more than 2 categories and is ordinal then the data is contained in a 2×n table. Denote the response levels as C_1, C_2, ..., C_n, where C_1 is the most favourable and C_n is the least favourable. The data can be reconstituted, according to the 'constraints' above, into any of n-1 2×2 tables with response categories C_{j-1} vs C_j to C_n (where j takes any value between 2 and n inclusive), and for each of the n-1 tables an odds ratio can be calculated. The continuation odds assumption is that all of these n-1 odds ratios are equivalent, i.e. not statistically significantly different from each other.

If we say that n=3 for the sake of simplicity, the original 2×3 table would look like table 2.9 (Tables 2.8 and 2.9, and 2.8a and 2.9a are essentially identical and reproduced for convenience) :-

Table 2.9: Probability distribution for a binary covariate and an ordinal response with 3 levels

	C_1	C_2	C_3
1	p_{11}	p_{12}	p_{13}
2	p_{21}	p_{22}	p_{23}

Reconstituting the data as described above, table 2.9 would break down into tables 2.9a and 2.9b below

Table 2.9a: As table 2.9, except levels 2 and 3 of the response are collapsed

	C_1	$C_2 + C_3$
1	p_{11}	$p_{12}+p_{13}$
2	p_{21}	$p_{22}+p_{23}$

Under the continuation odds concept, Table 2.9a has an odds ratio of conditional odds :-

$$\theta_{2.9a} = \{(p_{11}/[p_{11}+p_{12}+p_{13}])(p_{22}+p_{23}/[p_{21}+p_{22}+p_{23}]) \,/\, (p_{21}/[\, p_{21}+p_{22}+p_{23}])(p_{12}+p_{13}/[p_{11}+p_{12}+p_{13}])\}$$

Table 2.9b: As table 2.9, with level 1 of the response deleted

	C_2	C_3
1	p_{12}	p_{13}
2	p_{22}	p_{23}

Table 2.9b also has an odds ratio of conditional odds :-

$$\theta_{2.9b} = \{(p_{12}/[p_{12}+p_{13}])(p_{23}/[p_{22}+p_{23}]) \,/\, (p_{22}/[\, p_{22}+p_{23}])(p_{13}/[p_{12}+p_{13}])\}$$

Under the Continuation Odds Assumption, $\theta_{2.9a} = \theta_{2.9b}$.

As for previous models, suppose individual i has a response Y_i and a specific set of covariates $\mathbf{x}_i$, so then response levels have probabilities $\pi_1(\mathbf{x}_i)$, $\pi_2(\mathbf{x}_i)$, ..., $\pi_k(\mathbf{x}_i)$. Now consider a conditional probability that the individual responds in category j given that categories j or 'worse' are available to him, denoted by $H_j(\mathbf{x}_i)$:-

$$H_j(\mathbf{x}_i) = P(\text{individual responds in } j / \text{ levels } j \text{ to } k \text{ available})$$

i.e. :-

$$H_j(\mathbf{x}_i) = P(Y_i = j \,/\, \text{levels} \geq j \text{ available}) \qquad (2.27)$$

Therefore :-

$$H_j(\mathbf{x}_i) = \pi_j(\mathbf{x}_i)/\{\pi_j(\mathbf{x}_i) + \pi_{j+1}(\mathbf{x}_i) + \ldots + \pi_k(\mathbf{x}_i)\} \qquad (2.28)$$

$$H_j(\mathbf{x}_i) = \pi_j(\mathbf{x}_i)/\{\sum_{z\geq j}^{k} \pi_z(\mathbf{x}_i)\} \qquad 1 \leq j \leq K\text{-}1 \qquad (2.29)$$

Note: $H_k(\mathbf{x}_i) \equiv 1$ and $H_1(\mathbf{x}_i) \equiv \pi_1(\mathbf{x}_i)$.

The Continuation Odds Model takes the form :-

$$\log\{H_j(\mathbf{x}_i)/[1-H_j(\mathbf{x}_i)]\} = \alpha_j + \beta'\mathbf{x}_i \qquad 1 \leq j \leq K\text{-}1 \qquad (2.30)$$

At the top of this section, it is mentioned that the concept of the continuation odds model is similar to that of the proportional odds model. It can be seen that the structure of the two models are similar also, as the continuation odds model does not model the response probabilities, $\pi_j(\mathbf{x}_i)$, but instead the conditional probabilities, $H_j(\mathbf{x}_i)$, via the logit transformation. The model takes its name from the continuation ratio introduced by Fienberg (1979), the continuation ratio for response category j is similar to $H_j(\mathbf{x}_i)$, and is given by :-

$$R_s(\mathbf{x}_i) = \pi_j(\mathbf{x}_i) \,/\, [\pi_{j+1}(\mathbf{x}_i) + \ldots + \pi_k(\mathbf{x}_i)]$$

Whereas :-

$$\{H_j(\mathbf{x}_i) / [1-H_j(\mathbf{x}_i)]\} = \frac{\pi_j(\mathbf{x}_i) / [\pi_j(\mathbf{x}_i) + \ldots + \pi_k(\mathbf{x}_i)]}{[\pi_{j+1}(\mathbf{x}_i) + \ldots + \pi_k(\mathbf{x}_i)] / [\pi_j(\mathbf{x}_i) + \ldots + \pi_k(\mathbf{x}_i)]}$$

therefore :-

$$\{H_j(\mathbf{x}_i) / [1-H_j(\mathbf{x}_i)]\} \equiv R_s(\mathbf{x}_i)$$

Which means that modelling the logit of $H_j(\mathbf{x}_i)$ is equivalent to modelling the log of the continuation ratio $R_s(\mathbf{x}_i)$.

The conditional probabilities $H_j(\mathbf{x}_i)$ can be calculated by writing $\eta_{(j)i} = \alpha_s + \beta'\mathbf{x}_i$ and rearranging expression 2.30 :-

$$H_j(\mathbf{x}_i) = \exp[\eta_{(j)i}]/\{1 + \exp[\eta_{(j)i}]\} \tag{2.31}$$

in the same way cumulative probabilities are obtained in the proportional odds model. Once $H_j(\mathbf{x}_i)$ for $1 \leq j \leq k$ have been calculated, individual response level probabilities can be found :-

First :-

$$H_1(\mathbf{x}_i) = \pi_1(\mathbf{x}_i)$$

$$H_2(\mathbf{x}_i) = \pi_2(\mathbf{x}_i)/\{\pi_2(\mathbf{x}_i) + \pi_3(\mathbf{x}_i) + \ldots + \pi_k(\mathbf{x}_i)\}$$

.

.

$$H_{k-1}(\mathbf{x}_i) = \pi_{k-1}(\mathbf{x}_i)/\{\pi_{k-1}(\mathbf{x}_i) + \pi_k(\mathbf{x}_i)\}$$

$$H_k(\mathbf{x}_i) = 1$$

Rearranging these gives :-

$$\pi_1(\mathbf{x}_i) = H_1(\mathbf{x}_i)$$

$$\pi_2(\mathbf{x}_i) = [1-H_1(\mathbf{x}_i)]H_2(\mathbf{x}_i)$$

.

.

$$\pi_{k-1}(\mathbf{x}_i) = [1-H_1(\mathbf{x}_i)]...[1-H_{k-2}(\mathbf{x}_i)]H_{k-1}(\mathbf{x}_i)$$

$$\pi_k(\mathbf{x}_i) = [1-H_1(\mathbf{x}_i)]...[1-H_{k-1}(\mathbf{x}_i)]H_k(\mathbf{x}_i)$$

Therefore :-

$$\pi_1(\mathbf{x}_i) = H_1(\mathbf{x}_i)$$

$$\pi_j(\mathbf{x}_i) = \prod_{z=1}^{j-1}[1-H_z(\mathbf{x}_i)]\ H_j(\mathbf{x}_i) \qquad 2 \le j \le k. \tag{2.32}$$

The continuation odds model is in effect simultaneously fitting k-1 'sub-models', and producing a final model combining these 'sub-models' using an averaged global odds ratio for the covariates :-

$$\log\{H_1(\mathbf{x}_i)/[1-H_1(\mathbf{x}_i)]\} = \alpha_1 + \boldsymbol{\beta}'\mathbf{x}_i$$

$$\log\{H_2(\mathbf{x}_i)/[1-H_2(\mathbf{x}_i)]\} = \alpha_2 + \boldsymbol{\beta}'\mathbf{x}_i$$

.

.

$$\log\{H_{k-1}(\mathbf{x}_i)/[1-H_{k-1}(\mathbf{x}_i)]\} = \alpha_{k-1} + \boldsymbol{\beta}'\mathbf{x}_i$$

This illustrates that the continuation odds model is fitting a sequence of binary logistic models to the sub-divisions of the dependent variable described earlier that correspond to membership of the most favourable response category available. The model gives an 'average' log odds ratio of the most 'favourable' or 'better' response over the categories.

2.4.4.1: The continuation odds model and survival analysis

A documented way of fitting the continuation odds model using standard packages is given by Iyer (1985) and Whitehead (1991). The ordered response is manipulated according to the structure of the model, and with pseudo parameters forces the binary logistic model to take the form of the continuation odds model.

There is also a relationship between Cox's proportional hazards model and the continuation odds model, mentioned by McCullagh (1980), Iyer (1985) and Whitehead (1991). This section looks at the relationship in order to exploit it and offer another way to fit the continuation odds model using standard packages. Both methods of fitting the model are discussed in section 2.4.4.3.

This section reviews some simple features of survival analysis, to aid understanding of the connection between the continuation odds and proportional hazards models. Full details can be found in Kalbfleisch and Prentice (1980).

Let T be variable with possible lifetimes t_1, ..., t_k. Then, using standard probability theory :-

$$f(t_j) = P(T = t_j) = \pi_j$$

$$F(t_j) = P(T \leq t_j) = \sum_{z=1}^{j} \pi_z$$

$$S(t_j) = 1 - F(t_j) = \sum_{z=j+1}^{k} \pi_z$$

Where $S(t_j)$ is known as a survivor function.

A hazard function, $\lambda(t_j)$, is the probability of failure at time t_j given that the subject has survived up to time t_j :-

$$\lambda(t_j) = P(T = t_j / T > t_{j-1})$$

$$\lambda(t_j) = f(t_j) / S(t_{j-1})$$

$$\lambda(t_j) = \pi_j / \sum_{z=j}^{k} \pi_z$$

Also, the hazard and survivor functions can be related by noting that $f(t_j) = F(t_j)-F(t_{j-1}) = S(t_{j-1})-S(t_j)$, therefore :-

$$\lambda(t_j) = [S(t_{j-1}) - S(t_j)] / S(t_{j-1})$$

so :-

$$S(t_j) = [1- \lambda(t_j)]\, S(t_{j-1})$$

It is logical that failure will occur at some time $\geq t_1$, so $S(t_0) \equiv 1$. The above expression then becomes:-

$$S(t_j) = \prod_{z=1}^{j} [1 - \lambda(t_z)]$$

Therefore, given the survivor function, failure probabilities can be calculated :-

$$\pi_j = S(t_{j-1}) - S(t_j) = \prod_{z=1}^{j-1} [1 - \lambda(t_z)]\, \lambda(t_j)$$

Comparing the theory above with the description of the continuation odds model, similarities can be seen. The conditional probability $H_j(\mathbf{x}_i)$ is of the same form as the hazard function $\lambda(t_j)$. Also, the relationship between the response category probabilities and $H_j(\mathbf{x}_i)$ is of the same form as the relationship between failure probabilities and the hazard function.

The following section describes the relationship between the continuation odds model and Cox's proportional hazards model.

2.4.4.2: The Continuation Odds model and Cox's Proportional hazards model

Details of Cox's proportional hazards model (Cox (1972)) can be found in Kalbfleisch and Prentice (1980).

For the discrete proportional hazards model, again let T be a variable with k possible lifetimes, t_1 to t_k. Suppose an individual i with response T_i has specific covariates $\mathbf{x}_i$, so that the lifetimes have probabilities $\pi_1(\mathbf{x}_i)$, $\pi_2(\mathbf{x}_i)$, ..., $\pi_k(\mathbf{x}_i)$. The hazard for this individual at time t_j is then given by $\lambda(t_j;\mathbf{x}_i)$. The proportional hazards model takes the form :-

$$\log \{\lambda(t_j;\mathbf{x}_i) / [1-\lambda(t_j;\mathbf{x}_i)]\} = \log \{\lambda(t_j;\mathbf{0}) / [1-\lambda(t_j;\mathbf{0})]\} + \boldsymbol{\beta}'\mathbf{x}_i$$

where $1 \leq j \leq k-1$.

The proportionality in the model refers to the effect of covariates remaining constant irrespective of the time a failure occurs. This is similar to the feature of the proportional odds model where the effect of covariates is constant regardless of response category involved.

The term $\log \{\lambda(t_j;\mathbf{0}) / [1-\lambda(t_j;\mathbf{0})]\}$ is the logit of the baseline hazard where $x_i \equiv 0$, therefore it is a function of time t_j alone. This is equivalent to an intercept term, and so if the term is denoted by α_j, Cox's proportional hazards model becomes equivalent in form to the continuation odds model.

If ordinal response levels are considered as 'time' intervals, $H_j(\mathbf{x}_i)$ can be interpreted as a discrete hazard function, but instead with categories ordered 'best' to 'worst' of an outcome. Therefore, the continuation odds model can be considered as a discrete version of Cox's proportional hazards model.

It can be seen that the continuation odds model is simultaneously modelling the log odds of events [Yi=j / levels $\geq$ j available], $1 \leq j \leq k-1$, for an individual with covariates $\mathbf{x}_i$. When $\mathbf{x}_i \equiv 0$, the α_j terms represent the baseline odds of these events, therefore the β terms gauge the effect of a non-zero vector of explanatory variables the baseline odds.

In order to compare 2 individuals a and b with covariates $\mathbf{x}_a$ and $\mathbf{x}_b$, the difference in log odds can be found :-

$$\log \{H_j(\mathbf{x}_a)/[1-H_j(\mathbf{x}_b)]\} - \log \{H_j(\mathbf{x}_b)/[1-H_j(\mathbf{x}_b)]\} = \beta'(\mathbf{x}_a - \mathbf{x}_b)$$

Thus :-

$$\log \frac{\{H_j(\mathbf{x}_a)/[1-H_j(\mathbf{x}_a)]\}}{\{H_j(\mathbf{x}_b)/[1-H_j(\mathbf{x}_b)]\}} = \beta'(\mathbf{x}_a - \mathbf{x}_b), \quad 1 \leq j \leq k-1$$

$\beta'(\mathbf{x}_a - \mathbf{x}_b)$ can equivalently be interpreted as the relative odds of a more favourable response for an individual with covariates $\mathbf{x}_a$ compared to an individual with covariates $\mathbf{x}_b$. The difference in log odds, i.e. the odds ratio, is independent of response category. The α_j term cancels to leave the odds ratio to be constant over the levels, which is referred to as the continuation odds assumption. In Cox's model, this is the same as the proportional hazards assumption.

2.4.4.3: Fitting the continuation odds model

No standard packages cater specifically for the continuation odds model. The following section outlines two methods of fitting the continuation odds model using the statistical package SAS.

Using the LOGISTIC procedure

PROC LOGISTIC in SAS can fit the continuation odds model. By manipulation of the data and the ordinal response, the procedure simultaneously fits the k-1 binary logistic sub-models described earlier. Berridge and Whitehead (1991) give a full account of this method.

Let the ordinal response have k levels, and that a dataset has a single row of information for each subject containing the observed response and covariate values. There are k-1 steps in the data manipulation :-

(1) For all individuals, 2 new variables are created - CUTPT and IND. Now, CUTPT=1 for all subjects, IND=0 if an individual's response is in category 1 and IND=1 otherwise (categories 2 to k). This first step sets up the first binary split of the response using CUTPT, and IND allows the comparison between [level 1/1 to k available] vs [levels 2 to k/1 to k available] using the binary logistic model.

(2) For those individuals responding with level 2 or 'worse', a second row of data is generated, with original information unchanged. Now CUTPT=2, and IND=0 if the subject's response is in category 2, IND=1 otherwise. This step relates to the binary logistic model allowing the comparison between [level 2/2 to k available] vs [levels 3 to k/2 to k available].

.

.

.

(k-1) For those responding in category k-1 or worse (k), a k-1th row of data is created, now assigning CUTPT=k-1. If an individual's response is in category k-1, then IND=0, otherwise IND=1. This k-1th step relates to the binary logistic model allowing comparison between [level k-1/k-1 to k available] vs [level k/k-1 to k available].

To illustrate the process, consider the following made up small subset of the South Yorkshire Police data described in Chapter 3 :-

obs	Communication	Job Satisfaction	Morale
1	good	yes	v high
2	neither	yes	high
3	good	yes	neither
4	neither	no	low
5	bad	no	v low

For this example k=5 so there are 4 steps in the data manipulation process. The data is amended to :-

obs	Communication	Job Satisfaction	Morale	CUTPT	IND
1	good	yes	v high	1	0
2	neither	yes	high	1	1
3	good	yes	neither	1	1
4	neither	no	low	1	1
5	bad	no	v low	1	1
2	neither	yes	high	2	0
3	good	yes	neither	2	1
4	neither	no	low	2	1
5	bad	no	v low	2	1
3	good	yes	neither	3	0
4	neither	no	low	3	1
5	bad	no	v low	3	1
4	neither	no	low	4	0
5	bad	no	v low	4	1

The continuation odds model is fitted to this data by using the binary logistic model, via PROC LOGISTIC, using IND as a binary response variable.

All models must include a term for CUTPT to ensure the same set of explanatory covariate parameters β to be estimated for each continuation odds sub-model.

To fit the null continuation odds model in PROC LOGISTIC, ie without explanatory covariates, the binary response IND is modelled using CUTPT only, and the baseline odds, the α_j terms, can be estimated as :-

$$\alpha_j = \text{Intercept} + j(\text{CUTPT})$$

The disadvantages associated with the above approach include the fact that the data manipulation can be cumbersome, and that the standard errors of the intercepts are not easily obtained, as the α_j terms are not estimated directly.

Using the PHREG procedure

Earlier it was shown that the continuation odds model is equivalent to a discrete Cox's proportional hazards model, therefore the SAS survival analysis procedure PROC PHREG can be used.

PHREG avoids the need for data manipulation. Response categories are labelled 1 through to k to indicate the relative ordering. As the continuation odds and proportional hazards models are of the same form, one only needs to specify the correct SAS code to fit the continuation odds model :-

```
PROC PHREG DATA=dataset;
MODEL response = x1 x2, ..., xm / TIES = DISCRETE;
BASELINE COVARIATES = covariate dataset;
OUTPUT OUT = dataset for survivor function  SURVIVAL= label for survivor
function;
```

The 'TIES =' option indicates that ties in the response are not by chance but due to the discrete nature of the data and invokes the discrete logistic model for the proportional hazards.

The BASELINE option calculates the survivor function for a set of covariate values defined in a COVARIATES= dataset which must be constructed prior to the analysis. Details of the PHREG procedure can be found in 'Extended Help' within the package, or an up to date SAS/STAT user guide.

The parameter estimates can be used to estimate odds ratios as described earlier. Estimated probabilities, $p_j(\mathbf{x}_i)$, for a specific set of covariates, $\mathbf{x}_i$, can be calculated from the estimated survivor functions (specify covariate values in the COVARIATES = dataset) by applying :-

$$p_j(\mathbf{x}_i) = S_{j-1}(\mathbf{x}_i) - S_j(\mathbf{x}_i) \text{ where } S_0(\mathbf{x}_i) \equiv 1.$$

The PHREG procedure does not estimate intercept terms. In survival analysis the intercept represents the baseline hazard function, not estimated for proportional hazards model. However, this does not represent a problem as the intercepts need only be estimated to calculate category response probabilities, but as these can be obtained vie the estimated survivor function, there is no real need to estimate them using this procedure.

The PHREG procedure is not employed in this investigation, due to a sample size constraint currently within the procedure.

2.4.4.4: The Continuation Odds Assumption

The continuation odds assumption is vital to the success and usefulness of the continuation odds model. Theoretically, the assumption can be tested as described for the proportional odds model using a score statistic test. However, neither of the methods of fitting the model using SAS test the assumption automatically, and this inability to easily test the continuation odds assumption represents a drawback to the application of the model. The construction of a score test for the continuation odds assumption is very complex, therefore other methods of examining the appropriateness of the model are explored.

When fitting the continuation odds using PROC LOGISTIC when there is a single independent covariate, Iyer (1985) and Whitehead (1991) note that the assumption can be tested by including a CUTPT by explanatory covariate interaction term. If these are shown to contribute significantly to the description of the binary response (IND) then the assumption is deemed to have failed. In the case where there's more than one covariate, ie nearly all cases in practice, there is no explanation of how to test the assumption.

The most accessible method by which to examine whether the assumption is acceptable, is to reconstitute the data, as described above in section 2.4.4, and fit the k-1 sub-models mentioned separately using the binary logistic model. The assumption is examined by comparing parameter estimates (log odds ratios) for corresponding covariates, and checking for consistency of direction and magnitude, similarly to the approach adopted when the proportional odds assumption fails. An extremely helpful preliminary procedure is to examine the separate binary logistic models, for two proportional odds models for which the proportional odds assumptions are satisfied and violated respectively. This can give insight into the conditions under which such an assumption fails and succeeds, which can be applied to the examination of the continuation odds assumption.

2.4.4.5: Parameter Estimation for the continuation odds model

The parameter estimation methods for the continuation odds model are the same as for the proportional odds and binary logistic models, i.e. maximum likelihood or iteratively reweighted least squares. Given the $\pi_j(\mathbf{x}_i)$, the likelihood for a sample of N subjects is as for the proportional odds model.

The procedures PHREG and LOGISTIC in SAS actually use different estimation techniques. PHREG uses maximum likelihood, while LOGISTIC uses iteratively reweighted least squares. The two approaches will not yield exactly the same estimates, although they should be close due to the asymptotic equivalence of the techniques.

2.5: Criteria for assessing fit

Methods to assess the adequacy of the models described above are relatively limited. Rough indicators of the fit of the models can be found easily within standard packages, along with statistics testing the significance of the parameters, but tests to show how well the models fit the data are not widely prevalent.

As discussed in previous sections, the assumptions of proportional and continuation odds can be examined to see whether the type of model is appropriate or feasible.

2.5.1: Indicators of model adequacy

The raw deviance or scaled deviance, D, of a model is defined as minus twice the log likelihood evaluated at the parameter estimates obtained from the data :-

$$D = -2 \log (L)$$

Deviance assesses the lack of fit of the model, so that the poorer fit a model gives the higher the deviance (Agresti (1990)). On its own as an absolute measure, deviance is not very informative, however, it is a useful tool with which to compare two models for the same data (McCullagh and Nelder (1989)).

Within Proc LOGISTIC in SAS, the raw or scaled deviance for a model with intercept only, and the deviance for a model with intercept plus covariates is given. From these measures, one can assess whether the addition of covariate parameters offers a significant improvement to the fit of the model. If m_1 is the model with intercept term(s) only, with d_1 number of parameters, and m_2 is the model with intercept plus covariate terms, with d_2 number of parameters, then to assess whether the extra parameters are worthwhile, the difference in deviances approximately follows a χ^2 distribution with d_2-d_1 df :-

$$Dm_1 - Dm_2 \approx \chi^2_{(d2-d1)},$$

in SAS the procedure also gives a p-value so one can instantly see whether the covariate parameters as a whole are significant.

The same information can be used to compare two or more models for the same data. After deciding the covariate parameters are significant in a model, one can compare the deviances for two different models. If m_3 is a model for the same data as above, and has the same form as model m_2 except for the addition of one extra covariate, so the number of parameters is $d_3 = d_2+1$, we can see if this extra parameter provides a significant improvement in fit over model m_2 by comparing the deviances :-

$$Dm_3 - Dm_2 \approx \chi^2_{(d3-d2)}$$

so if $Dm_3 - Dm_2$ is significant on 1 degree of freedom, then the model m_3 with the extra parameter provides a better fit to the data than model m_2. This technique of comparing models for the same data by their deviances is called analysis of deviance (ANODEV). Further examples and instruction of this technique are given in McCullagh (1980), Hastie et al (1989) and Agresti (1990).

For each of the parameter estimates in a model, a Wald Chi-square statistic can be calculated to test whether the parameter is significant, i.e. statistically different from zero. These statistics with p-values are produced automatically in most packages. If a parameter is not significant, it should be removed from the model, in the case where more than one parameter is not significant, then one would remove the parameter with the highest p-value first, and then refit the model to see if this alters the significance of other parameters. In the case where one is using dummy variables, the significance of each parameter estimate is not paramount. If only one dummy variable parameter from a set representing a single covariate is significant, all dummy variables pertaining to that covariate should be kept in the model (Wald (1943)).

Proc LOGISTIC offers measures of association between observed responses and predicted probabilities, which assess the predictive ability of the model. The procedure produces percentage figures for the number of concordant and discordant pairs (see section 2.2.6). In Proc LOGISTIC, if an event response is defined as the response whose ordered value is 1, then a pair of observations, with different responses, is said to be concordant if the larger response has a lower predicted event probability than the smaller response. Similarly the pair is discordant if the larger response has a higher probability. If the pair is neither concordant or discordant, it is a tie. The predicted probabilities are categorised into intervals of length 0.002 in order to allow for enumeration of concordant and discordant pairs (SAS institute (1989)).

N is the total number of observations, and there are a total of t pairs with different responses, nc of them are concordant, nd are discordant, and t-nc-nd are tied. From these values, some indices of rank correlation are calculated :-

Somer's D = (nc-nd)/t

Goodman-Kruskal Gamma = (nc-nd)/(nc+nd)

Kendall's Tau-a = (nc-nd)/(0.5N(N-1))

Somers' D, presented by Somers (1962), and Gamma, given by Goodman and Kruskal (1954), are similar measures, almost identical when the number of tied pairs is very small. Somers' D measures the difference in the proportion of concordant and discordant pairs from all pairs with different responses. Gamma is the more commonly used statistic of those above, and measures the difference between proportions of concordant and discordant pairs, from all pairs that are untied. The higher the values of these measures, the better indication for the adequacy of the model. Their interpretation should take into account the number of tied pairs, as for a large proportion of tied values, Somer's D is likely to be smaller than Gamma, and the value of Gamma could possibly exaggerate the predictive ability of the model when there are large numbers of tied pairs.

Kendall's Tau-a (Kendall (1938)) is the difference between the proportion of concordant and discordant pairs out of all pairs of observations, again, it is desirable for this measure to be high in value.

The actual proportion of concordant pairs itself is an indicator of the behaviour of the model, and we wish this to be as high as possible if the model is to provide a good fit to the data.

The usefulness of the above measures as regression model diagnostics is questionable. They are indicators, but one could not base conclusions about a model of these statistics. The measures are based on ranking procedures rather than absolute numbers, therefore the model may estimate well in the desired direction , but not actually fit the data well. The need for a measure of fit for ordinal models is quite strong, and in a recent paper, an alternative way to assess ordinal models is documented, the following section discusses this technique.

2.5.2: Goodness-of-fit statistics introduced by Lipstiz et al (1996)

A paper by Lipsitz, Fitzmaurice and Molenberghs (1996) details a global goodness of fit statistic for ordinal regression models. The concept of the test is based on a statistic proposed for binary responses by Hosmer and Lemeshow (1980, 1989), and is considered an extension of Hosmer and Lemeshow's method for ordinal responses. The method is described below.

The first step is to assign a score s_k to each response category k. Methods of scoring are mentioned in section 2.3.1, and discussed by Agresti (1984, 1990), Koch et al (1977) and Thomas and Kiwanga (1993). The observed score for an individual i is :-

$$Z_i = \sum_{k=1}^{K} s_k Y_{ik}$$

The model is then fitted to obtain the predicted probabilities of individual i responding in each response category, 1 to k, p_{ik}. With these probabilities one can construct a predicted mean score for each individual, given by :-

$$\mu_i = \sum_{k=1}^{K} s_k p_{ik}$$

To form the goodness of fit statistic, the data is sorted in ascending order of predicted mean score, μ_i, and then partitioned based on percentiles of these. It is suggested that 10 groups of approximately equal size be formed, with the first group containing the n/10 (n is the total sample size) subjects with the lowest predicted mean scores, and the last group containing the n/10 individuals with the highest predicted mean scores. In general one can form G groups, with the gth group containing $n_g = n/G$ subjects.

Given the partition of the data, the goodness of fit statistic is formulated by defining the G-1 group indicators :-

$$\begin{aligned} I_{ig} &= 1 \text{ if } \mu_i \text{ is in group g} \\ &= 0 \text{ otherwise,} \end{aligned} \qquad (2.32)$$

for g = 1, ..., G-1. Then to assess the goodness of fit of model, say, (2.22) :-

$$\log\left[\gamma_j(\mathbf{x}_i)/\{1-\gamma_j(\mathbf{x}_i)\}\right] = \alpha_j + \boldsymbol{\beta}'\mathbf{x}_i \qquad 1 \le j \le K-1$$

we consider the alternative model :-

$$\log\left[\gamma_j(\mathbf{x}_i)/\{1-\gamma_j(\mathbf{x}_i)\}\right] = \alpha_j + \boldsymbol{\beta}'\mathbf{x}_i + \sum_{g=1}^{G-1} I_{ig}\gamma_g \qquad (2.33)$$

If 2.22 is correctly specified, then the extra parameters, $\gamma_1, \ldots, \gamma_{G-1}$, will all equal zero. To test the hypothesis H_0: $\gamma_1 = \ldots = \gamma_{G-1} = 0$ one can use a Wald statistic, given as standard in many computer packages, with a p-value indicating the significance of each parameter.

To further examine the fit of the model to the data, Lipstitz et al detail a method to calculate the difference between observed and expected counts for each response level, k, within each of the g groups :-

$$(O_{gk} - E_{gk})$$

for g = 1, ..., G and k = 1, ..., K, where :-

$$O_{gk} = \sum_{i=1}^{n} I_{ig} Y_{ik} \tag{2.34}$$

and :-

$$E_{gk} = \sum_{i=1}^{n} I_{ig} p_{ik} \tag{2.35}$$

Therefore enabling residuals for the groups and response levels to be computed.

General sample size guidelines suggested for discrete data should be applied (see Freeman (1987), i.e. all estimated expected counts E_{gk} should be greater than 1, and at least 80% should be greater than 5. If these guidelines do not hold, the χ^2 approximation may be poor. A possible solution to the situation where the above conditions are not met is to use fewer groups to partition the data, so that the proportion of estimated expected counts greater than 5 is 80%. For instance, if initially we use G=10 groups, but this causes more than 20% of the E_{gk} values to be less than 5, we could try G=9 groups and so on. The general feeling is that G=6 would probably be the minimum number of groups (Hosmer and Lemeshow (1989)). A test statistic calculated from fewer than 6 groups will usually have very low power (Freeman (1987)), and thus indicate a model fits well when it perhaps does not. For a given set of data, the average number of observations in group g giving response k is n/GK. Therefore, to try and ensure most $E_{gk} > 5$ we would choose G so that $n/GK > 5$ or $G < n/5K$. Thus, as a general rule, we should choose G so that $6 \leq G < n/5K$.

The standardised residual for response level k in group g is defined as :-

$$R_{gk} = \frac{O_{gk} - E_{gk}}{\sqrt{var(O_{gk}-E_{gk})}} \qquad (2.36)$$

The asymptotic variance $var(O_{gk} - E_{gk})$ can be obtained by applying a Taylor series approximation to the difference between observed and expected counts :-

$$\mathbf{O}_g - \mathbf{E}_g = (O_{g1}, \ldots, O_{g(K-1)})' - (E_{g1}, \ldots, E_{g(K-1)})' = \sum_{i=1}^{n} I_{ig}\{\mathbf{y}_i - \mathbf{p}_i(\beta)\}$$

where the last element of both $\mathbf{Y}_i$ and $\mathbf{p}_i$ has been dropped, i.e. $\mathbf{Y}_i = (Y_{i1}, \ldots, Y_{i,K-1})'$ and similarly $\mathbf{p}_i = E(\mathbf{Y}_i) = (p_{i1}, \ldots, p_{i,K-1})'$, as the covariance matrix of $(\mathbf{O}_g - \mathbf{E}_g)$ is less than full rank if the last element is not deleted.

A Taylor expansion gives :-

$$var(\mathbf{O}_g - \mathbf{E}_g) = \sum_{i=1}^{n} A_i V_i A_i' \qquad (2.37)$$

where :-

$$A_i = I_{ig}\mathbf{I} - (\sum_{j=1}^{n} I_{jg}D_j)(\sum_{j=1}^{n} D_j'V_j^{-1}D_j)D_i'V_i^{-1},$$

I is a $(K-1)\times(K-1)$ identity matrix,

$$V_i = var(\mathbf{Y}_i) = diag(\mathbf{p}_i) - \mathbf{p}_i\mathbf{p}_i'$$

and

$$D_i = \frac{\partial \mathbf{p}_i'}{\partial(\alpha', \beta_1')'}.$$

Then, var $(O_{gk} - E_{gk})$ is the jth diagonal element in expression (2.37). This only gives the estimate of variance when $k<K$, as $\mathbf{O}_g - \mathbf{E}_g$ contains only the first K-1 differences. Using :-

$$O_{gk} - E_{gk} = (n_g - \sum_{k=1}^{K-1} O_{gk}) - (n_g - \sum_{k=1}^{K-1} E_{gk})$$

$$= \sum_{k=1}^{K-1} (O_{gk} - E_{gk})$$

$$= \mathbf{1}'(\mathbf{O}_g - \mathbf{E}_g),$$

where **1** is a (K-1)×1 vector of 1s, then :-

$$\mathrm{var}(O_{gk} - E_{gk}) = \mathbf{1}'\mathrm{var}(\mathbf{O}_g - \mathbf{E}_g)\mathbf{1}.$$

As seen above, an estimate of $\mathrm{var}(O_{gk}-E_{gk})$ is fairly complicated to compute, therefore the paper describes a simple approximation :-

$$\mathrm{var}(O_{gk}-E_{gk}) = \mathrm{var}\left\{\sum_{i=1}^{n} I_{ig}(Y_{ik} - p_{ik}\right\} \qquad (2.38)$$

with an estimate of :-

$$\mathrm{var}\left\{\sum_{i=1}^{n} I_{ig}(Y_{ik} - p_{ik}\right\} = \sum_{i=1}^{n} I_{ig}p_{ik}(1-p_{ik}) \approx n_g \bar{p}_{gk}(1-\bar{p}_{gk}) \qquad (2.39)$$

where $\bar{p}_{gk} = (\sum I_{ig}p_{ik}/n_g)$ i.e. the mean predicted probability for response k in group g.

Then, we could use an approximate residual :-

$$\frac{O_{gk} - E_{gk}}{\sqrt{\{n_g \bar{E}_{gk}(1-\bar{E}_{gk})\}}} \qquad (2.40)$$

where $\bar{E}_{gk} = E_{gk}/n_g$.

This approximation is motivated by the fact that the predicted probabilities $\bar{p}_{ik}$ should be fairly similar in each group and thus $\bar{p}_{ik} \approx \bar{p}_{gk}$. The term $n_g \bar{E}_{gk}(1-\bar{E}_{gk})$ tends to overestimate the variance of $O_{gk} - E_{gk}$, therefore Lipsitz et al suggest an approximate residual which more closely approximates a Normal distribution with mean 0 and variance 1 (N(0,1)) :-

$$R_{gk}^* = \frac{O_{gk} - E_{gk}}{\sigma\sqrt{\{ n_g \bar{E}_{gk}(1-\bar{E}_{gk})\}}} \qquad (2.41)$$

where :-

$$\sigma = \sqrt{(\sum_{g=1}^{G} \sum_{k=1}^{K}(R_{gk}^2/GK))}.$$

σ can be thought of as an estimate of the 'common' standard deviation of expression (2.40) over all groups and response levels. As a general rule, if more than 5% of the $\{|R_{gk}|\}$ (or $\{|R_{gk}^*|\}$) are greater than 2, this may indicate poor fit by the model. The profile of individuals within the regions where the fit is poor should be examined in terms of covariates, response and predicted response for insight.

The sums of squares of the approximate residuals in expression (2.40) would give Pearson's χ^2 for the observed, O_{gk}, and expected, E_{gk}, counts. For a binary response, i.e. K=2, rather than an ordinal response, Pearson's χ^2 would identically be Hosmer and Lemeshow's (1980) goodness of fit statistic, which is approximately χ^2 with G-2 df when the given model fits.

To summarise the goodness-of-fit test by Lipsitz et al (1996), fitting the model with the group indicator dummy variables gives us a criterion with which to deem a model adequate, which is fairly simple to apply and interpret. After this measure has been interpreted one can further scrutinize the performance of the model, by producing residuals for the observed and expected counts within each group, within each response level. These residuals can help to identify where the model fits well and badly. The groups pertain to partitions of the data in ascending order of expected mean score, therefore if the model fits the data well in the first group, but progressive poorly in subsequent groups, one might conclude that the predictive ability of the model decreases as the expected response (value, i.e. score assigned to levels) increases, and so on. Therefore the goodness-of-fit test proposed by Lipsitz, Fitzmaurice and Molenberghs indicates the adequacy of fit of the model, and can also help highlight where, if at all, a model is deficient, so that something may be done about it, or the information considered, when interpreting results or making substantive conclusions.

2.5.3: Modification to guidelines for using Lipsitz et al (1996) goodness-of-fit statistics

The method for testing goodness-of-fit introduced by Lipstiz et al (described in the previous section), involves the ordering of the data based on predicted mean scores for the ordinal response. The data is then partitioned into approximately equal sized groups, to form indicator variables with which the model's goodness-of-fit is tested. This section discusses the effect of ordering the data, and/or partitioning it in different ways, especially when using discrete or categorical explanatory variables, where tied predicted mean scores may be likely. Specific guidelines for the application of the Lipsitz et al (1996) goodness-of-fit test are suggested.

When using the Lipsitz method for testing a model that is using only categorical or discrete explanatory variables, a model may produce many tied values for predicted mean scores, especially if the number of variables or number of categories per variable is small. Consequently, when partitioning the data, if the partitions are made purely on sample size considerations, observations with identical predicted mean scores may be partitioned into different groups. Depending on how many tied predicted mean scores are separated, this could cause the results to become dependent on the ordering of the data. This is because although the observations with tied predicted mean scores obviously have the same characteristics, they do not necessarily have the same response, and thus the parameters for the group indicator variables will fit differently when the groups' constituents are different. However, if the partitioning is made according to values of predicted mean scores closest to the desired partition sample size, without separating tied observations, the order of these tied observations is not important, as they will all be contained in the same partitioned group.

To illustrate this, subsequent sections are cross-referenced. A proportional odds model is fitted to a subset of the South Yorkshire Police data, containing only civil staff. The model is equivalent to model 4.4c (section 4.2.1) with a term for an interaction between the variables commsen and promearn (see chapter 3 for definition of variables), referred to as model 4.41c (see Appendix 4a) given by :-

$$\log [\gamma_j(\mathbf{x}_i)/\{1-\gamma_j(\mathbf{x}_i)\}] = \alpha_j + \beta_{job}(jobnew) + \beta_{c1}(com1) + \beta_{c2}(com2) + \beta_{pu1}(pub1) + \beta_{pu2}(pub2) + \beta_{pr1}(prom1) + \beta_{pr2}(prom2) + \beta_l(lendum) + \beta_{pr1c1}(prmcom1) + \beta_{pr2c1}(prmcom2) + \beta_{pr1c2}(prmcom3) + \beta_{pr2c2}(prmcom4) \quad (4.41c)$$

for j=1,2. Where $\gamma_j(\mathbf{x}_i)/\{1-\gamma_j(\mathbf{x}_i)\} = P(Y_i \le j)/[1-P(Y_i \le j)]$

Note that the specification of the model is not crucial to the understanding of the modifications to the guidelines for using the goodness-of-fit test by Lipsitz et al.

The proportional odds assumption is satisfied for model 4.41c and all parameters are significant. The predicted probabilities from the model are used to create the predicted mean scores. The data is then ordered by ascending predicted mean score. As the data set for civil staff only consists of 399 respondents, the data is partitioned into 6 groups of approximately equal size - 5 groups of 66 and 1 of 69 - based purely on sample size considerations. For this dataset, there are many tied predicted mean scores, up to 20 on a number of occasions. For all groups except I2, the partition boundaries separate tied values. When the observations are ordered by predicted mean score only, so that observations with tied scores appear in the relative order they do in the original dataset, the goodness-of-fit parameters for indicator variables I1 to I5 are not significant, suggesting the model is an adequate fit.

Table 2.10 : Goodness-of-fit test parameter estimates for model 4.41c
Groups assigned by sample size only

Parameter	Estimate	s.e.	p-value
$\gamma 1$	0.36	1.69	0.8331*
$\gamma 2$	0.50	1.31	0.7031*
$\gamma 3$	0.30	1.12	0.7893*
$\gamma 4$	-0.07	0.90	0.9392*
$\gamma 5$	-0.30	0.57	0.5914*

* denotes parameter not significant

If, however, the data are primarily ordered by predicted mean score, and secondarily by response value, therefore affecting the order of tied predicted mean scores, the results are very different. The groups indicated by I1 to I5, now consist, in part, of different observations. When added to model 4.41c to test the goodness-of-fit, the parameters now suggest that the model is not a good fit to the data, as 3 of the 5 terms are statistically significant :-

Table 2.11: Goodness-of-fit test parameter estimates for model 4.41c
Groups ordered secondarily by response level
and assigned by sample size only

Parameter	Estimate	s.e.	p-value
$\gamma 1$	3.45	1.68	0.0404
$\gamma 2$	2.85	1.30	0.0287
$\gamma 3$	2.54	1.12	0.0234
$\gamma 4$	1.34	0.90	0.1381*
$\gamma 5$	0.70	0.57	0.2241*

* denotes parameter not significant

Therefore when partitioning the data by sample size, the order of tied predicted mean scores may significantly affect the results of the goodness-of-fit test of a model. In this instance, the reason the model fails may be because when the data is ordered by response, as well as predicted mean score. If tied observations are split by the partition, the lower group will contain all the observations with lower responses, possibly giving the indicator variable parameter more accuracy for prediction. For example, let's say group 1 contains the 50 observations with the lowest predicted mean scores, and either side of the partition between groups 1 and 2 there are 20 tied observations with various responses. If the tied values are ordered by response, so that the 20 ties in group 1 all have response 1, instead of randomly distributed, and similarly the responses for the 20 ties in group 2 will be more uniform, this gives greater accuracy to the predictive ability of parameters for I1 and I2. Therefore these parameters will account for more variation in the data than they would otherwise, possibly rejecting a model when it may be adequate.

If the data is ordered by predicted mean score only, so tied observations are ordered as originally, the partitions should be made at a point closest to the desired sample size, where predicted mean scores are not tied. Using this guideline, the groups formed for assessing the fit of model 4.41c, are in the ratio 65:67:73:63:60:71, thus different from the 66:66:66:66:66:69 before. This difference in the ratios of the partitioned groups should make no technical difference to the method, as the partitions described by Lipsitz et al (1996) are based on approximately equally sized groups. The parameters for the extra group indicator variables in this case are non-significant, showing that model 4.41c does seem to fit the data adequately :-

Table 2.12: Goodness-of-fit test parameter estimates for model 4.41c

Groups assigned by not separating respondents with tied scores

Parameter	Estimate	s.e.	p-value
$\gamma1$	0.45	1.77	0.8012*
$\gamma2$	0.70	1.36	0.6060*
$\gamma3$	0.30	1.16	0.7922*
$\gamma4$	0.25	0.92	0.7896*
$\gamma5$	-0.28	0.57	0.6207*

* denotes parameter not significant

If the observations tied on predicted mean scores are ordered in any secondary way, for the same method of partitioning as previously, the results of fitting the extra parameters for I1 to I5 do not alter, as the respondents that constitute each of the partitioned groups will not differ.

Therefore, when using discrete or categorical independent variables, or if any ties for predicted mean scores occur, to apply the Lipsitz et al (1996) goodness-of-fit test, it is essential that the tied observations are not split between partitioned groups. The model specification given as an example by Lipsitz et al involves continuous variables, therefore the presence of tied values may not have been an issue, as there are potentially many more possibilities for permutations of independent variable values. The indications above can therefore be seen as a modification to the guidelines for using the Lipsitz, Fitzmaurice and Molenberghs (1996) goodness-of-fit test, for discrete independent variables, or conditions where tied predicted mean scores occur.

2.5.4: Diagnostic plotting

As an illustration of the performance of a model, the predicted mean scores, as calculated for the Lipsitz et al goodness-of-fit test, may be plotted against the actual observed responses. For a model that fits perfectly, one would be able to join the points with a diagonal line through the origin. More realistically, a plot that shows large concentrations of observations along the areas of the diagonal, with a definite pattern of increasing predicted mean score as observed response increases, would indicate an adequate model. Discussion on plotting diagnostics can be found in Hosmer and Lemeshow (1989), and Landwehr, Pregibon and Shoemaker (1984).

Chapter 3: The Data and Exploratory Data Analysis

3.1: The South Yorkshire Police data

3.1.1: The South Yorkshire Police staff survey

A study on the South Yorkshire Police was carried out by the Survey and Statistical Research Centre at Sheffield Hallam University (SSRC (1994)). The objective of the survey being to examine the quality of service provided by the police, experienced by the public, and to match this information against the opinions of South Yorkshire Police staff themselves.

This research involves the data pertaining to the responses to the questionnaire in Appendix 1, by South Yorkshire Police staff, regarding their opinions on various aspects of their job and the South Yorkshire Police on the whole. The data was collected by a postal survey given to all South Yorkshire Police staff. Roughly 50% returned the questionnaire resulting in a total sample of 2031 respondents, this rate of response was less than that anticipated. It is believed this was due to factors such as concerns about confidentiality, several internal surveys being carried out at the same time (therefore deflecting motivation away from completing the SSRC survey), and feelings that the 1994 survey was too soon after a similar study carried out in 1992.

The questionnaire given to the Police staff was designed to address the following topics :-

Quality of service
The structure of South Yorkshire Police
Career development
Staffing and resources
Morale

3.1.2: Background information on the South Yorkshire Police

The following information gives a description of the demographic make-up of the South Yorkshire Police, from the sample obtained.

Table 3.1: Percentages of Men and Women in the Survey Sample

	Percent
Male	75.8
Female	24.2
Total	1999*

* 32 missing observations

As table 3.1 shows, the sample is heavily male dominated, this is similar to the actual population of the South Yorkshire Police, although there was a higher proportion of female non-responders (response rate = 38.4%) than male (response rate = 50.0%).

Table 3.2: Percentages of Respondents' Ethnic Origins in the Survey Sample

	Percent
White	98.5
Black	0.3
Asian (Indian Subcontinent)	0.3
Other	0.9
Total	1992*

* 39 missing observations

The sample is hugely weighted towards whites, although whether this is a true reflection of the ethnic make up of the South Yorkshire Police is uncertain, as no population information is given. It would seem that any differences in attitude between whites and non-whites would either not be detected, or be unreliable due to the sparseness of data for the non-white ethnic groups.

Table 3.3: Percentages of Respondents' Lengths of Service in the Survey Sample

	Percent
Less than 2 yrs	10.7
2 - 5 yrs	14.7
6 - 10 yrs	20.3
11 - 20 yrs	33.2
21+ yrs	21.1
Total	1998*

* 33 missing observations

The police force is the sort of profession where long serving individuals are fairly likely. As can be seen, over half of respondents have served in the police force for 11 or more years.

Table 3.4: Percentages of Civil Support Staff and Police Officers in the Survey Sample

	Percent
Police Officers	76.6
Civil Support Staff	23.4
Total	1996*

* 35 missing observations

The sample contains a far greater proportion of police officers than civilian staff. The population figures are fairly similar 70% and 30% respectively. The breakdown of the actual ranks and grades of the respondents is given below.

Table 3.5: Percentages of Respondents in Police Officer grades in the Survey Sample

	Percent
Police Constable	68.8
Police Sergeant	18.9
Inspector	7.3
Chief Inspector/ Superintendent	4.1
More Senior than Superintendent	0.9
Total	1502*

Table 3.6: Percentages of Respondents in Civil Support staff grades in the Survey Sample

	Percent
Principal/Senior Officer	6.9
Scale 4 - 6	14.5
Scale 1 - 3	59.8
Hourly paid work member	12.2
Traffic Warden	6.5
Total	433*

* 96 missing observations

3.1.3: Morale and morale of the South Yorkshire Police

Morale, one of the key topics of interest from the survey, is the subject of this investigation. From the survey, morale is measured on an ordinal scale. It is the aim of this study to model morale, using the proportional odds and continuation odds models (sections 2.4.3 and 2.4.4), thus utilising the ordinality of the response.

'Morale is an attitude of satisfaction with, desire to continue in, and willingness to strive for the goals of a particular group or organisation' (Viteles (1954)).

Studies on morale are common in both organisational and psychological contexts. Most studies on morale are concerned with improving or examining the productivity of individuals and groups in industry. The term 'productivity' is perhaps not applicable when pertaining to the police force, although the concept of identifying variable factors which affect levels of morale is fairly general.

In a profession as high profile as the police, the consequences of improperly performed duties are far greater than most jobs, therefore morale and motivation of members of the police force is a very key issue.

Many personal factors can affect morale, however, this study makes the assumption that individuals can separate their personal life from their professional life, at least in terms of statistical information and interpretation, and concentrates on job related factors of interest.

With respect to the police force, aspects that can affect the morale of staff include, internally, things like promotion issues, relationships with other staff, communication within the force, recognition for the job done, as it is probably one of the most physically and mentally stressful and demanding careers one could choose, and possibly the amount of influence on decisions made. Externally, things like public opinion and media coverage could have an effect on the morale of police staff. Also differences in individuals, ie their demographic characteristics, may influence differences in attitudes.

An example of an horrific event that must have had quite an influence on morale within the South Yorkshire Police, both from internal and external sources, was the F.A. Cup semi-final between Liverpool and Nottingham Forest, at Sheffield Wednesday's home ground Hillsborough on April 15th, 1989. 'The Hillsborough disaster' left 95 people dead, crushed by overcrowding in the Leppings Lane end of the stadium. The investigation that followed this tragedy revealed that the South Yorkshire Police failed to cope properly, not only with the potential danger to the football supporters involved, but also the reality of the catastrophe. It was also found that the police actually tried to cover up their errors of judgement by attempting to shift the blame onto the supporters with assertions of drunkenness and 'misbehaviour', allegations deemed to be untrue by the investigation team headed by Lord Justice Taylor (Home Office (1990)).

When these facts were made public, public opinion of the South Yorkshire Police was surely low, exerting external pressure on morale, and from within the force. The morale of those concerned in the tragedy may have been low, and their colleagues' morale affected by the actions of certain officers in this case. This event alone could motivate an investigation into morale and general feeling within the police force.

The information collected on morale as a variable is ordinally scaled, and has factors associated with it logically, theoretically and statistically, which are also represented by some measurement in the survey. Therefore morale has some desirable properties in the context of this research.

3.1.4: Measures of morale

There are 2 measures of morale obtained from the survey. 'Respondents' own morale' is obtained from Q20a from the survey (Appendix 1) :-

> How would you describe your morale at the moment?

with the possible responses very high, high, neither high nor low, low or very low. Therefore the variable is ordinally scaled.

'Colleagues' perceived morale' is obtained similarly from Q20b of the survey :-

> How would you describe your colleagues' morale at the moment?

with the same response options.

Combining these measures, a third measure of morale, termed 'relative morale' can be derived by differencing 'own' and 'colleagues'' morale to give levels of 'own morale higher than colleagues'', 'both own and colleagues' morale the same' and 'colleagues' morale higher than own', so that relative morale keeps the ordinal nature of the two variables it is made from.

It is difficult to gauge which measure of morale from the survey (own or colleagues') is more true or accurate. The responses for the two variables may differ for an individual, in the sense that a respondent may overstate their morale in order not to give the impression of low or lower morale, but may give a truer reflection of morale when referring to his/her colleagues. On the other hand, a respondent may give an honest account of their own morale, but not wish to overestimate morale in general, and therefore give a lower response to the question of colleagues' morale. These hypothetical scenarios may not be a large cause for concern due to the emphasis on the confidentiality of respondents, but the possibility of responses in that nature is not unfeasible. Relative morale uses both measures, and without knowing all the reasons why a respondent may say their own morale is different to their colleagues', interpretation is perhaps not as straightforward as the direct measures of morale.

Table 3.7: Percentages of Respondents' Own Morale

Morale	Percent
Very High	7.7
High	33.7
Neither High nor Low	30.4
Low	20.6
Very Low	7.6
Total	2016

* 15 missing observations

Table 3.8: Percentages of Colleagues' Perceived Morale

Morale	Percent
Very High	1.4
High	18.9
Neither High nor Low	38.4
Low	31.5
Very Low	9.8
Total	2009*

* 22 missing observations

Table 3.9: Percentages of Relative Morale

Morale	Percent
Own Higher	38.9
Both Same	54.0
Colleagues' Higher	7.1
Total	2006*

* 25 missing observations

Respondent's own morale has a stronger association with most potential covariates than the other measures of morale, based on bivariate chi-squared tests for association. Therefore respondent's own morale is the most desirable variable of interest, with respect to interpreting relationships between influencing factors and morale, and possible model building.

3.1.5: Explanatory Variables

Most of the other variables in the study, the possible/potential explanatory covariates, have a significant statistical relationship (based on chi-squared test for association) with morale. Morale is a complex concept, however. A lot of the information in the questionnaire can be seen to represent similar theoretical factors. With so much statistical association between the response variable and potential explanatory variables, the selection process for a set of covariates becomes partly subjective and theoretical/logical, in the sense that one needs to decide which factors have a genuine association with morale, and which variable/s will represent each factor.

This section lists variables with a possible theoretical influence on morale, with justification for their inclusion in the investigation.

Table 3.10 : Respondents' Own Morale by Gender (Percentages)

	Respondents' own morale					
Gender	very high	high	neither	low	very low	**Row total**
Male	8.6	35.5	28.0	20.2	7.6	1512
Female	5.1	28.7	37.1	21.3	7.8	474
Col total	154	673	600	407	152	1986*

* 45 missing observations

Gender has a statistically significant relationship with Respondent's own morale, (χ^2 = 20.9, df=4). Examining the table, the departure from independence seems to stem from the fact that more men responded that their morale was high (35.5% compared to 28.7% of women). Similarly, more women felt their morale was neither high nor low (37.1% compared to 28% of men), the other levels of morale seem similarly dispersed between the sexes.

Colleagues' perceived morale actually has a slightly stronger statistical association with gender (χ^2=25.3, df=4), whilst relative morale is not influenced by sex.

The attitudes of men and women differ in a great many areas (Hollway (1991), therefore the effect of gender on morale seems a logical thing to investigate. Despite decades of debate, the workplace in general is a male dominated area, and the police force is probably a typical example of this. Looking at the breakdown of gender by whether the respondent is a police officer or a member of the civil staff (table 3.11), it can be seen that there is a far greater proportion of men in police officer ranks, and similarly larger proportions of women in the civil staff grades. Due to this, the effect of gender on morale may be seen through the effect of this factor.

Table 3.11 : Police Officer/Civil Staff by Gender (Percentages)

	Police Officer/Civil Staff		
Gender	Officer	Civil	**Row total**
Male	89.5	10.5	1511
Female	35.0	65.0	474
Col total	1519	466	1985*

* 46 missing observations

The association between whether a respondent is a police officer or member of the civilian staff and gender is highly statistically significant (χ^2=597.0, df=1). Therefore the effects of both on morale should be considered separately in terms of model building, and possibly when interpreting the effects of either on morale, the effect of the other should also be mentioned and considered.

Table 3.12: Respondents' Own Morale by Police Officer/Civil Staff (Percentages)

Police Officer/ Civil Staff	Respondents' own morale					
	very high	high	neither	low	very low	**Row total**
Officer	8.7	35.8	28.3	20.3	7.0	1525
Civil	4.4	27.0	36.4	22.2	10.0	459
Col total	153	670	598	411	152	1984

* 47 missing observations

The relationship between a respondent's own morale and whether they are a police officer or civilian staff is statistically significant (χ^2=29.5, df=4). This is slightly stronger (statistically) than the association between gender and morale. The major difference in the two groups' morale seems to be that more officers state their morale is high (note that a higher proportion of men are officers), and more civilian staff feel their morale is neither high nor low (note that a higher proportion of women are civilian staff).

Colleagues' perceived morale and relative morale also have a significant statistical association with whether the respondent is a police officer or member of the civilian staff (χ^2=18.2, df=4 and χ^2=10.9, df=2 respectively).

One could look at whether a respondent is a police officer or civilian worker as being a difference in department (although within the two groups there are many different sections). There are bound to be differences in not only tasks, but objectives and the nature of supervision and leadership, as well. The relationship between supervision/leadership and morale is well documented in texts about morale at work and work psychology (Hollway (1991), Viteles (1954)), although most concentrate on the personal aspect of specific supervisors as motivators, or an effect on morale in terms of productivity (see discussion for tables 3.19, 3.20 and 3.21).

Table 3.13: Respondents' Own Morale by Ethnicity (Percentages)

Ethnicity	Respondents' own morale					
	very high	high	neither	low	very low	**Row total**
White	7.8	34.0	30.2	20.5	7.5	1949
Non-white	3.3	23.3	40.0	26.7	6.7	30
Col total	153	670	600	407	149	1979*

* 52 missing observations

Whether a respondent is black or white (black in this context includes the ethnic minorities Black, Asian (Indian subcontinent) and other) may have an influence on their attitude in some areas (Burt (1924)). However, the data from the SYP survey indicates that ethnicity is not related, statistically, to a respondent's own morale (χ^2=3.3, df=4) or colleagues' perceived morale (χ^2=3.4, df=4). This result is possibly due to the very sparse representation of the ethnic minorities which can be seen in table 3.13. Without any population figures to compare, it is unclear whether this vast inequity is due to a larger proportion of non-responders from ethnic minority groups than white, or whether it is fair reflection of the characteristics of the members of the South Yorkshire Police staff.

Ethnicity is, however, related statistically to relative morale (χ^2=10.0, df=2). Those of a non-white ethnic origin have a greater tendency to feel their own morale is higher than their colleagues, whereas most white respondents feel that their morale is the same as their colleagues' (54.2%). As mentioned before, the sparse numbers of ethnic minority respondents may have a misleading influence on statistical conclusions drawn.

Table 3.14: Relative Morale by Ethnicity (Percentages)

Ethnicity	Relative morale			
	Own	Same	Colleagues'	**Row total**
White	38.9	54.2	6.9	1939
Non-white	46.7	33.3	20.0	30
Col total	768	1061	140	1969*

* 62 missing observations

Table 3.15: Respondents' Own Morale by Length of Service (Percentages)

Length of Service	Respondents' own morale					
	very high	high	neither	low	very low	**Row total**
< 2 yrs	16.2	49.0	22.4	9.5	2.9	210
2 - 5 yrs	8.9	34.4	32.3	18.9	5.5	291
6 - 10 yrs	3.7	34.1	32.6	22.6	7.0	402
11 - 20 yrs	6.4	29.1	31.4	24.1	9.1	660
21 + yrs	8.8	32.7	28.7	28.9	19.9	422
Col total	154	670	601	409	151	1985*

* 46 missing observations

The length of service of a respondent is statistically significantly associated with their morale (χ^2=85.6, df=16). Of those with 2 years or less service in the police force, over 65% felt their morale was either high or very high compared with just over 43% of those with between 2 and 5 years service. The subsequent groups, whilst of those with 21 or more years service, 41.5% said their morale was high or very high. A similar, equivalent, pattern can be observed for respondents with low or very low levels of morale.

Colleagues' perceived morale and relative morale are also significantly statistically related to length of service (χ^2=57.0, df=16 and χ^2=19.8, df=8 respectively).

The length of service in a job has been found to have a statistically significant association with morale in many studies (Viteles (1954)). Other particular studies have found a statistically significant relationship between the length of service of individuals and overall attitudes towards their organisation. Workers who had been in the job 5 years or more tended to have higher average attitude scores than those with 1 to 4 years service (University of Minnesota (1951)). In different studies, the pattern seemed to be that employees started with high morale, which seemed to diminish after a couple of years for a period, and then rose again with greater length of service (Viteles (1954)). Probable reasoning for the latter pattern of differing levels of morale with different lengths of service, is due to the worker starting filled with enthusiasm as the job is new, but after early progress there is not room for all ambitious newcomers to progress quickly, so morale levels drop. Subsequently, after a given number of years in a company, the employee will possibly be of two states of mind - if he/she is of great ability, then this ability may have been recognised and they have advanced, or if the individual has not progressed, they will probably be mature enough to accept that not everyone can advance to the top, and be resigned their fate which may not really be that bad (Hull (1939)). The SYP data, with respect to length of service and morale, shows similar trends to the scenario described above.

Table 3.16: Respondents' Own Morale by Job Satisfaction (Jobsat) (Is the respondent satisfied with their job, yes/no) (Percentages)

	Respondents' own morale					
Jobsat	very high	high	neither	low	very low	**Row total**
Yes	10.2	42.7	31.1	13.1	2.9	1487
No	0.6	8.1	27.6	42.7	21.1	508
Col total	155	676	602	412	150	1995*

* 36 missing values

Job satisfaction is the variable most significantly associated with morale in statistical terms in this study (χ^2=507.7, df=4). Both job satisfaction and morale are fairly general indicators of a worker's happiness, and perhaps influenced by the same things. However, they are not necessarily substitutable measures. Job satisfaction pertains more to an intrinsic aspect of the tasks performed, although the term may not always be perceived that way. An individual could feasibly be satisfied with their job, but have, say, low morale, especially 'at the moment' (as questioned in the SYP survey (Appendix 1)) as can be seen. Table 3.16 shows the distribution of data that gives rise to the statistical association. Aside from the fact that there are three times as many respondents satisfied with their job than there are not, the largest cell frequency for those satisfied with their job can be found for the group that feel their own morale is high (42.7%). Equivalently for those not satisfied with their job, the most populated group is those with low morale (also 42.7%), so the pattern is that which one might expect , i.e. that if an individual is satisfied with their job, then they are more likely to have high morale, or higher than someone who is not satisfied with their job and vice versa.

Colleagues' perceived morale and relative morale also have a statistical association with job satisfaction (χ^2=232.0, df=4 and χ^2=152.5, df=2 respectively).

Table 3.17: Respondents' Own Morale by Perceived Public View of the South Yorkshire Police (Percentages)

Public View	**Respondents' own morale**					
	very high	high	neither	low	very low	**Row total**
V. Positive	35.3	52.9	5.9	5.9	-	17
Positive	10.8	43.5	28.5	14.8	2.5	840
Neither	5.3	27.8	34.9	23.3	8.6	827
Negative	4.4	22.5	25.5	30.9	16.8	298
V. Negative	5.6	5.6	22.2	27.8	38.9	18
Col total	155	672	609	415	149	2000*

* 31 missing observations

How the respondent perceives the public view of the South Yorkshire Police has a highly significant statistical association with their own morale (χ^2=223.3, df=16). This factor represents, in a way, how the respondent feels about the South Yorkshire Police themselves. Unless they have had direct contact with the public, with an aim to find out how they view the SYP, the feelings will be their own. It is unsurprising that there is such a strong relationship, as the two variables are effectively measures of self esteem in this context. The extreme rows in the table, where the public view of SYP is perceived to be very positive and very negative, are sparsely populated, but the nature of the relationship is clear from the fact that 54% of those feeling the public's view is positive, said their morale was either high or very high, compared with 26.9% of those perceiving a negative public view. Similarly 47.7% of respondents who felt the public's view was negative, had low or very low morale, as opposed to only 17.3% of those with a positive perception of the public's view. It may be worth noting that 42% of respondents perceived a positive public view, with 41.4% perceiving neither positive nor negative accounting for over 83% of respondents, whilst 0.9% of respondents perceived very positive and very negative views. This may indicate a tendency to be cautious when speculating how the public feels, and/or a general reluctance to state that the public may have a negative view.

Colleagues' perceived morale has a slightly more significant statistical association with perceived public view than respondent's own morale (χ^2=228.7, df=16). Relative morale is also statistically related to the variable (χ^2=17.0, df=8).

Table 3.18: Respondents' Own Morale by Satisfaction with Service Provided (Percentages)

Service Provided	Respondents' own morale					
	very high	high	neither	low	very low	**Row total**
V. Satisfied	17.5	35.1	29.8	7.0	10.5	57
Satisfied	10.6	41.2	29.7	15.0	3.5	972
Neither	4.7	29.3	36.6	22.6	6.7	464
Dissatisfied	3.9	23.9	27.2	30.7	14.3	482
V. Dissatisfied	3.7	11.1	11.1	40.7	33.3	27
Col total	155	674	610	414	149	2002*

* 29 missing observations

How satisfied a respondent is with the service the SYP provide and his/her morale are statistically significantly associated (χ^2=206.0, df=16). This variable is similar to perceived public's view in what it represents, as it is a measure of how the respondent feels they deal with the public, and how well they feel they (possibly as an individual, section or the whole police force) are doing their job. The pattern of the data is similar to that for the previous factor. The extreme views of service provided - very satisfied and very dissatisfied - are again sparsely populated. 51.8% of those satisfied with the service the SYP provides, had high or high or very morale, and 45.1% of those dissatisfied with the service provided had low or very low morale.

Again, colleagues' perceived morale has a slightly stronger statistical relationship with how satisfied the respondent is with the service provided by the SYP (χ^2=232.2, df=16), whilst the relationship between the covariate and relative morale is not significant (χ^2=10.4, df=8).

Table 3.19: Respondents' Own Morale by Working Relationship with Line Manager (Percentages)

Working Relationship	Respondents' own morale					
	very high	high	neither	low	very low	**Row total**
V. Satisfactory	12.1	43.5	25.8	14.5	4.0	751
Satisfactory	5.0	31.6	36.9	21.1	5.4	857
Neither	4.8	18.3	30.1	32.8	14.0	186
Dissatisfactory	3.8	13.9	20.3	36.7	25.3	79
V. Dissatisfactory	2.9	-	25.7	22.9	48.6	35
Col total	147	643	591	388	139	1908*

* 123 missing observations

Table 3.20: Respondents' Own Morale by Communication with Immediate Supervisors (Percentages)

Communication	Respondents' own morale					
	very high	high	neither	low	very low	**Row total**
V. Good	12.1	43.6	25.0	15.0	4.3	917
Good	4.1	29.3	38.1	22.0	6.6	788
Neither	3.5	15.0	31.5	35.0	15.0	200
Bad	5.2	18.2	22.1	33.8	20.8	77
V. Bad	4.2	-	12.5	25.0	58.3	24
Col total	155	675	612	413	151	2006*

* 25 missing observations

Table 3.21: Respondents' Own Morale by Communication with More Senior Managers/Officers (Percentages)

Communication	Respondents' own morale					
	very high	high	neither	low	very low	**Row total**
V. Good	16.8	46.7	23.4	10.9	2.3	304
Good	7.6	39.7	32.1	15.8	4.9	761
Neither	5.8	28.1	36.2	23.1	6.8	602
Bad	1.9	20.4	24.5	38.9	14.3	265
V. Bad	1.7	8.3	23.3	21.7	45.0	60
Col total	150	672	612	408	150	1992*

* 39 missing observations

The variables 'working relationship with line manager', 'communication with immediate supervisors' and 'communication with more senior managers/officers' offer a different, probably better representation of the factor 'supervision and management', than whether the respondent is civilian staff or police officer. 'Working relationship with line manager' and 'communication with immediate supervisor' are effectively substitutes, as their definitions are almost identical. Both are highly associated with each other, statistically (χ^2=1661.7, df=16). All three measures are highly correlated, therefore when model building it is probably advisable that only one of the measures be used in any single model. All are highly statistically related to respondent's own morale (χ^2=289.3, 306.6 and 335.9 respectively, df=16). In each case, the nature of the relationship is the same, high proportions of respondents who felt the relationship was very satisfactory, or that communication was very good, felt their morale was high or very high. Equivalently, high proportions feeling the relationship was very unsatisfactory, or communication was very bad, also thought their morale was low or very low.

The association between colleagues' perceived morale and 'working relationship with line manager', 'communication with immediate supervisors' and 'communication with more senior managers/officers' is also statistically significant (χ^2=197.9, 199.9 and 238.1, df=16 respectively). The same applies for relative morale (χ^2=43.9, 44.5 and 29.9, df=8 respectively).

The hypothesis that the motivation and morale of an individual at work are influenced by the quality of supervision is one frequently tested, although most, if not all documentation refers to the effect on productivity, usually in terms of profit or suchlike. For example, in a study where 22 sections of a company were assessed on profit over a period of time, the supervisors whose sections had achieved the greatest increase in profit, in the first period, were transferred to the sections where the lowest increases in profit were recorded (and vice versa). The assessment was repeated. The supervisors who achieved the largest increases in profit during the first period, also achieved the biggest increases in profit with their new sections, that had managed the lowest profit increases previously, therefore showing an association between supervision and motivation or morale (Feldman (1937)).

Table 3.22: Respondents' Own Morale by Promotions Earned (Percentages)

Promotions Earned	Respondents' own morale					
	very high	high	neither	low	very low	**Row total**
Strongly Agree	27.5	32.5	25.0	10.0	5.0	40
Agree	11.0	46.7	27.5	14.3	0.4	454
Neither	8.0	34.5	35.3	18.0	4.3	678
Disagree	4.4	27.1	29.9	29.5	9.1	572
Strongly Disagree	3.3	22.3	21.2	21.2	32.1	184
Col total	146	655	584	399	144	1928*

* 103 missing observations

Table 3.23: Respondents' Own Morale by 'It's Not What You Know, It's Who You know'(Percentages)

Who You Know	**Respondents' own morale**					
	very high	high	neither	low	very low	**Row total**
Strongly Agree	4.3	17.1	29.8	25.4	23.4	299
Agree	5.5	29.1	33.2	25.9	6.3	745
Neither	7.7	41.0	30.6	16.4	4.3	585
Disagree	12.4	49.0	26.1	10.8	1.6	249
Strongly Disagree	23.9	38.0	22.5	12.7	2.8	71
Col total	147	657	596	401	148	1949*

* 82 missing observations

Promotion issues have a theoretical association with morale, as well as a highly significant statistical one in terms of the variables 'Promotions are given to those who earn them' and 'It's not what you know, it's who you know' (χ^2= 315.5 and 257.2, df=16). Those with first hand experience of some sort of promotion situation will probably be affected more than those without. For example, someone not receiving a promotion they felt they ought to have got, may have lower morale, and so probably also respond negatively to the statements above (ie disagree/strongly disagree and agree/strongly agree respectively). Someone receiving a promotion may feel the opposite. The departure from independence between the promotion issue variables and respondent's own morale shows no unusual traits in both cases. For the statement 'Promotions are given to those who earn them' large proportions of respondents who said they agree or agree strongly had high or very high morale. Large proportions of those who disagree strongly had low or very low morale. For 'It's not what you know, who you know' there is the same pattern in reverse due to the contrary nature of the statement.

Colleagues' perceived morale is associated statistically with the promotion issue variables (χ^2=211.4, df=16 for 'promotions are given to those who earn them' and

χ^2=211.1, df=16 for 'it's not what you know, it's who you know'). Relative morale is also statistically related to the factors (χ^2=36.4 and 30.1, df=8 respectively).

3.2: Application of CHAID to SYP data

This section illustrates the use of the SPSS package CHAID - CHi-square Automatic Interaction Detection, by using the technique to analyse the South Yorkshire Police data. The methodology for the package is described in section 2.2.3. For the following exploratory analyses, integer scores for both the response and explanatory variables, i.e. not calibrated scores, are used.

3.2.1: Variable Selection

The variables included in the analysis are those discussed earlier in sections 3.1.4 and 3.1.5. In order that all the variables are used in an analysis, the predictors are divided into subsets, due to the fact that it is not efficient to use all at once as some represent the same theoretical factor. There are 3 subsets of variables due to there being 3 representations of the theoretical factor 'supervision and management'.

Predictors subset 1

Job Satisfaction (jobsat)
Length of Service (lenserv)
Ethnicity (borw)
Police Officer/Civil Staff (officer)
Perceived Public View of SYP (pubview)
Communication with senior managers/officers (commsen)
Promotions given to those who earn them (promearn)

Predictors Subset 2

Job Satisfaction (jobsat)
Length of Service (lenserv)
Ethnicity (borw)
Gender (gender)
Service Provided by SYP (service)
Communication with Immediate Supervisors (commimm)
'It's not what you know, it's who you know' (whouknow)

Predictors Subset 3

Job Satisfaction (jobsat)
Length of Service (lenserv)
Ethnicity (borw)
Police Officer/Civil Staff (officer)
Perceived Public View of SYP (pubview)
Working Relationship with Line Manager (linemgr)
'It's not what you know, it's who you know' (whouknow)

3.2.2: CHAID Analysis of the South Yorkshire Police data

The dependent variable in the analysis is morale in its different guises, therefore as these variables are ordinal the analysis will be based on results using the ordinal method in CHAID.

3.2.2a: Respondent's Own Morale and Predictors subset 1

Diagram 3.1a shows the result of the CHAID analysis of respondent's own morale with predictors subset 1. The parent node at the top of the diagram is at depth 0, and represents the full sample of 2016 respondents who gave a valid response about their own morale. The score of 2.87 inside the node is the mean morale score for the respondents, indicating an overall level of neither high nor low morale in general, tending slightly more towards high than low.

The variable with the strongest association with omor is jobsat, so CHAID splits the sample into two subsets, corresponding to the levels of jobsat. For example, taking the first 'child' node, which contains only respondents who are satisfied with their job, n=1487 of the original 2016, and their mean morale score is 2.56, indicating their average level of morale is between high and neither high nor low. The general morale of this subgroup is higher than that for the whole sample (morale score is lower), as one might expect for respondents who are satisfied with their job.

The second 'child' node at depth 1 contains information on 529 respondents who stated they were either not satisfied with their job or unsure or did not answer the question. CHAID combines the missing value category with the second jobsat category due to the minimum subgroup size constraint (nodes must contain at least 50 observations - this setting can be varied). The respondents dissatisfied with their job or unsure have a mean morale score of 3.76 (n=508). The group with missing values for jobsat, containing just 21 respondents, has a mean morale score of 3.19, therefore this group's mean level of morale lies fairly centrally, between the general levels of morale of the dissatisfied or unsure group and the satisfied subset, but is merged with the former due to the subgroup size constraints. This category merger can be prevented by setting the merge level for jobsat to 1, so no categories are combined. The average morale score for this group is 3.73, i.e. the level of morale in general is between low and neither high nor low, tending more towards low. This level of morale is below the overall average, again logically, as most of the group (excluding the non-responders to the question) have admitted they are dissatisfied with their job or unsure.

CHAID then works off these two subgroups independently to find significant predictors with the response omor. Taking the first node at depth 1, where respondents are satisfied with their job (jobsat=1), the variable commsen (Communication with more senior managers/officers) is the one with the strongest statistical association with the response. The split produces 4 depth 2 child nodes (rather than the 6 which would correspond the levels of commsen plus the missing category), due to the minimum subgroup size constraints. Categories 4 and 5 (where communication is bad and very bad respectively) have been merged, and now represent a single subgroup, and the same applies to the first category, where communication is good, and the category for missing responses to commsen.

The first node at depth 2 contains 281 respondents, who said that communication between them and more senior managers was very good (or gave no response). These respondents also stated that they are satisfied with their job. The mean morale score for the subgroup is 2.21, reflecting relatively high morale in terms of the sample as a whole. This level of morale is between high and neither high nor low.

The two nodes pertaining to levels 2 and 3 of commsen (at depth 2) can be interpreted straightforwardly, similarly to above.

Within the node pertaining to communication with more senior managers/officers being bad and very bad (node 4 at depth 2), the categories are merged to give a single level, due again to the minimum subgroup size constraints - level 5 of commsen, for those who are satisfied with their job, contains only 22 observations. The mean morale score for this subgroup is 3.05, reflecting neither high nor low morale in general. The mean morale score for those who felt communication was bad (category 4) is 2.99, i.e. neither high nor low morale, with a subgroup size of 135. The 22 responding in category 5 of the variable have mean morale score of 3.45, indicating a general level of morale between low and neither high nor low. The mean morale levels of these groups are not necessarily statistically similar, but the size of the group that felt communication was very bad necessitated the merge. The subgroup size constraints could be adjusted to avoid this. In this case, the size of the group in question may cause misleading results, and possibly problems when modelling, therefore the category merge is probably beneficial.

Returning to depth 1, the subgroup containing respondents who are dissatisfied with their job, or unsure, and those with missing values for the variable jobsat, identifies 'promotions are given to those who earn them' (promearn) as the most significant predictor for respondent's own morale. Contrary to the job satisfied subgroup at depth 1. This illustrates CHAID's intricacy, when exploring significant relationships between variables, and examining the specific patterns in the dataset that would be incredibly painstaking and difficult to detect by without the package. Categories 1 and 2 of promearn, where respondents agree strongly and agree with the statement, respectively, are merged to give a single level, again due to the after merge subgroup size constraint - level 1 has just 7 respondents. The mean morale scores are fairly similar for the two levels. For level 1 the mean morale level is 3.43, whilst the 75 respondents in category 2 have a mean score of 3.20. Therefore it is feasible that the categories may have been merged by their statistical similarity. Also, the nature of the categories, i.e. they are representing the states 'agree strongly' and 'agree', suggests that a merger is desirable given the subgroup sizes. The resultant subgroup has contains 82 respondents with a mean morale score of 3.22, reflecting neither high nor low morale in general, tending towards low. From this subgroup, no significant predictors are identified, so the node is a completed path, i.e. a segment.

The node pertaining to neither agree nor disagree with the statement 'Promotions are given to those who earn them' (level 3) can be interpreted simply, and is a segment. The node containing those disagreeing with the statement, and missing values, is a merged level, due to the after merge subgroup size constraint (note the mean scores for both categories are very similar -for level 4, mean morale score = 3.84, n = 178, for missing, mean score = 3.78, n=27). This segment has a mean morale score of 3.83, indicating a low general level of morale, tending slightly towards neither high nor low.

The segment containing 84 respondents who are dissatisfied with their job, or unsure, and strongly disagree with the statement 'Promotions are given to those who earn them', have the lowest general level of morale in the analysis (mean score =4.30), reflecting between low and very low morale. This is no surprise, as the group exhibit the most negatively natured characteristics, in the context of the available explanatory variables.

As the analysis gets deeper, in terms of depth on the diagram, there is more propensity for categories of predictors to be combined, due to the subgroups getting smaller. The dataset is split up on more significant predictors, with the minimum subgroup size constraints in effect, to help keep a bit more stability in the analysis regarding general sample size guidelines (Freeman (1987)). What is interesting to note about CHAID diagram 3.1a, is that despite the large amount of category merging at depth 3 and 4, no merge infringes on the logical distinction between levels of the ordinal variables. For example, for the perceived public view of SYP, the responses representing positive and negative views, i.e. levels 1 and 2 and levels 4 and 5, are never combined. The same applies to the statement variable 'Promotions are given to those who earn them', those who agree or agree strongly are never contained in the same subgroup as those who disagree or disagree strongly. Those giving neutral response to the variables, when combined with respondents in other categories, tend to be grouped with the negatively natured levels of the predictors. This may be due perhaps to the idea that people will tend more to hide negative feelings with impartiality than positive feelings. This general result is fairly desirable, in the sense that it is more or less equivalent to how one would merge categories of these, or other likert scale variables, by intuitively combining categories by the nature of the responses they represent.

Segment 1 on the diagram suggests the possibility of a statistically significant interaction between the independent variables jobsat, commsen and officer, that is to say the effect of each of these variables on respondent's own morale, omor, is different at different levels of the others. The segment is a subgroup of respondents satisfied with their job, feel communication with more senior managers is very good, and are civil support staff, who have a distinct statistical relationship with omor. It should be noted that the relationships depicted by segment 1, are based on partial associations, pertaining to a relatively small subset of the data, therefore the implications of this descriptive analysis may not apply, or may not be true for the full sample. If the relationships are assumed to be true, the path could be equated to a model expression, for instance if forming a logit model (section 2.3.3), the segment would represent a 3 way interaction term :-

$$\ln(Y^*) = \lambda + \lambda_{(job)1} + \lambda_{(com)1} + \lambda_{(off)c} + \lambda_{(job*com*off)11c}$$

Where $\ln(Y^*)$ is the function of the response, omor, and the subscripts for $\lambda_{(com)}$ and $\lambda_{(off)}$ of 1 and c ignore the presence of missing values as these are not normally included when modelling.

The group contains 86 respondents with a mean morale score of 2.57, indicating a level of morale between high and neither high nor low in general.

The analysis produces 18 segments which could harbour useful information when model building, in terms of indicating variables with the strongest statistical association with respondent's own morale. Therefore the package at least offers a starting point, if not a full list of variables to include in a model, and also a head start with the inclusion of interaction terms which may improve a model, if necessary.

The purely descriptive side to the segments allows identification of groups whose morale is particularly good or bad, or similarly to describe the morale of a group with given characteristics of interest. It is this property of the package which is probably of most practical use.

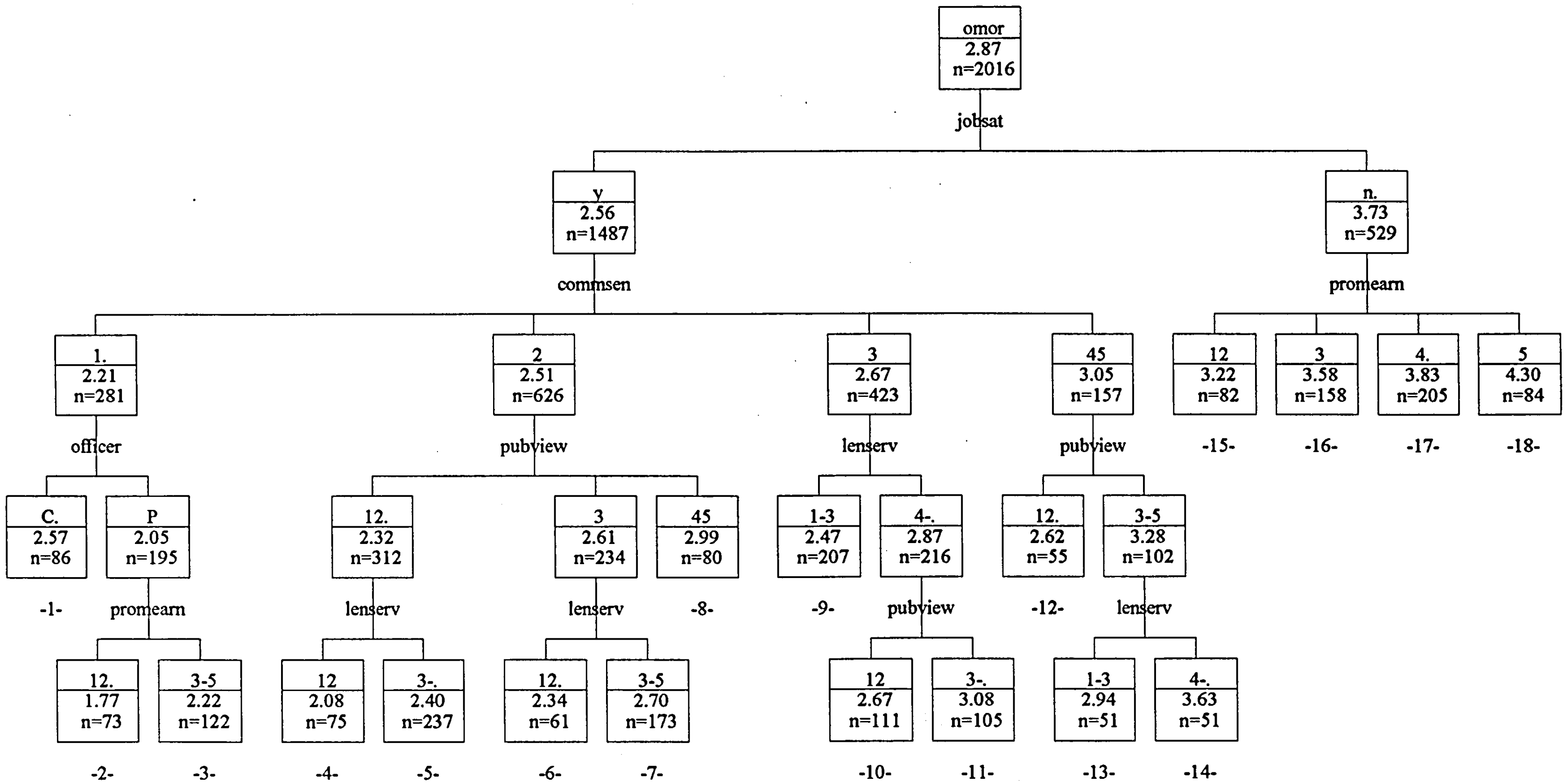
The segmented subgroup with the highest general level of morale is 73 respondents who are satisfied with their job, feel communication with more senior managers/officers is very good, are police officers and responded to the statement 'Promotions are given to those who earn them' with agree or agree strongly (segment 2 on the diagram). Their mean morale score is 1.77, between high and very high morale. This group exhibit the most positive characteristics (excluding those with missing values), in the context of the variables measured, therefore the effect of these traits on an individual's morale is as expected.

The segment with the lowest morale in general, as mentioned earlier, is the 84 respondents dissatisfied with their job or unsure and strongly disagree with the statement 'Promotions are given to those who earn them', their mean morale score is

4.30, reflecting between low and very low morale in general. Again, this result is as one would expect from the nature of the characteristics of this group.

The possible implications for model building are that all the variables in predictors subset 1, except ethnicity, have a statistically significant association with respondent's own morale in some context. Job satisfaction is the variable with the strongest statistical association. The relationships depicted have intuitive interpretations, i.e. the direction of the associations, between the explanatory variables and morale, is as one would expect, so that positive characteristics are associated with higher general levels of morale than negative characteristics, including potential interactions between the independent variables. Collapsing of the levels of all or some of the ordinal variables may be beneficial, in terms of sufficient sample sizes in the groups when modelling (see Freeman (1987) for discussion on sample size guidelines).

CHAID Diagram 3.1a: Respondent's Own Morale against Predictors Subset 1



CHAID segment index for Diagram 3.1a

id	count	score	vars...			
-1-	86	2.57	jobsat=y	commsen=1.	officer=C.	
-2-	73	1.77	jobsat=y	commsen=1.	officer=P	promearn=12.
-3-	122	2.22	jobsat=y	commsen=1.	officer=P	promearn=3-5
-4-	75	2.08	jobsat=y	commsen=2	pubview=12.	lenserv=12
-5-	237	2.40	jobsat=y	commsen=2	pubview=12.	lenserv=3-.
-6-	61	2.34	jobsat=y	commsen=2	pubview=3	lenserv=12.
-7-	173	2.70	jobsat=y	commsen=2	pubview=3	lenserv=3-5
-8-	80	2.99	jobsat=y	commsen=2	pubview=45	
-9-	207	2.47	jobsat=y	commsen=3	lenserv=1-3	
-10-	111	2.67	jobsat=y	commsen=3	lenserv=4-.	pubview=12
-11-	105	3.08	jobsat=y	commsen=3	lenserv=4-.	pubview=3-.
-12-	55	2.62	jobsat=y	commsen=45	pubview=12.	
-13-	51	2.94	jobsat=y	commsen=45	pubview=3-5	lenserv=1-3
-14-	51	3.63	jobsat=y	commsen=45	pubview=3-5	lenserv=4-.
-15-	82	3.22	jobsat=n.	promearn=12		
-16-	158	3.58	jobsat=n.	promearn=3		
-17-	205	3.83	jobsat=n.	promearn=4.		
-18-	84	4.30	jobsat=n.	promearn=5		

3.2.2b: Respondent's Own Morale and Predictors subset 2

CHAID Diagram 3.1b shows the result of analysing respondent's own morale with predictors subset 2 listed above. The initial stages of the analysis correspond to the analysis using predictors subset 1. The full sample splits on job satisfaction, and the group satisfied with their job identify the factor representing 'supervision and management' (communication with immediate supervisors, commimm) as the most significant predictor with the response. Those dissatisfied or unsure or with missing response for jobsat, find the variable pertaining to 'promotion issues' (response to the statement 'It's not what you know, it's who you know', whouknow) as the one with the strongest statistical association with omor.

Beyond this, the segmentation takes on a slightly different shape to the previous analysis, possibly due to the fact that predictors subset 1 contains all the eligible variables with the strongest direct statistical association with respondent's own morale. For predictors subset 2, the theoretical factors that affect morale, discussed in sections 3.13, 3.14 and 3.15, are represented by different, effectively 'weaker' indicators. Having said this, the analysis carries on to depth 5 and 6 on the left hand side, where respondents are satisfied with their job, indicating some interaction between 5 and 6 variables. On the right hand side, for those not job satisfied or unsure and those with missing values for jobsat, the analysis goes 1 depth further than before, splitting on two different predictors at different levels of whouknow, showing differing associations between morale and the explanatory variables commimm and service at different levels of whouknow. The paths to segments 9 and 10 on the diagram represent a 6 way interaction between specific levels of jobsat, commimm, lenserv, whouknow, service and gender. Whilst this is statistically significant and very comprehensive in descriptive terms, with respect to modelling, the terms representing these segments would make a model very intricate and more difficult to interpret. The relationships described are based on partial associations and small subsets of the data, and may not be useful or applicable when modelling respondent's own morale.

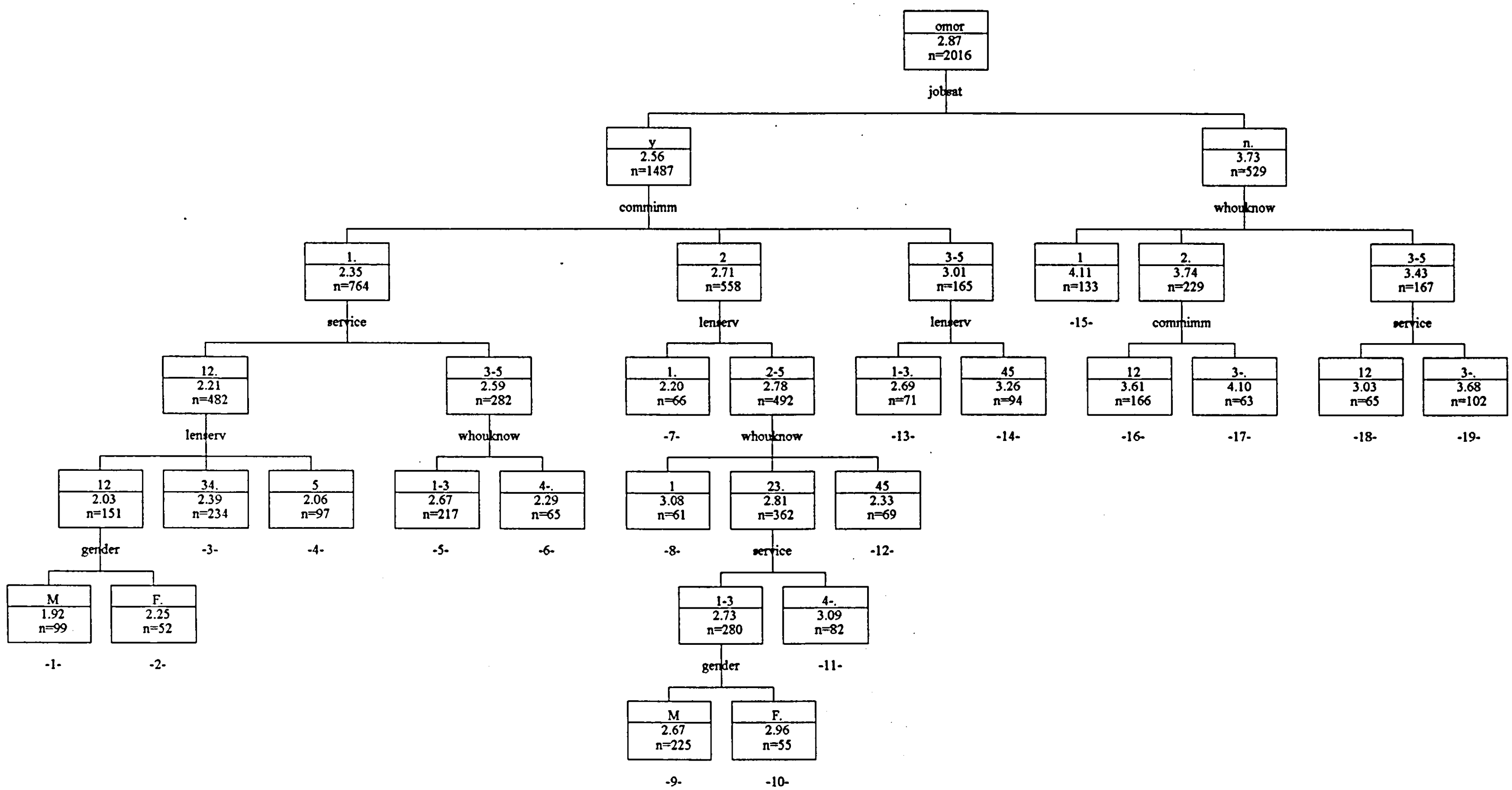
As for the previous analysis, all category merging leaves levels of predictors that represent contrary responses separate.

The group with the highest general level of morale is 99 respondents who are satisfied with their job, feel communication with immediate supervisors is very good or have missing response for commimm, are satisfied or very satisfied with the service provided by SYP or have missing response for service, have served in the police force for 5 years or less and are male. Their average level of morale is high (mean score = 1.92). The characteristics of the levels of jobsat, commimm and service the most positive for those predictors, so the high morale of the group is no surprise. As described in section 3.1.5, a larger proportion of those who have served in the police force for less than 2 years or 2 - 5 years, said their own morale was high or very high, than any other length of service group. Also, approximately 10% more males said that their morale was high or very high than did females.

The segmented group with the worst average level of morale corresponds with that identified using predictors subset 1. The subgroup contains 133 respondents who are dissatisfied with their job ,or gave no response for jobsat, and agree strongly with the statement 'It's not what you know, it's who you know'. Their general level of morale is low (mean score = 4.11). These traits represent the negative feelings towards job satisfaction and promotion issues, as displayed by the equivalent group in the previous analysis.

The most intricately defined groups in the analysis are located in segments 9 and 10 on CHAID Diagram 3.1b. The two subsets are identical to depth 5 on the diagram, the respondents are satisfied with their job, feel communication with immediate supervisors is good, have served in the police force for over 2 years, they agree or neither agree nor disagree with the statement 'It's not what you know, it's who you know' and they are either satisfied, very satisfied or neither satisfied nor dissatisfied, with the service provided by SYP. The 225 respondents in segment 9 are male, with a general level of morale between high and neither high nor low (mean score = 2.67), whilst the 55 in segment 10 are female with average morale neither high nor low (mean score = 2.96). These groups, although intricately defined considering the depth of the analysis, are not very interpretationally distinct. For example the width of the length of service band, agreeing or neither agreeing nor disagreeing with the statement 'It's not what you know, it's who you know', and being satisfied, very satisfied or neither satisfied nor dissatisfied with the service provided by SYP, does not really allow any specific conclusions to be made about what characteristics may influence morale. This results from the category merging of some levels that maybe should stay exclusive. The problem can be overcome by changing the subgroup size constraints, that were possibly instrumental in the merging, or setting the merge level for the predictors of interest so that the categories cannot be combined. Enforcing either of those will change the results of the analysis, and therefore the groups described would probably not exist, however, it is important to be able to interpret segments sensibly.

CHAID Diagram 3.1b: Respondent's Own Morale against Predictor Subset 2



CHAID segment index for Diagram 3.1b

id	count	score	vars...			
-1-	99	1.92	jobsat=y	commimm=1.	service=12.	lenserv=12
	gender=M					
-2-	52	2.25	jobsat=y	commimm=1.	service=12.	lenserv=12
	gender=F.					
-3-	234	2.39	jobsat=y	commimm=1.	service=12.	lenserv=34.
-4-	97	2.06	jobsat=y	commimm=1.	service=12.	lenserv=5
-5-	217	2.67	jobsat=y	commimm=1.	service=3-5	whouknow=1-3
-6-	65	2.29	jobsat=y	commimm=1.	service=3-5	whouknow=4-.
-7-	66	2.20	jobsat=y	commimm=2	lenserv=1.	
-8-	61	3.08	jobsat=y	commimm=2	lenserv=2-5	whouknow=1
-9-	225	2.67	jobsat=y	commimm=2	lenserv=2-5	whouknow=23.
	service=1-3		gender=M			
-10-	55	2.96	jobsat=y	commimm=2	lenserv=2-5	whouknow=23.
	service=1-3		gender=F.			
-11-	82	3.09	jobsat=y	commimm=2	lenserv=2-5	whouknow=23.
	service=4-.					
-12-	69	2.33	jobsat=y	commimm=2	lenserv=2-5	whouknow=45
-13-	71	2.69	jobsat=y	commimm=3-5	lenserv=1-3.	
-14-	94	3.26	jobsat=y	commimm=3-5	lenserv=45	
-15-	133	4.11	jobsat=n.	whouknow=1		
-16-	166	3.61	jobsat=n.	whouknow=2.	commimm=12	
-17-	63	4.10	jobsat=n.	whouknow=2.	commimm=3-.	
-18-	65	3.03	jobsat=n.	whouknow=3-5	service=12	
-19-	102	3.68	jobsat=n.	whouknow=3-5	service=3-.	

3.2.2c: Respondent's Own Morale and Predictors subset 3

CHAID Diagram 3.1c shows the result of analysing respondent's own morale using predictors subset 3. The Initial split of the full sample again identifies jobsat as the most significant predictor.

In the previous two analyses, the variable representing supervision and management was chosen as the most significant predictor for respondent's own morale, using the subgroup where respondents are satisfied with their job. The variable linemgr, however, has a slightly weaker direct statistical association with omor, which may be why perceived public view of SYP (pubview) is the most significant predictor for the job satisfied group in this case. The variable linemgr is, however, identified as a significant predictor in the analysis

Those dissatisfied with their job or with missing values for jobsat, divide on the statement variable 'It's not what you know, it's who you know' (whouknow). It is interesting to note that from the subgroups created by this split, linemgr and pubview are identified as significant predictors at depth 2, in a directly correspondent manner to commimm and service in the previous analysis. The variables linemgr and commimm, and pubview and service, are substitutes for each other, as each pair depicts the same theoretical factor related to morale.

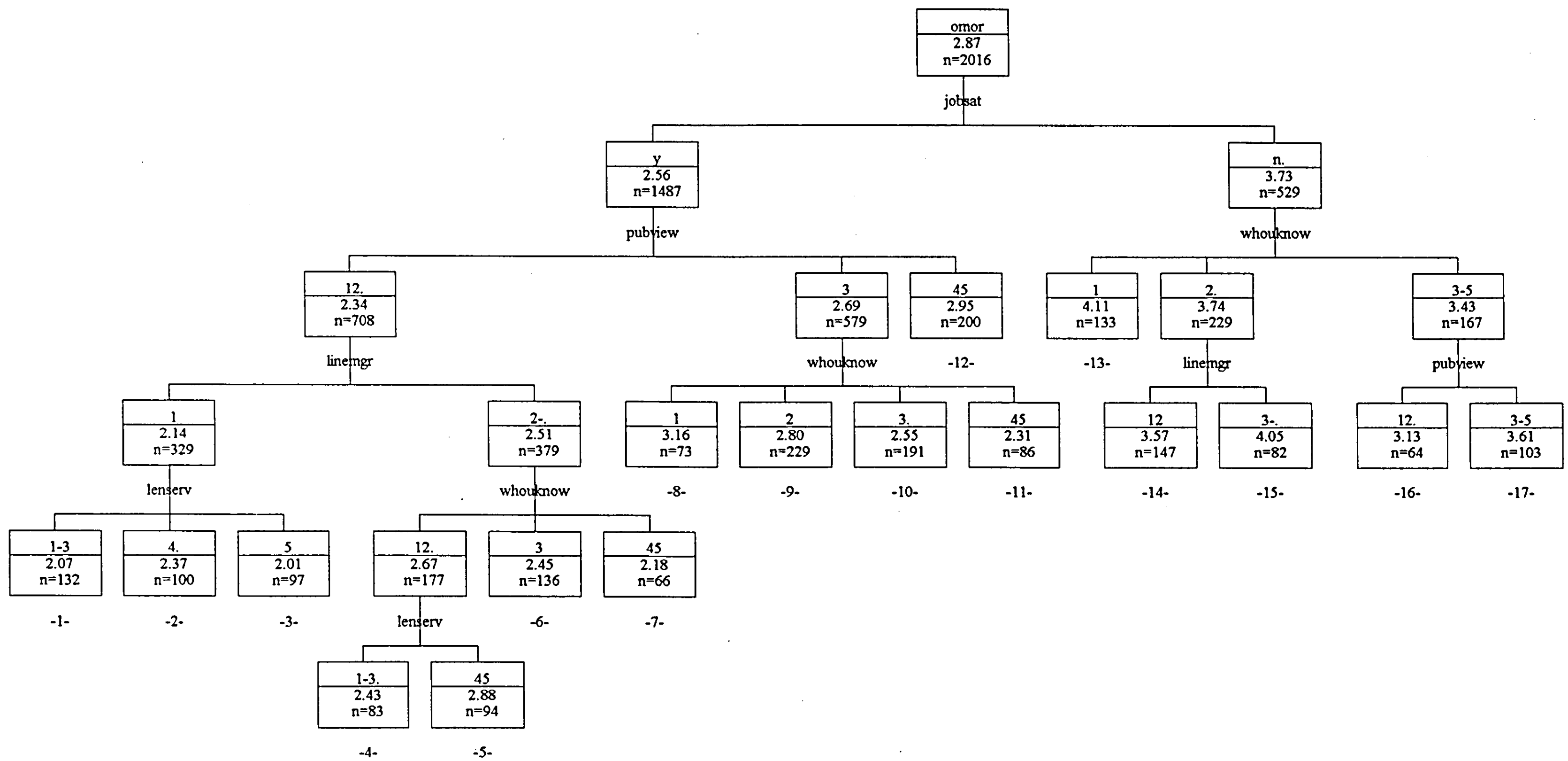
A possibly unexpected exclusion from the analysis is the variable officer, as this variable is significant in the subgroup with the highest general level of morale in the first analysis. Its absence may be due to the interaction with commsen previously, whereas the variable linemgr behaves differently with respect to morale.

Working relationship with line manager splits on the group satisfied with their job, and satisfied or very satisfied with the service provided by SYP. This split leads to categories of linemgr with contrary definitions being merged. Whilst there are no technical rules being broken, the resultant category, containing responses for linemgr of satisfactory, neither satisfactory nor dissatisfactory, dissatisfactory and very dissatisfactory (as well as missing values), is almost impossible to interpret usefully. As discussed above in the previous analysis, the subgroup size constraints could be changed, or the merge level set to prevent the combining of categories for a predictor or predictors. Merging categories manually before performing a CHAID analysis using the variables, and then using the constraints to aid sensible interpretation of results is also feasible.

The subgroup with the highest general level of morale is those in segment 3 on CHAID Diagram 3.1c. The 97 respondents are satisfed with their job, satisfied or very satisfied with the service provided by SYP, or with missing value for pubview, feel the working relationship with their line manager is very satisfactory and have served in the police force for 21 or more years. Their average morale is high (mean score = 2.01).

The group with the lowest general morale is as for the analysis using predictors subset 2. The subgroup contains 133 respondents who are dissatisfied with their job, or gave no response for jobsat, and agree strongly with the statement 'It's not what you know, it's who you know', and their general level of morale is low (mean score = 4.11).

CHAID Diagram 3.1c: Respondent's Own Morale against Predictor Subset 3



CHAID segment index for Diagram 3.1c

id	count	score	vars...			
-1-	132	2.07	jobsat=y	pubview=12.	linemgr=1	lenserv=1-3
-2-	100	2.37	jobsat=y	pubview=12.	linemgr=1	lenserv=4.
-3-	97	2.01	jobsat=y	pubview=12.	linemgr=1	lenserv=5
-4-	83	2.43	jobsat=y	pubview=12.	linemgr=2-.	whouknow=12. lenserv=1-3.
-5-	94	2.88	jobsat=y	pubview=12.	linemgr=2-.	whouknow=12. lenserv=45
-6-	136	2.45	jobsat=y	pubview=12.	linemgr=2-.	whouknow=3
-7-	66	2.18	jobsat=y	pubview=12.	linemgr=2-.	whouknow=45
-8-	73	3.16	jobsat=y	pubview=3	whouknow=1	
-9-	229	2.80	jobsat=y	pubview=3	whouknow=2	
-10-	191	2.55	jobsat=y	pubview=3	whouknow=3.	
-11-	86	2.31	jobsat=y	pubview=3	whouknow=45	
-12-	200	2.95	jobsat=y	pubview=45		
-13-	133	4.11	jobsat=n.	whouknow=1		
-14-	147	3.57	jobsat=n.	whouknow=2.	linemgr=12	
-15-	82	4.05	jobsat=n.	whouknow=2.	linemgr=3-.	
-16-	64	3.13	jobsat=n.	whouknow=3-5	pubview=12.	
-17-	103	3.61	jobsat=n.	whouknow=3-5	pubview=3-5	

3.2.2d: Colleagues' Perceived Morale and Predictors subset 1

Colleagues' perceived morale is associated slightly differently with the predictor variables to respondent's own, in the sense that most of the predictors are more statistically significantly related to omor (perceived public view of SYP and satisfaction with service provided by SYP are more highly significantly statistically associated with cmor). Therefore, one might expect different results from the CHAID analyses on the different measures of morale.

CHAID Diagram 3.1d shows the result of analysing colleagues' perceived morale using predictors subset 1. Initially it is useful to notice that the overall average level of colleagues' perceived morale, for 2009 respondents, is lower than the respondents' own, mean score = 3.29, compared with 2.87 before, indicating between low and neither high nor low morale in general. If we assume that the equally spaced integer scores represent the theoretical 'distance' between the levels of morale, then colleagues' perceived morale, in general, is roughly half a level lower than respondents' own. With different levels of morale for the two measures, this again begs the question of which is 'truer'? Or what is the real level of morale? More specifically, what is also unknown about the discrepancy, is whether respondents are overestimating their own morale, or underestimating that of their colleagues, if we are to assume either level of morale to be true.

Job satisfaction is the most significant predictor for the response. The split produces groups similar to the corresponding analysis using respondents' own morale, except for the minor disparity that those with missing values for jobsat are merged with those job satisfied. Conforming to the observation made above, the group job satisfied or missing perceive colleagues' morale, in general, as roughly 'half a level' below their own (mean score for colleagues' perceived morale = 3.11, for respondent's own morale = 2.56). The concept of how respondents estimate colleagues' morale is not, however, a general thing. Those not satisfied with their job estimate colleagues' morale as low in general (mean score = 3.83), not too dissimilar to that of their own (mean score for omor = 3.73). This group splits on the variable promotions are given to those who earn them, as it does with own morale as the response. Examination of the corresponding segments produced for the two analyses hints that those with more positive characteristics, tend to estimate colleagues' morale as further away from their own, in the lower direction, than those with more negative opinions. This mechanism is as if the respondents with a more positive disposition, in any respect, tend to regard their 'higher' morale as relatively rare, and therefore feel others have lower morale. Those with a more negative disposition, in general, tend to feel their morale is more typical, whilst still, on the whole, estimating colleagues' morale as lower than their own.

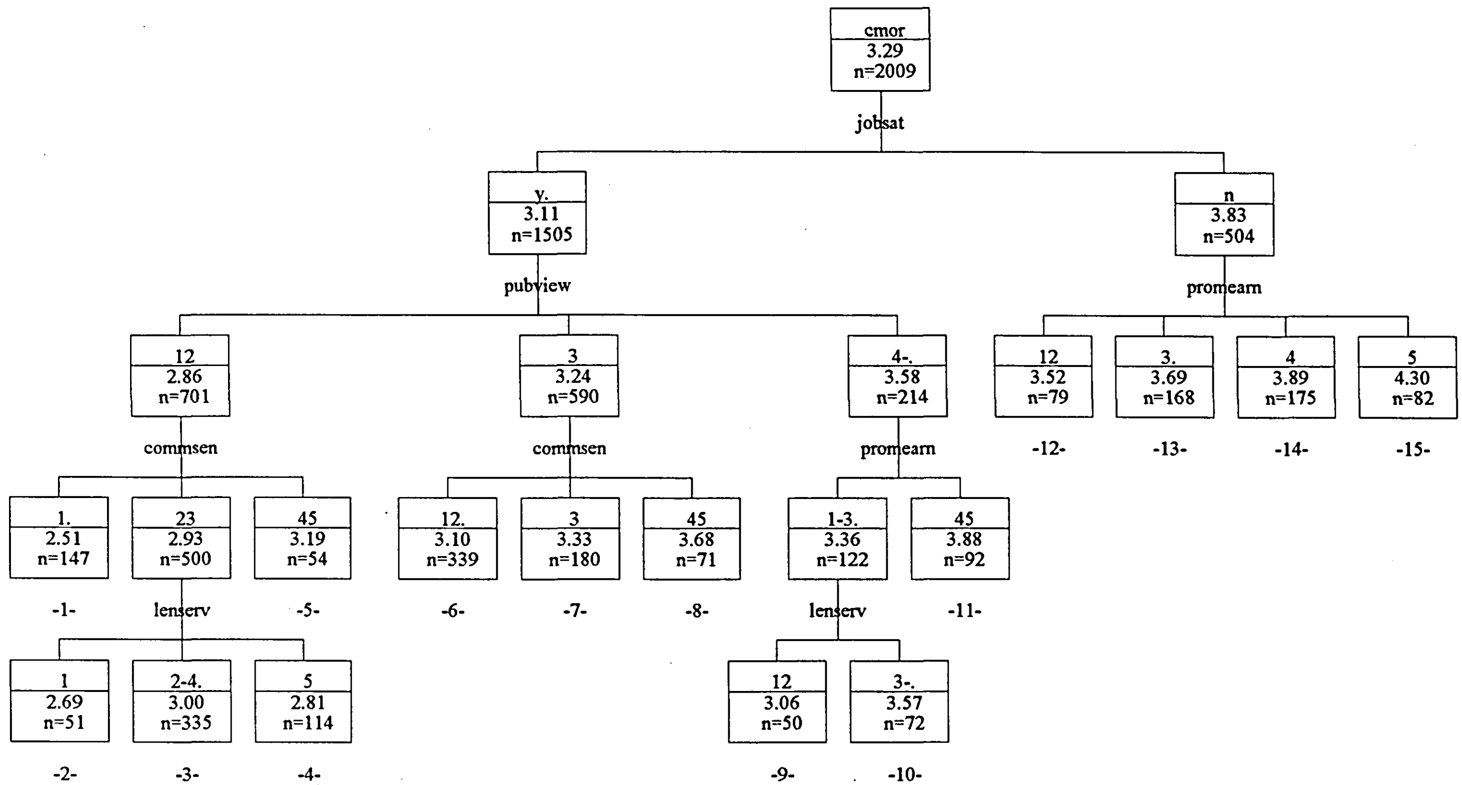
For the group satisfied with their job or with missing values for jobsat, the variable perceived public view of SYP has the most statistically significant relationship with the response. As mentioned earlier in this section, pubview is one of only two variables more significantly related to colleagues' perceived morale than respondent's own, so this result is as anticipated. The split produces good interpretational groups. The three possible directions of reply for pubview are kept distinct, i.e. those feeling the public's view of SYP is positive or very positive form one subgroup, those giving a neutral reply form another and those feeling public view is negative or very negative (plus those with missing values) form a third group.

Due to the different predictors identified using the different responses omor and cmor, there is no direct comparison to these subgroups. Examining the left hand side of both Diagrams 3.1a and 3.1d at depth 2, however, a general discrepancy of half a 'morale level' or more can be seen, from the predictors stemming from the job satisfied (plus missing for cmor) group.

At depth 3, there are 2 splits from the levels of pubview on the predictor communication with more senior managers (as there are 2 splits from different levels of commsen on pubview at depth 3 using respondent's own morale (Diagram 3.1a)). However, due to the different constituent levels of the two variables making up the subgroups, only one group of respondents is comparable across the two diagrams. Those satisfied with their job (plus missing for jobsat using cmor), feel the public's view of SYP is positive or very positive (plus missing for pubview using omor) and feel communication with more senior managers/officers is bad or very bad - segment 5 on diagram 3.1d and segment 12 on diagram 3.1a. Comparing the mean levels of colleagues' perceived morale and respondent's own respectively, mean scores 3.19 and 2.62, there is again evidence that, in general, respondents estimate their own morale roughly half a 'level' higher than that of their colleagues.

The only other noteworthy aspects of this analysis in relation to the first are, firstly, there are fewer segments defined, due to the merging of pubview, at depth 2, into 3 levels, where commsen kept 4 distinct groupings. Also, whether the respondent is civilian staff or police officer is not significant in this analysis, whilst measures of association show that there is a statistically significant difference between morale levels of civil staff or police officer. A possible reason that this factor does not feature in the analysis of colleagues' perceived morale, may be that while respondents' own morale differs between the groups, when respondents are estimating that of their colleagues, they may generalise across the sections.

CHAID Diagram 3.1d: Colleagues' Perceived Morale against Predictors Subset 1



CHAID segment index for Diagram 3.1d

id	count	score	vars...			
-1-	147	2.51	jobsat=y.	pubview=12	commsen=1.	
-2-	51	2.69	jobsat=y.	pubview=12	commsen=23	lenserv=1
-3-	335	3.00	jobsat=y.	pubview=12	commsen=23	lenserv=2-4.
-4-	114	2.81	jobsat=y.	pubview=12	commsen=23	lenserv=5
-5-	54	3.19	jobsat=y.	pubview=12	commsen=45	
-6-	339	3.10	jobsat=y.	pubview=3	commsen=12.	
-7-	180	3.33	jobsat=y.	pubview=3	commsen=3	
-8-	71	3.68	jobsat=y.	pubview=3	commsen=45	
-9-	50	3.06	jobsat=y.	pubview=4-.	promearn=1-3.lenserv=12	
-10-	72	3.57	jobsat=y.	pubview=4-.	promearn=1-3.lenserv=3-.	
-11-	92	3.88	jobsat=y.	pubview=4-.	promearn=45	
-12-	79	3.52	jobsat=n	promearn=12		
-13-	168	3.69	jobsat=n	promearn=3.		
-14-	175	3.89	jobsat=n	promearn=4		
-15-	82	4.30	jobsat=n	promearn=5		

3.2.2e: Colleagues' Perceived Morale and Predictors subset 2

This analysis is comparative to that pertaining to Diagram 3.1b.

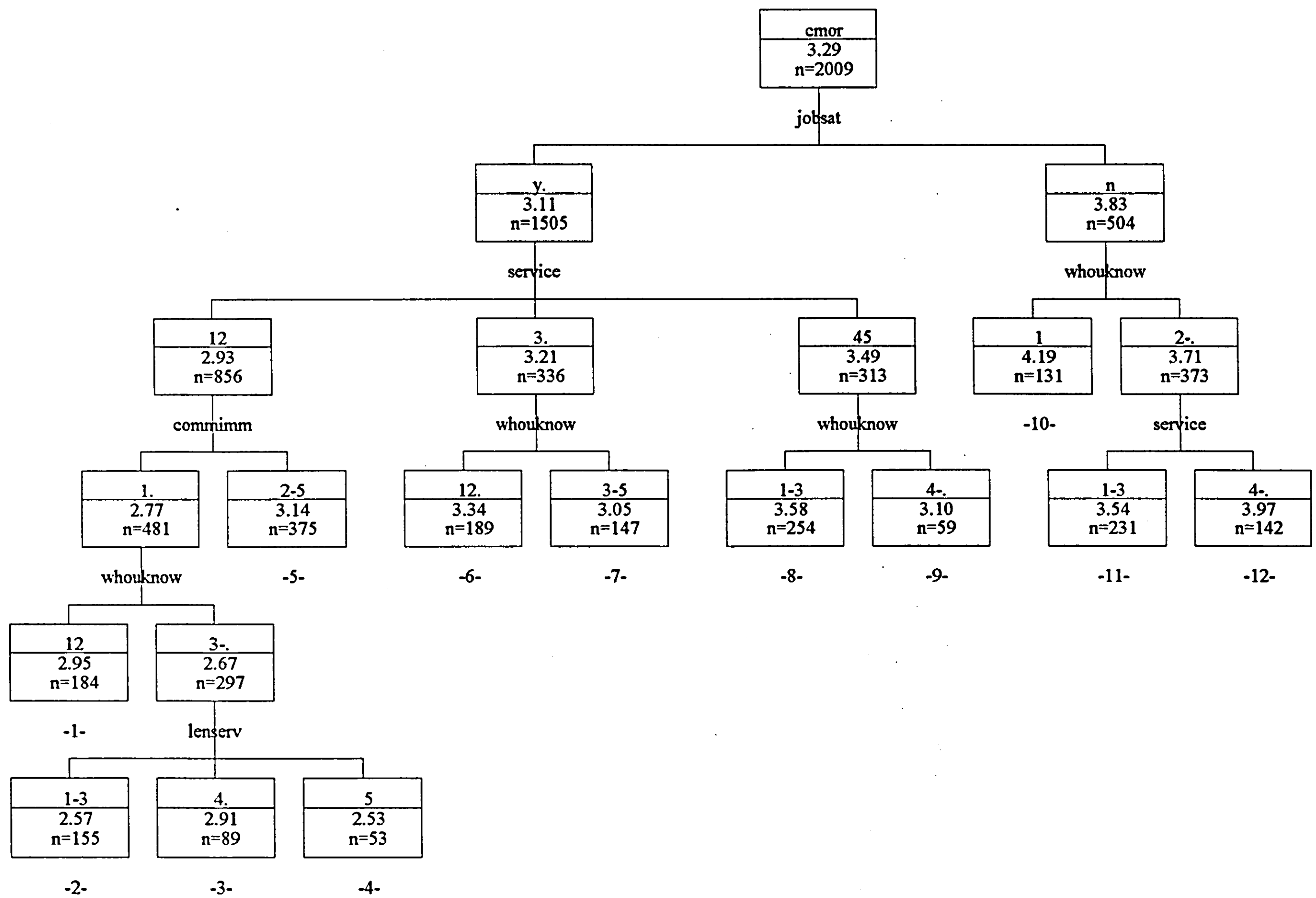
Job satisfaction is the most significant predictor, as it is for all the analyses. Similarly to the last analysis, the satisfaction of respondents with the service they are providing, is more significantly associated with colleagues' perceived morale than respondents own. This variable is the most significant predictor for the group satisfied with their job or with missing values for jobsat.

Comparing Diagram 3.1e below with Diagram 3.1b, each of the first three corresponding nodes on the left hand side show a discrepancy in mean score reflecting roughly half a 'level' of morale, consistent again with the general behaviour of respondents' estimates of colleagues' morale in relation to their own.

There are seven fewer segments in this analysis than using respondent's own morale, due to the less prominence of length of service and the absence of gender from the results. The reason for these differences could possibly be down to the fact that it seems personal characteristics, like length of service, gender, whether the respondent is civil staff or police officer etc., do not have as much relevance when estimating general morale, i.e. of colleagues, as they do when stating one's own. These characteristics do not reflect how a respondent feels about anything, whereas the opinion variables indicate more of a state of mind, which influence a person's perception of something more.

An undesirable result of this analysis is the merging of categories of the variable 'It's not what you know, it's who you know', where levels with contrary meaning are combined, these being agreeing with the statement and disagreeing and disagreeing strongly (with the neutral category, too), making the information gained more difficult to interpret (node 5, depth 2). This merge is due to the statistical similarity of the groups rather than the minimum subgroup size constraints, so to avoid this, one would have set the merge level higher until the categories were kept distinct.

CHAID Diagram 3.1e: Colleagues' Perceived Morale against Predictors Subset 2



CHAID segment index for Diagram 3.1e

id	count	score	vars...			
-1-	184	2.95	jobsat=y.	service=12	commimm=1.	whouknow=12
-2-	155	2.57	jobsat=y.	service=12	commimm=1.	whouknow=3-. lenserv=1-3
-3-	89	2.91	jobsat=y.	service=12	commimm=1.	whouknow=3-. lenserv=4.
-4-	53	2.53	jobsat=y.	service=12	commimm=1.	whouknow=3-. lenserv=5
-5-	375	3.14	jobsat=y.	service=12	commimm=2-5	
-6-	189	3.34	jobsat=y.	service=3.	whouknow=12.	
-7-	147	3.05	jobsat=y.	service=3.	whouknow=3-5	
-8-	254	3.58	jobsat=y.	service=45	whouknow=1-3	
-9-	59	3.10	jobsat=y.	service=45	whouknow=4-.	
-10-	131	4.19	jobsat=n	whouknow=1		
-11-	231	3.54	jobsat=n	whouknow=2-.	service=1-3	
-12-	142	3.97	jobsat=n	whouknow=2-.	service=4-.	

3.2.2f: Colleagues' Perceived Morale and Predictors subset 3

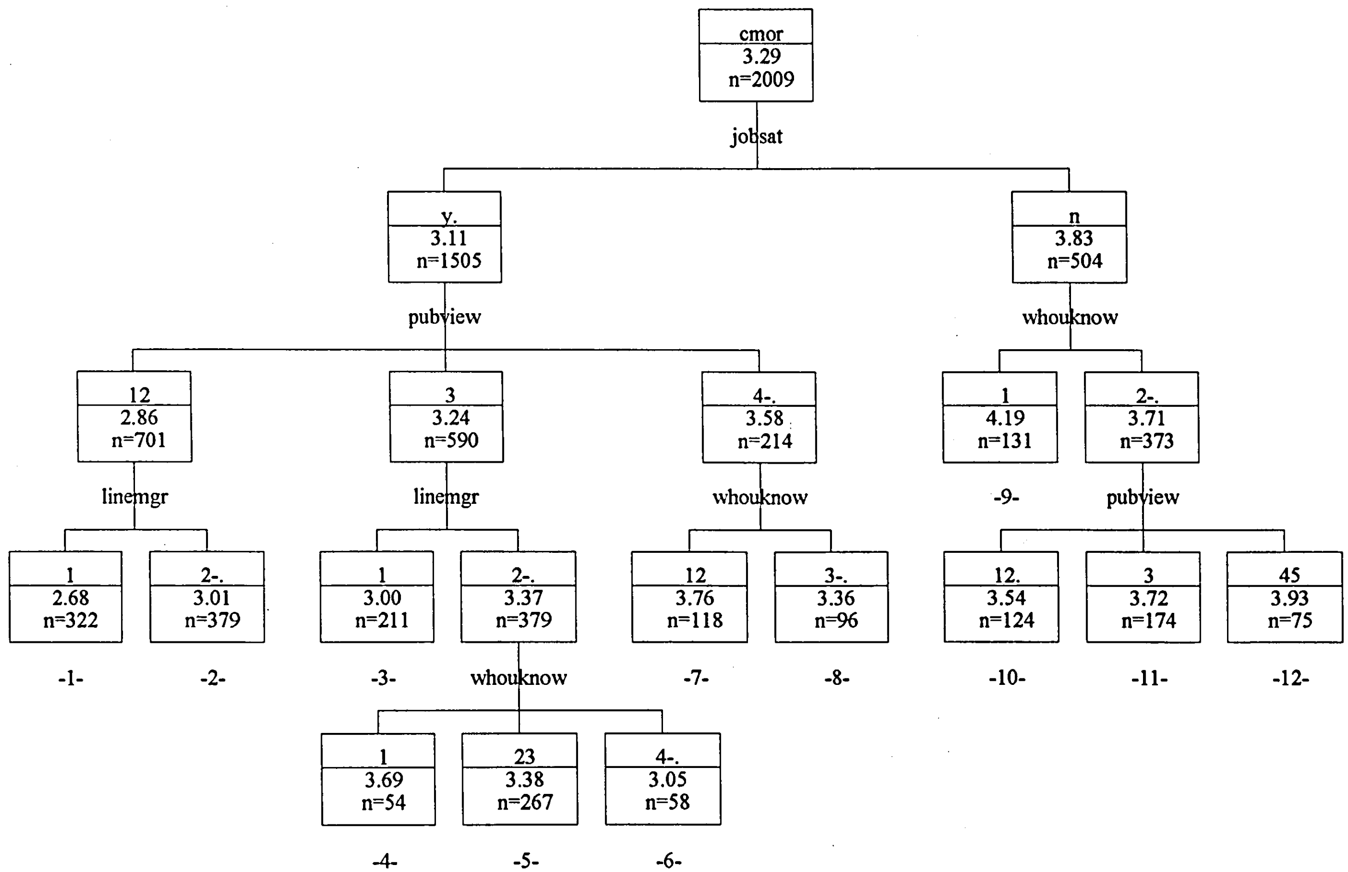
This analysis is comparative to that pertaining to Diagram 3.1c.

Job satisfaction is the most significant predictor. From the two groups produced, splits on the same variables as for respondent's own morale - perceived public view of SYP and 'It's not what you know, it's who you know' - although as for Diagram 3.1e, levels of whouknow with contradictory meanings are merged because of their statistical similarity. At depth 3, the same problem occurs with the variable working relationship with line manager, categories satisfactory to very unsatisfactory are combined together, however the merging of the contrary categories is due to the minimum subgroup size constraints, rather than statistical indistinctness.

It is again notable that length of service is absent from the results of the analysis using colleagues' perceived morale, whereas it featured using respondent's own morale. This perhaps adds weight to the thought that when estimating morale in general, personal or factual characteristics are not as relevant as opinion or state of mind characteristics.

Throughout the analyses with the three different subsets of predictors, the analyses using colleagues' perceived morale invariably create fewer segments, indicating less statistically significant association between the predictors and cmor as the response, in general. Possibly down to the notion given above that personal or factual characteristics do not carry as much weight when estimating morale. Maybe the mechanism by which a person comes to respond about colleagues' perceived morale is more subjective than responding about their own. If so, the information gained may be less accurate or true, which may reduce the statistical strength of association.

CHAID Diagram 3.1f Colleagues' Perceived Morale against Predictors Subset 3



CHAID segment index for Diagram 3.1f

id	count	score	vars...			
-1-	322	2.68	jobsat=y.	pubview=12	linemgr=1	
-2-	379	3.01	jobsat=y.	pubview=12	linemgr=2-.	
-3-	211	3.00	jobsat=y.	pubview=3	linemgr=1	
-4-	54	3.69	jobsat=y.	pubview=3	linemgr=2-.	whouknow=1
-5-	267	3.38	jobsat=y.	pubview=3	linemgr=2-.	whouknow=23
-6-	58	3.05	jobsat=y.	pubview=3	linemgr=2-.	whouknow=4-.
-7-	118	3.76	jobsat=y.	pubview=4-.	whouknow=12	
-8-	96	3.36	jobsat=y.	pubview=4-.	whouknow=3-.	
-9-	131	4.19	jobsat=n	whouknow=1		
-10-	124	3.54	jobsat=n	whouknow=2-.	pubview=12.	
-11-	174	3.72	jobsat=n	whouknow=2-.	pubview=3	
-12-	75	3.93	jobsat=n	whouknow=2-.	pubview=45	

3.2.2g: Relative Morale and Predictors subset 1

The analyses involving relative morale as the response variable are not comparable to those using respondent's own morale or colleagues' perceived morale, due to the main fact that the levels of relative morale have different interpretations to omor and cmor. There are only 3 states of relative morale compared to 5 levels of the other measures of morale, so the mean scores indicate different general states of morale. Also, the nature of the variable does not hold information about the level of morale of a respondent, only whether the respondent's own morale is higher, lower or the same as they perceive their colleagues' to be. Therefore there is no logical state of relative morale for a respondent with certain characteristics, i.e. if a respondent has 'positive' natured traits, e.g. they are satisfied with their job, there is no 'expected' level of relative morale.

Basic crosstabulations of the explanatory variables and relative morale (reported in section 3.1.5) show that there is less statistical association between the variables. This is reflected in the diagram below.

Overall, the mean relative morale score is 1.68 for 2006 individuals that gave responses for both previous measures of morale, showing that more respondents feel their own morale is higher than that of their colleagues' than vice versa, whilst the general feeling is closer to both respondent's and colleagues' perceived morale the same.

Job satisfaction is again the most significant predictor for relative morale. Those satisfied with their job (1486 respondents) have a mean relative morale score of 1.59. This reflects similar response behaviour as for the full sample, with either a slightly higher proportion feeling their morale is higher than their colleagues' or a slightly lower proportion feeling colleagues' morale is higher than their own or both. The group not satisfied with their job, or with missing values for jobsat, have an average relative morale score of 1.95, indicating, in general, that the subgroup have a tendency to feel that their own morale and that of their colleagues' is roughly the same.

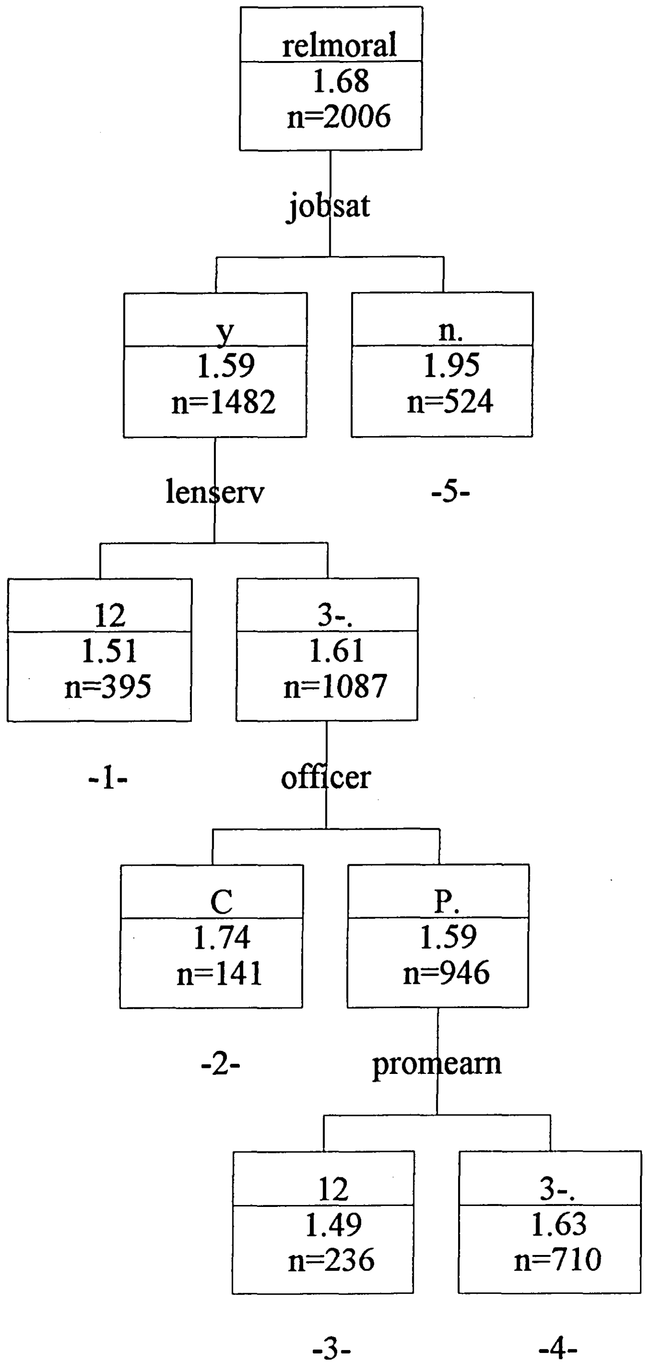
Referring to a point made in previous analyses - those with more 'positive' characteristics have a greater tendency to estimate their own morale as higher than their colleagues', than those with 'negative' traits. The relative morale of the different job satisfaction groups, depicted above, supports the observation. Due to the fact that relative morale does not take into account the magnitude of discrepancy between own morale and colleagues' perceived, it cannot help quantify the nature of the response behaviour.

The job satisfied group split on the next most significant predictor, length of service, with those serving the police for less than 2 years or between 2 and 5 years having a general level of relative morale between own higher than colleagues' and both the same (mean score = 1.51). Those serving 6 years or more have a similar level of relative morale, tending slightly more towards both own and colleagues' perceived morale the same, than the previous group (mean relative morale score = 1.61).

For the group unsatisfied with their job or with missing values for jobsat, no other explanatory variables are statistically associated with relative morale.

For the subgroup satisfied with their job, and with length of service 6 years or more, whether a respondent is an officer or civil support worker affects their relative morale differently, and for those who are police officers, agreement with the statement variable 'promotions are given to those who earn them' has a significant relationship with relative morale.

CHAID Diagram 3.1g: Relative Morale against Predictors Subset 1



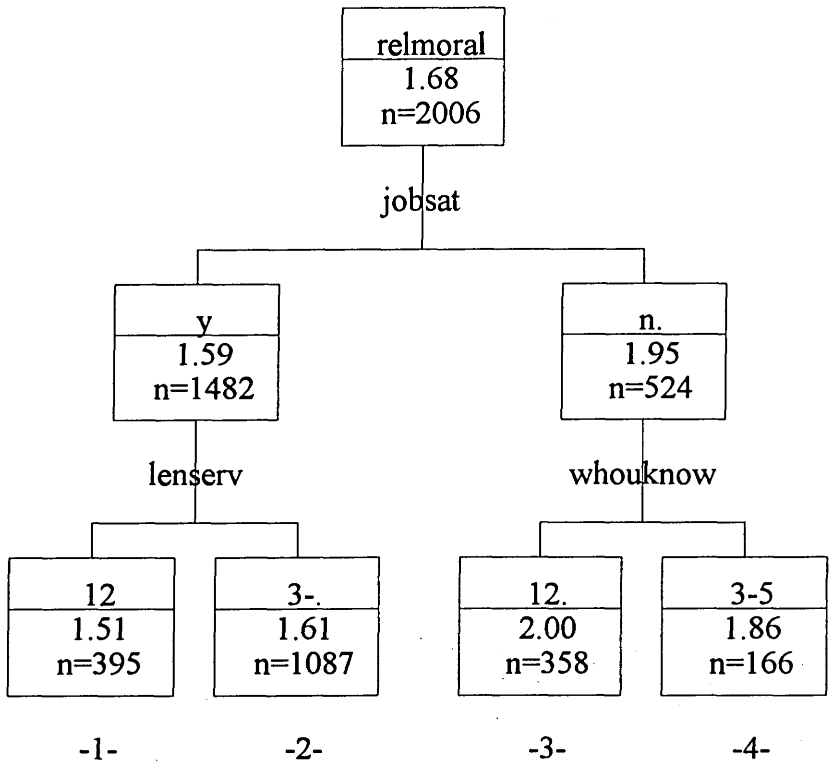
CHAID segment index for Diagram 3.1g

id	count	score	vars...			
-1-	395	1.51	jobsat=y	lenserv=12		
-2-	141	1.74	jobsat=y	lenserv=3-.	officer=C	
-3-	236	1.49	jobsat=y	lenserv=3-.	officer=P.	promearn=12
-4-	710	1.63	jobsat=y	lenserv=3-.	officer=P.	promearn=3-.
-5-	524	1.95	jobsat=n.			

3.2.2h: Relative Morale and Predictors subset 2

This analysis produces the same results as the previous one on the left hand side of the diagram down to depth 2, i.e. where respondents are satisfied with their job and that group splits on the variable length of service, and the analysis stops there. On the right hand side of the diagram, respondents are not satisfied with their job or have missing values for jobsat, this group identifies the promotion issues variable 'It's not what you know, it's who you know' as a significant predictor. Those with more negative characteristics tend to estimate their own morale as the either the same or lower than their colleagues', in general. Those who agree or agree strongly with the statement whouknow, i.e. have more negative characteristics, have a mean relative morale score of 2.00, reflecting that in general the group feel their own morale is the same as their colleagues'. For all other more positively dispositioned groups, relative morale is somewhere between own higher than colleagues and both the same.

CHAID Diagram 3.1h: Relative Morale against Predictors Subset 2



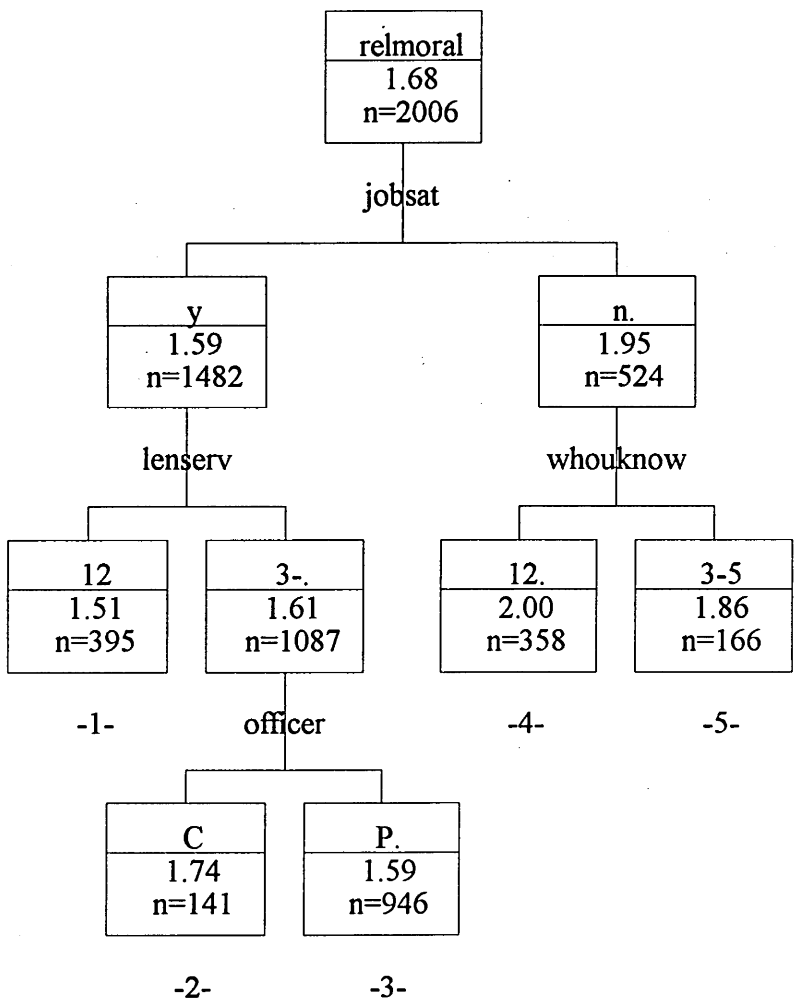
CHAID segment index for Diagram 3.1h

id	count	score	vars...	
-1-	395	1.51	jobsat=y	lenserv=12
-2-	1,087	1.61	jobsat=y	lenserv=3-.
-3-	358	2.00	jobsat=n.	whouknow=12.
-4-	166	1.86	jobsat=n.	whouknow=3-5

3.2.2i: Relative Morale and Predictors subset 3

The results of this analysis, given by Diagram 3.1i below, are basically a combination of the results of the previous two analyses down to depth 3, where the analysis stops. Therefore there is nothing to be gained by discussing diagram 3.1i, that is not already stated above.

CHAID Diagram 3.1i: Relative Morale against Predictors Subset 3



CHAID segment index for Diagram 3.1i

id	count	score	vars...		
-1-	395	1.51	jobsat=y	lenserv=12	
-2-	141	1.74	jobsat=y	lenserv=3-.	officer=C
-3-	946	1.59	jobsat=y	lenserv=3-.	officer=P.
-4-	358	2.00	jobsat=n.	whouknow=12.	
-5-	166	1.86	jobsat=n.	whouknow=3-5	

3.2.3: Summary of CHAID results

The package is most useful as a descriptive tool, and the analyses above help pinpoint specific groups with certain levels of morale that may be of interest. The characteristics of a group that has particularly high morale may be examined, and implications to improve morale in groups where it is particularly low may possibly be made. The CHAID analysis is an exploration of the data, and shows that relationships between the explanatory variables and morale are intuitive in their nature.

As an aid to modelling, implications are not necessarily instructions on what model to fit, as the relationships depicted are based on relatively small subsets of the data, and on partial associations, although any information gained will be useful and worth exploring. From the analyses above, job satisfaction is the variable with the biggest 'influence' on morale. The diagrams suggest that other variables identified may also have an effect on morale in some context. Other implications for model building are discussed for Diagram 3.1a, above. These include possible collapsing of some categories of explanatory variables, to avoid sparseness of data, and maybe helping to identify possible variables that may affect morale differently at different levels of another explanatory variable, i.e. interactions.

What CHAID does not do is give an indication of how good a fit the predictors may give in a hypothetical model. Similarly, perhaps more usefully, CHAID does not allow comparison between analyses, for example, it is impossible to say which of the above analyses is 'best', or best describes the variation within the respective morale measure.

Chapter 4: Modelling morale within the South Yorkshire Police

4.1: Defining the dependent variable, morale

The aim of modelling is to determine and quantify the effects of explanatory variables on respondents' morale.

The first step in the modelling process here is to decide what form of response, i.e. morale, is to be modelled. There are three measures of morale available, respondent's own morale (omor), colleagues' perceived morale (cmor) and relative morale (rmor) which is derived from the other two measures (see Chapter 3). The exploratory data analyses, including CHAID analyses, (Chapter 3) show that respondent's own morale is the one with the strongest statistical association with the covariates. This variable is possibly the most reliable or accurate measure of morale, and will be modelled as the dependent variable.

The response omor has 5 ordered levels (very high, high, neither high nor low, low, very low). The end categories of the variable, very high and very low morale, contain relatively few respondents, 7.7% and 7.6% respectively. Therefore in an effort to make analysis more efficient, the variable is collapsed to three categories, combining high and very high morale to form a 'higher' morale level, and low and very low morale are merged to make a 'lower' morale category. This also avoids a potential drawback of the methods. Creating a cumulative logit by dichotomising the response with 5 levels, as described for the proportional and continuation odds models, involves the collapsing of categories that have contrary definitions, i.e. the first of the simultaneously fit sub-models in both techniques opposes the category very high morale and the combined other categories, merging the level high morale with the low and very low levels of morale. The variable is assumed to be a continuum, therefore theoretically there is no problem with constructing a dichotomy of very high morale and not very high morale. However, the constituents of the latter group will have vastly differing characteristics, if the relationships expected between the explanatory variables and the response are observed. Interpreting this merged level could be difficult, and therefore collapsing the response to 3 categories avoids any unnecessary complexity. This may result in the loss of some information, but will aid interpretation of any implications of models.

The factors most likely to have an effect on morale, on the basis of exploratory analyses are deemed to be job satisfaction (jobsat), communication with more senior managers/officer (commsen), perceived public view of SYP (pubview), reaction to the statement 'promotions are given to those who earn them' (promearn), length of service (lenserv) and officer/civilian staff (officer). Therefore these variables will form the starting point for analysis using the proportional odds and continuation odds models.

4.2: The Proportional Odds model and SYP data

The independent variables are all categorical, either nominal or ordinal, however, the SAS procedure to fit the proportional odds model (LOGISTIC) has no facility for using categorical variables, so some form of recoding must be employed. Nominal variables, i.e. jobsat and officer in this case, can be represented as 0/1 binary variables. Ordinal variables (commsen, pubview, promearn, lenserv) can either be represented by dummy variables and therefore be treated as nominal, or have scores assigned to their levels and thus be treated more like continuous or interval scale variables. Either way they are not treated as ordered categorical variables as such, in the way the ordinal response is catered for, i.e. using purely the adjacency or ordering of the levels of the variable.

Initially the use of dummy variables to represent all variables is explored, and subsequently the assigning of scores to the levels of the ordinal variables using CHAID to estimate scores by maximum likelihood, and also assigning integer scores.

4.2.1: Modelling morale using dummy variables for ordinal variables

The use of dummy variables effectively turns a polytomous variable into a series of dichotomous variables. For instance a 3 category nominal variable would be represented by 2 dummy variables, pertaining to characteristic 1 or not and characteristic 2 or not, respectively, with the third level depicted by not being either of the previous 2.

The variables from the South Yorkshire Police are recoded below and interpretation should be clearer.

Jobsat is binary and therefore is represented by a single dummy variable, which is identical to simply recoding level 2 of the variable to zero:-

Jobsat level	jobnew
1 (yes)	1
2 (no)	0

The variable could be used in its original state, but is recoded to 0/1 binary in order to be consistent with the recoding of the ordinal variables. Recoded, the dummy variable is interpreted as satisfied with job, compared with not satisfied. The effect of the latter level of the variable is quantified in the intercept of a model, with the parameter pertaining to jobnew being interpreted as the additional log odds for those satisfied with their job, compared to those not, of being in a more favourable response category (proportional odds), or the most favourable response category available (continuation odds).

The variable officer can be recoded and interpreted equivalently to jobsat

Officer level	offnew
1 (civil staff)	1
2 (officer)	0

The variable communication with more senior managers/officers has 5 levels and therefore needs 4 dummy variables to represent it. Each of the first four levels has a term to show the effect of that level of the variable on morale, relative to the fifth level.

As for jobsat the effect of the last (fifth) level is explained in the intercept term(s). The last level acts as a sort of base effect, whereby the parameters pertaining to dummy variables, for the other levels of the variable, represent the additional log odds, as above, to the log odds when communication is deemed very bad (level 5).

Commsen level	com1	com2	com3	com4
1 (v. good)	1	0	0	0
2 (good)	0	1	0	0
3 (neither)	0	0	1	0
4 (bad)	0	0	0	1
5 (v. bad)	0	0	0	0

The other ordinal variables can be recoded and interpreted similarly to commsen.

Promearn level	prom1	prom2	prom3	prom4
1 (strongly agree)	1	0	0	0
2 (agree)	0	1	0	0
3 (neither)	0	0	1	0
4 (disagree)	0	0	0	1
5 (strongly disagree)	0	0	0	0

Pubview level	pub1	pub2	pub3	pub4
1 (v. positive)	1	0	0	0
2 (positive)	0	1	0	0
3 (neither)	0	0	1	0
4 (negative)	0	0	0	1
5 (v. negative)	0	0	0	0

Lenserv level	len1	len2	len3	len4
1 (<2 yrs)	1	0	0	0
2 (2-5 yrs)	0	1	0	0
3 (6-10 yrs)	0	0	1	0
4 (11-20 yrs)	0	0	0	1
5 (21+ yrs)	0	0	0	0

The initial model contains only main effects terms, depicted as above in dummy variable form :-

$$\log [\gamma_j(\mathbf{x}_i)/\{1-\gamma_j(\mathbf{x}_i)\}] = \alpha_j + \beta_{job}(jobnew) + \beta_{c1}(com1) + \beta_{c2}(com2) + \beta_{c3}(com3) + \beta_{c4}(com4) + \beta_{pu1}(pub1) + \beta_{pu2}(pub2) + \beta_{p3}(pub3) + \beta_{p4}(pub4) + \beta_{pr1}(prom1) + \beta_{pr2}(prom2) + \beta_{pr3}(prom3) + \beta_{pr4}(prom4) + \beta_{l1}(len1) + \beta_{l2}(len2) + \beta_{l3}(len3) + \beta_{l4}(len4) + \beta_o(offnew) \qquad (4.1)$$

for j=1,2. Where $\gamma_j(\mathbf{x}_i)/\{1-\gamma_j(\mathbf{x}_i)\} = P(Y_i \leq j)/[1-P(Y_i \leq j)]$.

(The number of valid observations, i.e. those without missing values for any of the variables involved in the analysis is 1837.)

Using SAS Proc LOGISTIC to fit the model, the proportional odds assumption is violated for the above model. The score test statistic given :-

$\chi^2 = 53.39$ with 18 df (p=0.0001)

Therefore the assumption, that the log odds, for individual covariates, for the dichotomies of the morale scale (having 'higher' morale as opposed to 'neither high nor low' or 'lower', and having 'higher' or 'neither high nor low' morale as opposed to 'lower'), are equivalent is not satisfied, given the explanatory variables in the model.

Applying or interpreting the model is of little use as the violation of the proportional odds assumption renders the model invalid.

To ascertain why the model is invalid, examining the parameter estimates may give insight, or provoke ideas for further analysis or modification. The parameters estimated for the above model are given below :-

Table 4.1: Parameter estimates, their standard errors and p-values, for model 4.1

Parameter	Estimate	s.e.	p-value
$\alpha 1$	-6.25	1.056	0.0001
$\alpha 2$	-4.44	1.051	0.0001
βjob	1.94	0.122	0.0001
$\beta c1$	1.4	0.336	0.0001
$\beta c2$	0.91	0.318	0.0043
$\beta c3$	0.54	0.319	0.0926*
$\beta c4$	0.08	0.334	0.8213*
$\beta pu1$	4.51	1.265	0.0004
$\beta pu2$	3.46	1.017	0.0007
$\beta pu3$	2.79	1.016	0.0061
$\beta pu4$	2.42	1.020	0.0178
$\beta pr1$	0.92	0.404	0.0226
$\beta pr2$	0.74	0.194	0.0001
$\beta pr3$	0.39	0.181	0.0322
$\beta pr4$	0.03	0.182	0.8697*
$\beta l1$	1.53	0.215	0.0001
$\beta l2$	0.51	0.168	0.0024
$\beta l3$	0.19	0.151	0.2125*
$\beta l4$	0.08	0.132	0.5455*
βo	-0.59	0.130	0.0001

* denotes parameter not significant

From the table above, it is evident that all the explanatory variables make a significant contribution to the fit of the model. Even though not all the parameter estimates are significant, at least one dummy variable pertaining to each covariate is significant ($p<0.05$). The significant dummy variable causes a significant decrease in deviance, due to the inclusion of each explanatory variable, therefore all dummy variables must be included.

Where a dummy variable parameter is not significant, it implies that the effect of the respective level is not different from the effect of the last level, of the explanatory variable. Parameters for levels 3 and 4 of commsen, 4 of promearn, and 3 and 4 of lenserv are not significant, indicating that the individual parameters for these levels are not contributing statistically to the model. From this information, it may be possible to collapse some levels of the variables, as the CHAID analyses also suggest. In the case of the variables commsen, pubview and promearn listed above, it is feasible to collapse the extreme categories. For example, for commsen one could logically merge the category very good with good and similarly very bad with bad, as the response variable respondent's own morale was recoded. For length of service the merging of any adjacent levels is feasible. The collapsing of some of the categories of the independent variables reduces the number of variables in the model when using dummy variables. This may improve the validity of the proportional odds assumption if, say, a parameter causing, in whole or part, the violation of the assumption is replaced by a term representing a dummy variable for the new merged level of the variable, though this is not necessarily the case. Also, from sections 3.1.2 to 3.1.5, it can be seen that the proportions of respondents answering in the extreme ends of the 5 category variables are notably smaller than the numbers in the other categories - commsen has only 3% of respondents in the very bad category, the very positive and very negative levels of pubview contain only 0.8% and 0.9% of respondents respectively, and only 2% of respondents strongly agree with the promearn statement. This sparseness of data also provides some incentive for reducing the number of levels of the variables. Therefore, the variables commsen, pubview and promearn are collapsed to 3 levels, and the variable lenserv collapsed to binary, in an attempt to avoid including redundant parameters in the model, and sparse groups of data.

The variables are recoded and dummy variables assigned as follows :-

Commsen level	newcom	com1	com2
1 & 2	1	1	0
3	2	0	1
4 & 5	3	0	0

Pubview level	newpub	pub1	pub2
1 & 2	1	1	0
3	2	0	1
4 & 5	3	0	0

Promearn level	newprom	prom1	prom2
1 & 2	1	1	0
3	2	0	1
4 & 5	3	0	0

Lenserv level	newlen	lendum
1	1	1
2 - 5	2	0

Note that dummy variables for jobsat and officer are unchanged.

Therefore the new main effects model is given below :-

$$\log [\gamma_j(\mathbf{x}_i)/\{1-\gamma_j(\mathbf{x}_i)\}] = \alpha_j + \beta_{job}(\text{jobnew}) + \beta_{c1}(\text{com1}) + \beta_{c2}(\text{com2}) + \beta_{pu1}(\text{pub1}) + \beta_{pu2}(\text{pub2}) + \beta_{pr1}(\text{prom1}) + \beta_{pr2}(\text{prom2}) + \beta_l(\text{lendum}) + \beta_o(\text{offnew}) \quad (4.2)$$

The proportional odds assumption does not benefit from category reduction, the score test statistic, $\chi^2 = 44.74$ with 9 df ($p=0.0001$). Therefore this model, too, is invalid.

In order to determine why the assumption of proportional odds is not accepted, the corresponding binary logistic models for the binary splits of respondent's own morale must be fitted. The response is dichotomised to form a logit of higher morale versus neither high nor low and lower for model 4.3a, and a logit of higher and neither high nor low morale versus lower for model 4.3b, whilst the explanatory covariate specifications stay the same.

The parameter estimates and their standard errors and p-values are given below in table 4.2:-

Table 4.2: Parameter estimates, their standard errors and p-values for binary logistic models 4.3a and 4.3b

	Model 4.3a			Model 4.3b		
Parameter	Estimate	s.e.	p-value	Estimate	s.e.	p-value
α	-3.30	0.262	0.0001	-2.42	0.220	0.0001
βjob	2.16	0.181	0.0001	1.96	0.133	0.0001
βc1	0.81	0.179	0.0001	0.99	0.169	0.0001
βc2	0.22	0.193	0.2474*	0.66	0.176	0.0002
βpu1	0.85	0.177	0.0001	1.33	0.182	0.0001
βpu2	0.10	0.178	0.5767*	0.76	0.170	0.0001
βpr1	0.65	0.142	0.0001	0.90	0.172	0.0001
βpr2	0.20	0.130	0.1206*	0.69	0.142	0.0001
βl	1.51	0.207	0.0001	1.37	0.292	0.0001
βo	-0.78	0.155	0.0001	-0.09	0.156	0.5286*

* denotes parameter not significant

Comparing the parameter estimates for the two models, there are some obvious areas where the discrepancy between corresponding estimates is, perhaps, too large for a single parameter in the proportional odds model to be adequate.

The instances where the parameter from one model is significant, whilst the corresponding parameter from the other model is not, are the probable cause of the violation of the proportional odds assumption. For example, the parameter estimates for dummy variables com2, pub2, prom2 and offnew differ noticeably between models.

The realistic possibilities in this situation are fairly limited, but among the options are the following.

The explanatory variables commsen, pubview and promearn could be collapsed further to binary to see if this improves the proportional odds assumption. This poses a problem in terms of logically how should the variables be dichotomised? With which level should the neutral category be combined? There is no obvious choice to this dilemma so this possibility is perhaps not the most desirable.

One or more of the 'offending' variables could be omitted and the proportional odds model refitted, to see if this satisfies proportional odds assumption. This is not desirable, as all the variables are seen to contribute significantly to the proportional odds model, and omitting information, to compensate the proportional odds assumption, would be at the expense of some level of goodness-of-fit of the model. Therefore this issue, i.e. what is more important? - the fit of the model or the proportional odds assumption - must be decided.

The ordinal independent variables in the model could instead be used in a different form, e.g. with scored categories instead of as dummy variables. This option is explored in section 4.3, as an alternative approach, and therefore will be not be employed to solve the current problem.

At this stage, if any of the above options were not feasible, one may have to settle for the results from the separate binary logistic models. One would examine the goodness-of-fit of these models, and depending on their adequacy, interpret and use the results.

If the proportional odds model is not appropriate or adequate, one could try applying a different model for an ordinal response, such as the continuation odds model. The continuation odds model is applied to the South Yorkshire Police data in section 4.5.

In this instance, the variable officer has two distinct levels, and especially as it is a personal characteristic, there is the option of splitting the dataset into officers and civilian staff and treat them separately. This has the attraction of removing the variable officer from the model, which appears to contribute heavily to the violation of the proportional odds assumption, without losing any information. Therefore the proportional odds model is refitted on the two new subsets of the data, model 4.4o for officers and model 4.4c for civil staff, as specified for model 4.2, obviously excluding the variable officer (parameter offnew).

Of the 1837 valid respondents, 1438 are police officers and 399 are civilian staff.

The model fitted to the officers data does not satisfy the proportional odds assumption, the score test statistic χ^2 = 36.35 with 8 df (p=0.0001) (see Appendix 4a). The addition of two, three and four way interaction terms between independent variables failed to improve the model, therefore model 4.4o is discarded. The proportional odds model using the variable information in a different form, using scored explanatory variables for the ordinal variables, instead of dummy variables, is attempted in section 4.3. The continuation odds model for dummy variables and scored ordinal variables is fitted to the data in section 4.5, therefore it is not necessary to proceed with, or interpret, a proportional odds model for officers only using dummy variables for ordinal variables, as alternatives are explored.

The model fitted to the data including civilian staff only, 4.4c, however, does satisfy the assumption of proportional odds. The variable communication with more senior managers/officers is found to be non-significant in the model, i.e. the variable has no statistically significant linear association with respondent's own morale for civil staff respondents. Therefore the parameters pertaining to commsen should be removed from the model and the model refitted, as these parameters are not contributing to the fit of the model to the data. The parameter for prom2 is also non-significant, but the parameter for prom1 is significant so prom2 must stay in the model.

Communication with more senior officers/managers may be something that is less applicable to civil staff. This group contains traffic wardens and hourly paid members of staff and so communication with more senior officers is possibly non-existent. Not to suggest these peoples' jobs are less important than police officers, but the consequences of their work are probably less severe. Depending on the structure of the work of the- civilian staff, less communication with more senior managers may be required. With this in mind, the exclusion of the variable commsen from the model is not as surprising as it may first appear.

The model is refitted excluding the dummy variables for commsen (com1, com2). The new model is referred to as model 4.5c. Model 4.5c comfortably satisfies the proportional odds assumption, score test statistic, $\chi^2 = 3.706$ with 6 df (p=0.7164).

These results confirm that the effect of the explanatory variables on respondent's own morale is different for officers and civil staff, and therefore treating the two groups of respondents separately in this case is justified.

The fundamental assumption of the model is satisfied, therefore the parameters, and goodness-of-fit of the model must be examined.

The decrease in deviance from the intercept only model is $\chi^2 = 168.72$ with 6 df which is highly significant, (-2 Log (L) deviance for model with intercept only = 874.495, with covariates, deviance = 705.775). Compared with the corresponding statistic for model 4.4c, 170.52 with 8 df, there is a difference of 2.2 with 2 df, thus confirming that the loss of information by excluding commsen is not statistically significant. Note that this decrease in deviance would not be comparable to that of model 4.1 or 4.2, as these models are applied, essentially, to different data, and this measure is only useful when comparing models for the same data.

The parameters for model 4.5c are given below :-

Table 4.3: Parameter estimates, their standard errors and p-values, for model 4.5c

Parameter	Estimate	s.e.	p-value
$\alpha 1$	-3.78	0.34	0.0001
$\alpha 2$	-1.61	0.28	0.0001
βjob	1.98	0.23	0.0001
$\beta pu1$	1.44	0.29	0.0001
$\beta pu2$	0.59	0.25	0.0192
$\beta pr1$	0.69	0.26	0.0073
$\beta pr2$	0.33	0.24	0.1700*
βl	1.26	0.25	0.0001

* denotes parameter not significant

All explanatory variables included in the model have significant parameters. The interpretation of the effects of the independent variables make intuitive sense. For example, the estimate value of the parameter β_{job}, is the increase in log odds of a 'more favourable' response for those satisfied with their job over those not, ceteris paribus. In this case the parameter estimate is 1.98, indicating that if someone is satisfied with their job, the odds of them having 'more favourable' morale is increased by roughly 7 times ($\exp\{1.98\} = 7.24$), compared to someone not satisfied with their job, if all other variable information is equal. This is as we would logically expect, at least in terms of the nature of the association, if not the magnitude. Of the variables with 3 categories, the parameter estimates are the increase in log odds of a 'more favourable' response for those in the respective category 1 and 2 of the explanatory variable compared to those in category 3, all else the same. For example, we would expect an individual who feels the public view of SYP is positive, to have greater odds of a 'more favourable' level of morale than those who feel the public view of SYP is negative. According to the model this is true, as $\beta_{pu1} = 1.44$, indicating an increase in odds of approximately 4 times ($\exp\{1.44\} = 4.22$). The subsequent dummy parameters for pubview and the other ordinal variables can be interpreted similarly.

The goodness-of-fit of model 4.5c must be assessed, i.e. how well it describes the data patterns. To do this the method introduced by Lipsitz, Fitzmaurice and Molenberghs (1996) can be used, as described in section 2.5.2 and 2.5.3.

The first step of the Lipsitz et al method involves assigning scores to the levels of the ordinal response. In this case integer scores are used for simplicity, and the assumption that the inter-category distances between higher and neither high nor low, and lower and neither high nor low should be fairly similar, so as not to infer that the neutral group are more like one of the non-neutral response groups than the other. Therefore the response is coded as :-

Level of morale	Score
Higher	1
Neither high nor low	2
Lower	3

In order to construct a predicted mean score, the individual response probabilities are calculated (given by SAS) and multiplied by the respective response category scores, so the predicted mean score, μ, for an individual i would be :-

$$\mu_i = 1(p_1) + 2(p_2) + 3(p_3)$$

where p_1, p_2 and p_3 are the probabilities of responding in categories 1, 2 and 3 respectively of respondent's own morale, for each individual, as estimated by the model.

The data is sorted in ascending order of the predicted mean scores and partitioned into approximately equally sized percentile groups. In this analysis the number of groups used is g = 6, as using more than 6 groups leads to more than 20% of expected counts in each response level within each group being less than 5, for which the χ^2 approximation to the data may be poor. Using fewer than 6 groups will give a test statistic with fairly low power, which may give misleading results (see section 2.5.2). Therefore the 399 observations are divided into groups as indicated below, where the first 57 observations have the lowest predicted mean scores and the last 64 have the highest, the reason for using unequally weighted groups is so as not to separate observations with tied predicted mean scores (see section 2.5.3). Once the data is partitioned, the g-1 group indicators are defined as :-

$I_{ig} = 1$ if μ_i is in group g,

$I_{ig} = 0$ otherwise,

for g = 1, ..., 5. These indicators are constructed to act as dummy variables in the model, and their assignation to the data is as follows :-

obs no.	n_g	group, g	I_1	I_2	I_3	I_4	I_5
1 - 57	57	1	1	0	0	0	0
58 - 130	73	2	0	1	0	0	0
131 - 189	59	3	0	0	1	0	0
190 - 262	73	4	0	0	0	1	0
263 - 335	73	5	0	0	0	0	1
336 - 399	64	6	0	0	0	0	0

The model 4.5c is now refitted with these group indicator dummy variables, and for the model to fit the data adequately, the parameter estimates for these group indicators should not be statistically significant, i.e. we hope for a p-value of > 0.05.

Model 4.6c is therefore given as :-

$$\log [\gamma_j(\mathbf{x}_i)/\{1-\gamma_j(\mathbf{x}_i)\}] = \alpha_j + \beta_{job}(\text{jobnew}) + \beta_{pu1}(\text{pub1}) + \beta_{pu2}(\text{pub2}) + \beta_{pr1}(\text{prom1}) + \beta_{pr2}(\text{prom2}) + \beta_l(\text{lendum}) + \gamma_1(I_1) + \gamma_2(I_2) + \gamma_3(I_3) + \gamma_4(I_4) + \gamma_5(I_5) \qquad (4.6c)$$

The proportional odds assumption is intact with the addition of the extra parameters, $\chi^2 = 12.59$ with 11df. However, when testing the goodness-of-fit using Lipsitz et al's method, the form of the model, i.e. proportional odds model, is assumed to be correct before the goodness-of-fit is assessed, therefore it is the contribution of the parameter estimates for the group indicators that are of primary concern.

The difference in deviance for model 4.6.c from the intercept only model is χ^2 = 178.70 with 11 df (for intercept only, -2 Log (L) = 874.495, deviance with covariates + g-o-f = 695.795). Therefore the decrease in deviance for model 4.6c over model 4.5c is χ^2 =9.98 with 5 df (p>0.05), indicating that the parameters pertaining to the group indicators do not make a significant contribution to the fit of the model. Model 4.5c accounts for a large enough proportion of the variation within the data, so that the extra group indicator parameters are not required. From the evidence of the Lipsitz et al goodness-of-fit test the model fits the data adequately.

The parameter estimates for model 4.6c are given below :-

Table 4.4: Parameter estimates, their standard errors and p-values, for model 4.6c

Parameter	Estimate	s.e.	p-value
$\alpha 1$	-3.33	0.40	0.0001
$\alpha 2$	-1.08	0.37	0.0033
βjob	0.67	0.73	0.3549*
βpu1	0.39	0.59	0.5124*
βpu2	0.16	0.36	0.6510*
βpr1	0.25	0.38	0.5174*
βpr2	-0.02	0.34	0.9492*
βl	0.60	0.56	0.2815*
$\gamma 1$	2.36	1.66	0.1569*
$\gamma 2$	2.46	1.28	0.0556*
$\gamma 3$	1.64	1.06	0.1202*
$\gamma 4$	1.08	0.84	0.1961*
$\gamma 5$	0.17	0.56	0.7663*

* denotes parameter not significant

The parameters $\gamma 1$ to $\gamma 5$ are not statistically significant in the model, although $\gamma 2$ is borderline significant. Therefore providing evidence that the model 4.5c gives an adequate fit to the data. The group indicator variables have a confounding effect on the other parameters as there is likely to be correlation between the independent variables and the group indicator variables as both sets are trying to explain the same pattern in the data. For example, the parameters for jobnew, pub1 and prom1 are trying to account for higher morale response, as that is what we expect for individual with the characteristics job satisfied, feel public's view of SYP is positive, and agree that promotions are given to those who earn them. Parameter $\gamma 1$, and possibly $\gamma 2$, are doing the same, as they are coded to those observations with the lowest predicted mean scores, i.e. higher expected morale. If the group indicator parameters are not significant, then the explanatory variable parameters in the model are explaining the data patterns sufficiently. The group indicators are systematically assigned after the initial model, and therefore estimates of their effect will fit the data. If the parameters for the independent variables are not fitting the data adequately, the group indicators will be accounting for a large proportion of the variation within the data.

To further assess the fit of the model, in terms of perhaps where the model doesn't fit well, observed and expected frequencies for the response levels within each of the 6 partitioned groups can be calculated, and thus approximate standardised residuals computed for the resultant 18 'cells' (6 groups, 3 response levels). The observed frequencies are simply the counts of observations in each morale category for each group, calculated by :-

$$O_{gk} = \sum_{i=1}^{n} I_{ig} Y_{ik} \quad \text{for } g = 1, \ldots, 6, \text{ and } k = 1, 2, 3.$$

This involves the simple addition of the group indicator I_6, which can be coded as 1 where I_1 to $I_5 = 0$, and 0 otherwise, in order for the expression to be true for all observed counts. The same applies to expected counts, as calculated below :-

$$E_{gk} = \sum_{i=1}^{n} I_{ig} p_{ik} \quad \text{for } g = 1, \ldots, 6, \text{ and } k = 1, 2, 3.$$

The expected counts are simply the sums of the respective individual response probabilities, for the individuals within each group.

The approximate residuals, R_{gk}* and adjusted approximate residuals, R_{gk}** are the differences between the observed and expected values, standardised by an error term (see section 2.5.2).

For example, the observed count for response 1, higher morale, in group 1 is 37, whilst the expected count is 39.99. The approximate standardised residual for this set of respondents is :-

$$R_{11}^* = -2.99 / \sqrt{57(0.7015)(1-0.7015)} = -0.865$$

The adjusted approximate residuals, R_{gk}^{**}, are the R_{gk}^*'s divided by an estimate of their 'common' standard deviation, $\sigma = \sqrt{(\Sigma R_{gk}^{*2}/GK)} = \sqrt{(24.909/18)} = 1.176$, so that :-

$$R_{11}^{**} = -0.865/1.176 = -0.74$$

Therefore all approximate residuals are scaled down by this measure.

Table 4.5 contains observed counts, expected counts, approximate residuals and adjusted approximate residuals for each group and response level, as computed from the results of fitting model 4.5c :-

Table 4.5: Observed and Expected values with standardised residuals for model 4.5c

		Morale		
G		Higher	Neither	Lower
1	O	37	19	1
	E	39.99	14.26	2.75
	R*	-0.87	1.45	-1.08
	R**	-0.74	1.23	-0.92
2	O	43	24	6
	E	36.92	28.57	7.50
	R*	1.42	-1.10	-0.58
	R**	1.21	-0.94	-0.49
3	O	17	35	7
	E	19.07	28.31	11.62
	R*	-0.58	1.74	-1.51
	R**	-0.49	1.48	-1.28

		Morale		
G		Higher	Neither	Lower
4	O	16	32	25
	E	15.06	35.49	22.44
	R*	0.27	-0.82	0.65
	R**	0.23	-0.69	0.55
5	O	8	19	46
	E	7.22	27.40	38.38
	R*	0.31	-2.03*	1.78
	R**	0.26	-1.73	1.51
6	O	1	17	46
	E	2.19	12.84	48.97
	R*	-0.82	1.30	-0.88
	R**	-0.70	1.11	-0.75

* denotes significantly large residual

The adjusted approximate residuals R** show little or no inadequacy of fit from the model, however, each of these residuals is standardised and scaled down within the context of the general magnitude of the approximate residuals, R*, as a whole. For the

R*'s, there is not much evidence against the model. However, in group 5, the discrepancy between the observed and expected counts in response 2, neither high nor low morale, is significantly large, but when adjusted for the general size of the approximate residuals, the magnitude of the residual is decreased to non-significance. Half of the R* standardised residuals are greater than 1, although as stated before, only 1 is significantly large. Of the 18 residuals, 10 are negative and 8 are positive, and there seems to be no pattern to whether a residual is positive or negative. The evidence of the residual analysis above gives no indication that the model is not an adequate fit to the data. It must be noted that the data, 399 observations are aggregated into a fairly small table of 18 cells, whereas there are 108 different combinations of values for each of the independent variables and the response (jobsat (yes,no), pubview (positive, neither, negative), promearn (agree, neither, disagree) lenserv, (<2 yrs, 2 yrs+), and morale (high, neither, low)), thus the loss of some information is inevitable, which may or may not make a difference in the results of a test for the goodness-of-fit of a model.

To illustrate the performance of the model, a useful exercise is to plot the predicted mean scores, as calculated for the Lipsitz et al goodness-of-fit test, against the observed responses, this is done below for model 4.5c :-

Figure 4.1: Plot of Predicted mean scores vs Observed responses for model 4.5c

```
  Legend: A = 1 obs, B = 2 obs, etc.

       3 |          A           C  AA A  D     A B K  JCAKAAD B GF    NBW F O
Observed |
Response |
         |
         |
         |
         |
       2 |          C C  AF  FH   BH AE Q  A K F S  I DGA B A BC    CCJ C A
         |
         |
         |
         |
         |
       1 |          F J   J  KN CEF CLBE     F D J  BBBD  A         C   A
         ---------------------------------------------------------------------
           1.00     1.25     1.50     1.75     2.00     2.25     2.50     2.75     3.00

                                   Predicted Mean Score
```

For a model that fits the data really well, we would expect a large concentration of points along the ascending diagonal (the diagonal being a line where observed response is equal to predicted mean score, through the origin) to denote that the predicted mean scores are close to the observed responses in general. This is not quite the case for model 4.5c, although in figure 4.1, above, it can be seen that there is a trend of increasing predicted mean scores as the observed response increases in general. The spread of the points is fairly large along all three response levels, although for morale levels 2 and 3 (neither high nor low and low respectively) there is a concentration of points around the areas on the graph where observed response and predicted mean score are equal, denoted by the latter letters of the alphabet. For higher morale, it seems the predicted mean scores are not as accurate, with the majority between 1.25 and 1.5. The plot does not specifically suggest that model fits well, but does not disprove the other diagnostic measures. Other measures suggest the fit of the model may be adequate enough.

Models involving interactions were explored to examine if the models given above may be improved. One model in particular made an improvement to the fit of the model statistically, producing satisfactory results for all the above criteria. However, the main effects model was preferred ultimately. The improved model includes an main effects terms as for model 4.4c with interaction term for the variables commsen and promearn (referred to as model 4.41c - see Appendix 4a), meaning that the effect on morale, of how a respondent feels about communication with more senior officers/managers, is dependent on their level of agreement with the statement 'promotions are given to those who earn them'. From model 4.4c, it can be seen that commsen has no direct statistical association with morale for this subset of the data, the civil support staff only. Examination of the basic crosstabulation for commsen (collapsed to 3 categories) versus respondent's own morale, omor (3 levels) controlling for promearn (3 categories) (see Appendix 2), shows that when promearn = 1, i.e. the respondent agrees with the statement, there is still no statistical association between the variables commsen and omor as whole variables. However, the proportion of respondents feeling communication is bad, who have higher morale, is found to be greater than the proportion who feel communication is good, who have higher morale. The former group is based on small numbers which seem spurious. The point of describing this behaviour of the data, is that the model containing the interaction between commsen and promearn, implies that for two respondents, who both agree with the promearn statement, if one feels communication is bad compared to the other who feels communication is good, the individual who feels communication is bad is more likely to have higher morale than the individual who feels communication is good. The model is not wrong in it's specification, it is describing the data pattern that exists. However, it is assumed that this data pattern is an unusual occurrence, and emanates from the lack of respondents with the particular characteristics of feeling communication is good, and agreeing with the statement 'promotions are given to those who earn them'. Therefore, the main effects model is preferred to the interaction model, to avoid unnecessary complexity, with respect to interpretation.

With the adequacy of fit of model 4.5c tested and accepted, the implications of the model must be examined. Refer to table 4.3 for the parameter estimates (log odds ratios) for the explanatory variables.

The intercept terms give the baseline odds of higher morale versus neither high nor low or lower (α_1), and higher or neither high nor low morale versus lower (α_2). The baseline odds, as we are using dummy variables, contain the effects of an individual having the following characteristics :-

not satisfied with their job
feel the public's view of SYP is negative
disagree that promotions are given to those who earn them
have served the police for more than 2 years

An individual with the above characteristics is roughly 44 times more likely to have neither high nor low or lower morale than higher morale (-3.78 is the log odds of higher morale, therefore $\exp\{-3.78\} = 0.023$ is the odds of higher morale, thus $1/0.023$ is the odds of neither high nor low or lower morale). Similarly, the same individual is approximately 5 times more likely to have lower morale than neither high nor low or higher.

The proportional odds model allows us to make statements about the effects of explanatory factors, regardless of level of morale. The explanatory variable parameters are estimates of the log odds of a 'more favourable' response, defined as higher morale versus neither or lower, or higher or neither versus lower morale.

To illustrate the numerical interpretation of the parameters, for two individuals with identical characteristics except that the first says he/she is satisfied with their job, whilst the second is not, the odds of the first individual having more favourable morale are 7.24. The first respondent is over 7 times more likely to have a more favourable level of morale.

The implication made above is fairly obvious, in order to try and increase morale within the civil support staff of the South Yorkshire Police, job satisfaction must be promoted. This could come in the way of pay incentives or bonuses, or greater variety of tasks for example. The feasibility of either of these options depends on the structure within the jobs the civil staff perform. Ways of increasing job satisfaction are the subject of many studies in their own right (Feldman (1937), Viteles (1954), Hollway (1991)), and depend heavily on what options are feasible within an organisation such as the police.

Interpreting other implications of the model, an individual who feels the public's view of the South Yorkshire Police is positive, compared to an individual who feels the opposite, with all other characteristics the same, is more likely to have more favourable morale. In order to improve respondents' perception of the view of the public, and thus consequently improve morale, the police might campaign for the support of the public, or organise events to improve public relations.

How a respondent from the civil staff feels about promotion issues also has an effect on his/her morale. A respondent who has the same characteristics as a colleague, except that they agree with the statement 'promotions are given to those who earn them', whereas their colleague disagrees, is twice as likely to have more favourable morale according to the model. A modification of promotion policies may help to improve morale via this issue, or maybe greater communication regarding promotion issues. Another way of interpreting the information gained from this variable may be that those in direct contact with someone who earned a promotion, or earned one themselves, and got it, will be very likely to agree with the statement, whereas an individual may have an opposite experience and reply conversely. The effect of this factor may, therefore, be a personal thing, and down to the philosophy of the individual. The morale of a respondent who has perhaps missed out on promotion, compared to one who has been given promotion, may be worth examining for further insight.

Finally, it seems that those who have served in the civil support staff of the SYP for less than 2 years are more likely to have more favourable morale. This may be due to the novelty of the job or maybe the fact that after less than 2 years, an individual may still be learning a lot about his/her job. Alternatively, before a certain length of service, an individual may be less likely to be affected by the politics of an organisation which may contribute to worsening morale. Also, in general, an employee of less than 2 years service is less likely to have direct responsibility for others, with which a certain amount of extra burden may come, and subsequently morale may be affected. From the model, it could be concluded that improving the morale of those who have served longer should be more of a priority, compared with those with less than 2 years service.

It is important to stress that the implications interpreted from model 4.5c are kept in the context of the data used to construct the model, the civil support staff, pertaining to the specific variable information collected.

A reason for not being able to find a suitable proportional odds for the full dataset, or officers only, (using dummy variables) may be due to the loss of some information by using ordinal explanatory variables as nominal. This issue is addressed in section 4.3, as the use of scored categories for ordinal explanatory variables is explored, thus treating them as interval or continuous variables. The proportional odds model may not be appropriate to describe the behaviour of officers' morale. The continuation odds model is applied as an alternative to the proportional odds in section 4.5.

The fact that the data patterns for officers and civil staff are different is possibly only explainable, without further insight, by the differing proportions of respondents in each classification. It maybe that with fewer respondents, the variation within the data for civil staff is decreased, therefore allowing more general odds ratios for a more favourable response to be accepted. Alternatively, it may be that the morale of civil staff is affected by different factors than that of officers, or maybe the same things in differing dimensions.

4.3: Modelling morale using scored levels for ordinal variables

Using the ordinal variables as continuous or interval variables may have an advantage over using dummy variables in that the quantitative nature of the variables is utilised. However, the fact that there is a discrete number of categories that a respondent may choose, when answering the questions from the survey, is lost and it is assumed by any modelling techniques that there are no restrictions on values the variable may take. Note that any model constructed will only be valid for the range of the data used to create it.

Whilst this investigation describes sophisticated methods that account for an ordinal response, there are no methods in common use that account for ordinality in explanatory variables without assigning scores. Therefore when using ordinal independent variables, one immediately faces a dilemma, in that one must decide whether to use the categorical nature of the variables, or their quantitative property.

Section 4.2.1 explored the former option of the two using dummy variables to represent the levels of the ordinal variables. This section explores the option of scoring the categories of ordinal explanatory variables.

Two sets of scores are assigned to the ordinal variable for comparison. Integer scores and CHAID estimated scores (see section 2.2.3) are used. The values of the CHAID scores are computed as those most likely to be associated with the response, respondent's own morale. The motivation for estimating the scores in this way, is to estimate the distance between the categories of a variable, that is to say we know that there is a natural ordering to the levels of the variable, but we do not know the magnitude of the difference between adjacent categories.

Using CHAID to estimate scores for the independent variables in a model presents somewhat of a methodological problem with the interpretation of any results. The CHAID scores are estimated as those most likely to be associated with the response variable morale, therefore data for the response must be known in order to estimate the scores. By taking these scored variables and attempting to model morale with them, the fact that the scores were obtained using the response is ignored, and the response and explanatory variables are assumed to be independent in their conception, when this is not the case, as the values of the scored explanatory variables are computed to fit the association between them and the response.

Therefore, modelling the response using ordinal explanatory variables with scored levels estimated in this way is similar to using a log-multiplicative model (see Agresti (1984)), with the difference that the scores that give the best fit, for each explanatory variable, are estimated independently of anything else, and externally from the modelling procedure.

It should be noted that the scores computed for levels of ordinal variables by CHAID are based purely on the relationship with one variable, i.e. in this case morale. Effectively, the CHAID scores are summarising the pattern of the association between the two variables involved, so that if the levels of the response is coded 1, 2, 3, to correspond to higher, neither high nor low and lower morale respectively, then when estimating scores for a variable using morale, a higher score will be assigned to the level of the variable that is associated with lower morale (larger value response), and similarly a lower score to that which is associated with higher morale.

Modelling morale using dummy variables showed that the data collected for police officers and civil support staff behave differently. Therefore, separate analyses for officers and civil staff are performed, using scores for the ordinal variable levels..

The computation of CHAID estimated scores for the ordinal variables commsen, pubview and promearn (lenserv is dichotomised) calibrated by respondent's own morale, for officers and civil staff separately, show up some interesting patterns in the data.

The variables pertaining to feelings about communication with more senior managers/officers, perception of public view of SYP and agreement with the statement that promotions are given to those who earn them, are measured on scales designed to be ordinal, with an association with morale expected to be linear. For officers, the expected relationship is supported for all variables. For civil staff, however, the expected association involving the variables commsen and promearn assumed, is not observed.

For civil staff, the scores assigned to the levels of commsen and promearn that are most likely to be associated with respondent's own morale are :-

Commsen	CHAID score	Promearn	CHAID score
1	0	1	46.70
2	29.95	2	0
3	44.41	3	42.89
4	100	4	59.52
5	94.60	5	100

The relationships between the explanatory variables observed from these scores are counter-intuitive. The scores imply the effects of communication and opinion on promotion issues on morale is not as expected.

Interpreting these scores, the pattern observed in the data for civil staff suggests that those who feel communication is very bad have slightly higher morale, in general, than those who feel communication is 'only' bad. Also, those who agree strongly with the statement that 'promotions are given to those who earn them' have lower morale in general than those who agree and those who neither agree nor disagree.

The association between communication and morale for civil staff is found not to be statistically significant in a model, therefore the pattern observed may be explained as random. There are also relatively few respondents who feel communication is very bad, 23 individuals, which may be insufficient to base conclusions on.

Possible explanations for the trait observed for the variable promearn could include the following:-

a) A problem with the data collection, in that the respondents may have interpreted the question differently to how it was designed. For officers, the relationship between the responses for promearn and respondent's own morale behave intuitively, or as expected. Therefore the difference in the nature of the civil staff data, may suggest a difference in perception between the two types of individual.

b) Similarly, the variable promearn is measured on a 5 point scale, and thus agreement with the statement is measured by direction and also strength of direction. It is possible that a respondent may agree or disagree with the statement, but at the same time be unsure of the strength of their opinion, especially if the idea is fairly new. If this is possible, then it is also possible that the dimension measuring whether the agreement/disagreement is strong or not may not be useful. The factor of interest may be that the respondent has agreed or disagreed, not the strength of their opinion.

c) A very small proportion of respondents strongly agree with the statement that promotions are given to those who earn them. This number (12 out of 399) may be insufficient to determine a general pattern of response, and the behaviour of the data observed may not reflect a true relationship.

d) The pattern observed may be genuine, and therefore the relationship between the variables, for this group of respondents (civil staff), may not be as expected or anticipated.

In order to proceed with modelling, considering c) above mainly, but also pertaining to b), the ordinal variables are collapsed to 3 categories to avoid sparse data, when estimating scores, and any misleading interpretations. This collapsing of the ordinal variables is also applied to officers' data as the proportions of respondents in the extreme categories of the variables are relatively small.

Estimating scores for the ordinal variables collapsed to 3 categories, so that categories in the same direction of opinion are combined, supports the assumption that the variables are ordinal in nature, and illustrates that the relationships between these variables and respondent's own morale are as intuitively expected.

4.3.1: Modelling the morale of officers

The variables jobsat and lenserv are used in the following models in the same form as for previous models, i.e. as binary. The ordinal variables are assigned scores to their levels.

Firstly, the ordinal variables are assigned integer scores so their value is that which is coded from the original survey, and depicts the order of the levels as 1, 2 and 3 (newcom, newpub and newprom below).

Scores estimated using CHAID are assigned separately, these scores are the ones most likely to be associated with respondent's own morale :-

Commsen level	newcom	CHAID score
1 & 2	1	1
3	2	45.85
4 & 5	3	100

Pubview level	newpub	CHAID score
1 & 2	1	1
3	2	62.10
4 & 5	3	100

Promearn level	newprom	CHAID score
1 & 2	1	1
3	2	40.87
4 & 5	3	100

Modelling morale for officers using the CHAID estimated scores for values of pubview, commsen, and promearn results in the violation of the proportional odds assumption. The score test statistic for proportional odds is $\chi^2 = 20.98$ with 5 df ($p<0.001$), therefore the model is discarded.

The corresponding model using integer scored ordinal variables is given by :-

$$\log[\gamma_j(\mathbf{x}_i)/\{1-\gamma_j(\mathbf{x}_i)\}] = \alpha_j + \beta_{job}(\text{jobnew}) + \beta_c(\text{newcom}) + \beta_{pu}(\text{newpub}) + \beta_{pr}(\text{newprom}) + \beta_l(\text{lendum}) \qquad (4.7o)$$

for j=1,2. Where $\gamma_j(\mathbf{x}_i)/\{1-\gamma_j(\mathbf{x}_i)\} = P(Y_i \le j)/[1-P(Y_i \le j)]$.

Using integer scores also treats the ordinal variables as if they were continuous variables, however, this model (4.7o) satisfies the proportional odds assumption. The score test statistic for proportional odds given as $\chi^2 = 7.84$ with 5 df ($p=0.165$).

The decrease in deviance from fitting the model with covariates, compared to intercept only is 562.43 with 5 df, following a χ^2 distribution (deviance for intercept only = 3081.64, deviance with covariates = 2519.21. Note that due to the sample size, the deviance for officers is much larger than for civil staff), so the contribution of the explanatory variables to the fit of the model to the data is highly statistically significant.

The parameters estimated for model 4.7o are tabulated below :-

Table 4.6: Parameter estimates, their standard errors and p-values, for model 4.7o

Parameter	Estimate	s.e.	p-value
$\alpha 1$	0.70	0.25	0.0065
$\alpha 2$	2.40	0.26	0.0001
βjob	2.00	0.14	0.0001
βc	-0.57	0.08	0.0001
βpu	-0.55	0.08	0.0001
βpr	-0.41	0.07	0.0001
βl	1.65	0.30	0.0001

* denotes parameter not significant

All the parameters in the model are statistically significant. The values of the estimates reflect intuitive relationships between the explanatory variables and morale. The estimates for the ordinal variables (treated as continuous) are negative indicating that a higher value for the variable, pertaining to a more negative characteristic, will decrease the odds of a more desirable level of morale.

To test the goodness-of-fit of model 4.7o, the Lipsitz et al (1996) method is applied, as for model 4.5c in section 4.2.1.

The data is ordered by predicted mean score, calculated using the response probabilities obtained from the model. The data is then partitioned in to 10 groups of approximately equal size, as close to 144 respondents in each, whilst not splitting any observations with tied values between groups :-

obs no. -	n_g	group, g
1 - 179	179	1
180 - 310	131	2
311 - 412	102	3
413 - 602	190	4
603 - 708	106	5
709 - 846	138	6
847 - 1017	171	7
1018 - 1149	132	8
1150 - 1291	142	9
1292 - 1438	147	10

Using dummy variables I1, ..., I9 as assigned for model 4.5c to indicate these groupings, similarly to model 4.6c, the model 4.7o is refitted with the extra goodness-of-fit parameters added.

The model used to test the goodness-of-fit of model 4.7o is therefore given by :-

$$\log [\gamma_j(\mathbf{x}_i)/\{1-\gamma_j(\mathbf{x}_i)\}] = \alpha_j + \beta_{job}(\text{jobnew}) + \beta_c(\text{newcom}) + \beta_{pu}(\text{newpub}) + \beta_{pr}(\text{newprom}) + \beta_l(\text{lendum}) + \gamma_1(I_1) + \ldots + \gamma_9(I_9) \quad (4.8o)$$

for j=1,2. Where $\gamma_j(\mathbf{x}_i)/\{1-\gamma_j(\mathbf{x}_i)\} = P(Y_i \leq j)/[1-P(Y_i \leq j)]$

The added group indicator parameters decrease the deviance of model 4.7o by 4.07 with 9 df, following a χ^2 distribution (deviance for intercept only = 3081.64, deviance with covariates + g-o-f = 2515.14. Decrease = 566.50). Collectively the goodness-of-fit terms make no statistically significant contribution to the fit of the model. This suggests that model 4.7o is an adequate fit to the data. The parameter estimates for the group indicator variables are given below :-

Table 4.7: Parameter estimates, their standard errors and p-values, for goodness-of-fit parameters obtained from model 4.8o

Parameter	Estimate	s.e.	p-value
γ_1	0.43	1.24	0.7289*
γ_2	0.32	1.10	0.7710*
γ_3	0.22	1.08	0.8347*
γ_4	0.40	0.98	0.6848*
γ_5	-0.01	0.94	0.9955*
γ_6	0.11	0.86	0.8971*
γ_7	0.04	0.75	0.9539*
γ_8	-0.07	0.61	0.9050*
γ_9	0.05	0.39	0.8996*

* denotes parameter not significant

The fact that the goodness-of-fit parameters offer no significant improvement indicates that the model accounts for a large enough proportion of the data, so that the systematically assigned variables are not required.

Table 4.8 contains observed counts, expected counts, approximate residuals and adjusted approximate residuals for each group and response level, as computed from the results of fitting model 4.7o :-

Table 4.8: Observed and Expected values with standardised residuals for model 4.7o

G		Morale			G		Morale		
		Higher	Neither	Lower			Higher	Neither	Lower
1	O	147	24	8	6	O	56	58	24
	E	144.09	27.11	7.8		E	61.16	50.99	25.85
	R*	0.55	-0.60	0.07		R*	-0.88	1.24	-0.40
	R**	0.68	-0.74	0.09		R**	-1.08	1.52	-0.49
2	O	90	31	10	7	O	66	52	53
	E	89.46	31.26	10.28		E	59.89	67.26	43.85
	R*	0.10	-0.05	-0.09		R*	0.98	-2.39*	1.60
	R**	0.12	-0.06	-0.11		R**	1.20	-2.94*	1.96
3	O	67	24	11	8	O	31	48	53
	E	66.04	26.70	9.26		E	31.90	51.31	48.80
	R*	0.20	-0.61	0.60		R*	-0.18	-0.59	0.76
	R**	0.25	-0.75	0.74		R**	-0.22	-0.72	0.93
4	O	112	63	15	9	O	13	47	82
	E	108.16	58.66	23.18		E	16.40	41.70	83.90
	R*	0.56	0.68	-1.81		R*	-0.89	0.98	-0.32
	R**	0.69	0.84	-2.22*		R**	-1.09	1.20	-0.39
5	O	51	40	15	10	O	7	24	116
	E	56.32	34.90	14.78		E	6.17	21.76	119.07
	R*	-1.04	1.05	0.06		R*	0.34	0.52	-0.65
	R**	-1.28	1.29	0.07		R**	0.42	0.64	-0.80

* denotes significantly large residual

Only 1 of the 30 approximate standardised residuals is significantly large, and only 2 of the 30 adjusted. In general, the residuals are fairly small, only 20% greater than 1. Of the 30 residuals, 16 are positive and 14 negative with no pattern to the differing polarity. The standardised residuals give little evidence of an ill-fitting model, therefore these diagnostics suggest the model is adequate. It should be noted that whilst the residuals indicate that the model is adequate, the data is compressed into 30 cells. The possible permutations for variable values is 324, with the model accounting for more possibilities than this, as it treats the ordinal variables as continuous and therefore able to take any value, therefore some loss of information is almost inevitable.

To illustrate the performance of model 4.7o, the predicted mean scores are plotted against the observed values.

Figure 4.2: Plot of Predicted mean scores vs Observed responses for model 4.7o

```
                Legend: A = 1 obs, B = 2 obs, etc.

Observed      |
Response   3  |        H AIK  IO F  WICBBAZFI B EZA PBJL  PPD ZQALZ U
              |
              |
              |
              |
              |
              |
           2  |     C A AS AZW  YZ P AZZEB  XGP B EPB F KD  MKA LE  F B
              |
              |
              |
              |
              |
              |
           1  |    QPCK IZ CZZA ZZ T AZXDD  ZEI A  P  C BB  DB  AB AC
              ----------------------------------------------------------------
                1.0            1.5            2.0            2.5            3.0

                                   Predicted Mean Score

NOTE: 395 obs hidden.
```

The plot shows clusters of observations around the pertinent areas of the diagonal, indicated by the Z's, and with 395 observations hidden behind these Z's (as they can only represent 26 respondents), the concentration of points in the approximately diagonal regions is more greater than displayed. The plot does show a general tendency of increasing predicted mean score as observed response is increased, supporting the evidence that the model is adequate. As was found using dummy variables, the plot is not conclusive, but at the same time does not disprove the assumption that the model fits the data.

Having accepted model 4.7o to describe the behaviour of respondent's own morale, with respect to the variable information gathered, the implications of the model must be examined. Refer to table 4.6 for log odds ratios (values of parameter estimates).

The terms for job satisfaction and length of service can be interpreted as for model 4.5c, using dummy variables on civil staff data. The magnitude of the estimates for these variables are also similar. If two police officers have identical explanatory characteristics, except that one is satisfied with their job and one is not, the one who is job satisfied is roughly 7.4 times more likely to have more favourable morale. Similarly, for two officers with identical characteristics except one has served the police for less than 2 years, whereas the other has served for 2 years or more, the officer with the shorter length of service is roughly 5 times more likely to have more favourable morale.

The ordinal variable terms are interpreted similarly to those in an ordinary regression model. Using the parameter estimate pertaining to communication with more senior managers/officers, for every unit increase in the value of the variable commsen (i.e. communication is deemed one level worse), remembering the variable may take the value 1, 2 or 3 only, the odds of more favourable morale are decreased by roughly 1.8 times. The parameter estimates for perceived public view of SYP, and agreement with the statement promotions are given to those who earn them, can be interpreted similarly.

The implications for this model are the same as for model 4.5c, pertaining to civil staff data using dummy variables to represent the ordinal variables, except that communication has a significant relationship with morale for officers, whereas it does not for civil support workers.

Discussion on improving job satisfaction, perceived public view of SYP, agreement that promotions are earned and comments on length of service, are given in section 4.2.1 for model 4.5c, and do not differ for model 4.7o, except to apply to officers instead of civil staff. According to the model, respondents who feel communication between them and more senior officers/managers is good are more likely to have more favourable morale in general. Ways of improving communication might include more personal contact, about the structure and objectives of the South Yorkshire Police, maybe even increased input into decision making for all officers. Improved feedback from more senior officers/managers may help. The nature of the variable is fairly self explanatory, and whilst the model constructed implies the relationship, someone with insight into communication within the South Yorkshire Police will be better equipped to discuss or act on the implication.

The implications interpreted from model 4.7o should be kept in the context of officers only, pertaining to the specific questions asked in the survey.

4.3.2: Modelling the morale of civil Staff

The variables jobsat and lenserv are coded as binary, as before. The ordinal variables are assigned scores in the same manner as previously. Integer scores depict the order of the levels as 1, 2 and 3 (newcom, newpub and newprom below), and scores estimated by CHAID are given below:-

Commsen level	newcom	CHAID score
1 & 2	1	1
3	2	26.37
4 & 5	3	100

Pubview level	newpub	CHAID score
1 & 2	1	1
3	2	63.11
4 & 5	3	100

Promearn level	newprom	CHAID score
1 & 2	1	1
3	2	56.73
4 & 5	3	100

The scores estimated by CHAID, for civil staff. show a monotonic relationship between the ordinal explanatory variables and morale, whereas using 5 categories led to a non-monotonic relationship, discussed above.

As for models using dummy variables to represent communication with more senior officers/manager, models using scored levels of commsen found the variable to be non-significant with respect to civil staff respondents. Therefore commsen is excluded from the analysis of main effects.

A model produced using the ordinal variables pubview and promearn, with CHAID scored levels, satisfies the proportional odds assumption (see Appendix 4b). The decrease in deviance due to the explanatory covariates is highly statistically significant, χ^2 = 156.12 with 4 df, (deviance with intercept only = 874.495, deviance with covariates = 718.375). However, the corresponding model using integers scores performs better, based on the decrease in deviance, therefore the model using integer scores is preferred.

The model using integer scores is given by :-

$$\log [\gamma_j(\mathbf{x}_i)/\{1-\gamma_j(\mathbf{x}_i)\}] = \alpha_j + \beta_{job}(\text{jobnew}) + \beta_{pu}(\text{newpub}) + \beta_{pr}(\text{newprom}) + \beta_l(\text{lendum}) \quad (4.7c)$$

for j=1,2. Where $\gamma_j(\mathbf{x}_i)/\{1-\gamma_j(\mathbf{x}_i)\} = P(Y_i \leq j)/[1-P(Y_i \leq j)]$

The proportional odds assumption is accepted for model 4.7c, the score test statistic, $\chi^2 = 2.79$ with 4 df (p=0.59).

The decrease in the deviance, for the model including the explanatory variables in the model is statistically significant, χ^2 = 168.27 with 4 df (deviance with intercept only = 874.495, deviance with covariates = 706.221). Note that is a decrease of 12.15 from the model using CHAID estimated scores for pubview and promearn, on no extra degrees of freedom.

The parameter estimates computed from model 4.7c are given in table 4.9 :-

Table 4.9: Parameter estimates, their standard errors and p-values, for model 4.7c

Parameter	Estimate	s.e.	p-value
$\alpha 1$	-0.63	0.42	0.1308
$\alpha 2$	1.54	0.43	0.0003
βjob	1.99	0.23	0.0001
βpu	-0.72	0.14	0.0001
βpr	-0.35	0.13	0.0063
βl	1.24	0.25	0.0001

* denotes parameter not significant

The explanatory variable terms are all highly statistically significant in the model, and the estimates given reflect logical relationships between the independent variables and the response.

The method by Lipsitz et al (1996) is again used to indicate the adequacy of the fit of model 4.7c.

The data is ordered by predicted mean score and then partitioned in to 6 groups of approximately equal size, as for model 4.5c. There are 399 civil staff, so the groups will consist approximately 66 respondents, whilst not splitting any observations with tied values between groups :-

obs no.	n_g	group, g
1 - 57	57	1
58 - 130	73	2
131 - 214	84	3
215 - 262	48	4
263 - 335	73	5
336 - 399	64	6

Dummy variables I1, ..., I5 are assigned to indicate the groups. The model 4.7c is refitted with the extra goodness-of-fit parameters added.

The model used to test the goodness-of-fit of model 4.7c is therefore given by :-

$$\log [\gamma_j(\mathbf{x}_i)/\{1-\gamma_j(\mathbf{x}_i)\}] = \alpha_j + \beta_{job}(\text{jobnew}) + \beta_{pu}(\text{newpub})$$
$$\beta_{pr}(\text{newprom}) + \beta_l(\text{lendum}) +$$
$$\gamma_1(I_1) + \ldots + \gamma_5(I_5) \qquad (4.8c)$$

for j=1,2. Where $\gamma_j(\mathbf{x}_i)/\{1-\gamma_j(\mathbf{x}_i)\} = P(Y_i \le j)/[1-P(Y_i \le j)]$.

The added group indicator parameters bring about a decrease in deviance, from model 4.7c, of 9.06 with 5 df following a χ^2 distribution (p>0.10) (deviance for intercept only = 874.495, deviance with covariates + g-o-f = 697.165. Decrease = 177.33). Collectively the goodness-of-fit terms make no statistically significant contribution to the fit of the model. This suggests that model 4.7c is an adequate fit to the data. The parameter estimates for the group indicator variables are given below :-

Table 4.10: Parameter estimates, their standard errors and p-values, for goodness-of-fit parameters obtained from model 4.8c

Parameter	Estimate	s.e.	p-value
γ_1	1.61	1.51	0.2854*
γ_2	1.97	1.21	0.1033*
γ_3	1.06	0.94	0.2580*
γ_4	0.75	0.74	0.3148*
γ_5	-0.02	0.52	0.9634*

* denotes parameter not significant

The goodness-of-fit parameters offer no significant improvement to model 4.7c, indicating that the model describes the patterns sufficiently, so that the systematically assigned variables are not required.

Table 4.11 contains observed counts, expected counts, approximate residuals and adjusted approximate residuals for each group and response level, as computed from the results of fitting model 4.7c :-

Table 4.11: Observed and Expected values with standardised residuals for model 4.7c

G		Morale		
		Higher	Neither	Lower
1	O	37	19	1
	E	40.23	14.10	2.67
	R*	-0.94	1.50	-1.05
	R**	-0.85	1.35	-0.95
2	O	43	24	6
	E	36.09	29.07	7.84
	R*	1.62	-1.21	-0.70
	R**	1.46	-1.09	-0.63
3	O	22	46	16
	E	25.59	40.47	17.94
	R*	-0.85	1.21	-0.52
	R**	-0.77	1.09	-0.47

G		Morale		
		Higher	Neither	Lower
4	O	11	21	16
	E	9.28	23.05	15.66
	R*	0.63	-0.59	0.10
	R**	0.57	-0.53	0.09
5	O	8	19	46
	E	7.02	27.03	38.68
	R*	0.39	-2.00*	1.72
	R**	0.35	-1.80	1.55
6	O	1	17	46
	E	2.21	12.88	48.91
	R*	-0.82	1.29	-0.86
	R**	-0.74	1.16	-0.77

* denotes significantly large residual

Only 1 of the 18 approximate standardised residuals is significantly large. None of the 18 adjusted residuals are significant, due to the scaling down of the residuals in relation to the general size of the R*'s. In general, the residuals are not overly large, less than 50% are greater than 1. Of the 18 residuals, 10 are positive and 8 negative with no pattern to the differing direction. There is nothing unusual or noteworthy about the standardised residuals, suggesting that the model is adequate to describe the data. It should be noted again that the data is aggregated into 18 cells, where there are 108 permutations of possible variable values, with the model accounting for a larger number than this, as it treats the ordinal variables as continuous and therefore able to take any value, thus almost certainly resulting in the loss of some information

To illustrate the performance of model 4.7c, the predicted mean scores are plotted against the observed values.

Figure 4.3: Plot of Predicted mean scores vs Observed responses for model 4.7c

```
                    Legend: A = 1 obs, B = 2 obs, etc.

Observed
       3 |          A         C A A A D    A   M    M A KBD  BL A N  Y F O
Response |
         |
         |
         |
       2 |          C C G   FH B H A S C  L    Y    I D GAB  AD A C  M C A
         |
         |
         |
         |
         |
       1 |          F J J   KN H F C PBA  F    N    D B D A       C    A
      -------------------------------------------------------------------------
         1.00     1.25     1.50     1.75     2.00     2.25     2.50     2.75     3.00

                                   Predicted Mean Score
```

It can be seen from the plot that there is a general trend of increasing predicted mean scores as the observed response increases. The spread of the points is fairly large along all three response levels, and for morale levels 2 and 3 the model seems to perform better, as the accuracy of predicted mean scores for higher morale (response = 1) is less. The points denoted by the latter letters of the alphabet show clusters of points in pertinent areas, which supports the validity of the model. As found for previous models, the plot does not necessarily suggest that model fits well, but does not disprove the other diagnostics.

The parameter estimates can be interpreted in the same way as for model 4.7o.

The implications for this model are identical to those for model 4.5c, which models the same data using the same variables, with the ordinal variables depicted by dummy variables and used as nominal, rather than used as continuous or interval scale variables. The effects on morale of the explanatory variables estimated by the two models, 4.5c and 4.7c, are identical in interpretation and very similar in magnitude. For example, referring to table 4.3, the parameter estimates for model 4.5c, if all else is equal, an individual who feels the public view of SYP is positive is 4.2 times more likely to have more favourable morale than an individual who feels the public view is negative (log odds parameter, $\beta_{pu1} = 1.44$). Equivalently, for model 4.7c, the numerical interpretation for the same comparison is $\exp\{1(\beta_{pu}) - 3(\beta_{pu})\} = \exp\{-0.72 - (-2.16)\} = \exp\{1.44\} = 4.2$, therefore this relationship is identical for both models. The other comparisons are not quite identical, but differences are negligible and the implications are the same.

4.4: Comparison of dummy variables and scored categories

For civil staff, the analyses using the variables commsen, pubview and promearn in different forms are analogous. Whether the variables are treated as nominal or interval scale makes no difference to the results and implications. In this eventuality, the model which uses the ordinal variables as interval may be preferred, as the ordinal variables, in their nature, are theoretically more like interval variables than nominal. The model 4.7c is more parsimonious, in terms of the number of parameters it uses, to describe the same degree of detail in the data, as the model with dummy variables.

For officers, the two analyses produce different results with the respect to the proportional odds modelling procedure. The model using dummy variables violates the proportional odds assumption, so that the odds of higher morale versus neither high nor low or low are not the same as the odds of higher or neither high nor low morale versus low. The model using scored categories for ordinal variables, however, satisfies the assumption.

Examining the binary logistic models using dummy variables that correspond to the two possible dichotomies of the response, there is discrepancy between the parameter estimates of pubview and promearn for the alternative models. This discrepancy is not evident for the corresponding binary logistic models using scored categories. A possible explanation for this may be the fact that using dummy variables employs two parameters per original variable, compared to a single parameter for scored categories, when modelling proportional odds. The dummy variables are separately estimating the log odds of more favourable morale, for levels 1 and 2 of the variables, compared with level 3, and thus the parameter estimates are independent of each other and therefore unconstrained. The scored categories are effectively saying that level has double the effect of level 1, level 3 has three times the effect of level 1 and level 3 has 1.5 times the effect of level 2, therefore a parameter that fits this constraint is estimated. The dummy variables for a particular factor do not necessarily have a linear relationship with the response. When estimating the odds of higher morale versus neither high nor low, and the odds of high or neither high nor low morale versus low, separately, the difference in effects on the response, of the levels of an ordinal variable, may be too great to use a global log odds ratio, for the dummy variables. When using scores, the effects are constrained to be linear, so the discrepancy may be averaged out to the value of the parameter that gives best fit. Therefore, for the different dichotomies of respondent's own morale, the discrepancy between corresponding parameters for dummy variables may be of opposite polarity. However, when using scores the discrepancies may be smoothed to a similar magnitude. Thus, the dummy variables violate the proportional odds assumption, whilst the linearity of the effects of pubview and promearn when categories are scored, satisfies the assumption.

The variables pubview and promearn are designed to be ordinal, and as ordinal variables more closely resemble interval variables than nominal, assigning scores to the categories may be appropriate. The analyses performed suggest that the use of scores offers some advantage, in the context of this dataset, to the modelling of morale using the proportional odds model. The advantages of using scores are more parsimonious models, and for officers specifically, the acceptance of the proportional odds assumption, allowing an adequate model for the data to be constructed.

4.5: The Continuation Odds model and SYP data

For analysis via the continuation odds model, the variables were coded identically as for the corresponding proportional odds models (sections 4.2 and 4.3). The morale of officers and civil staff was modelled separately for the same reasons given in the previously.

Fitting the continuation odds model in SAS, there is no automated test for the continuation odds assumption. In order to ascertain whether a global odds ratio parameter which measures the log odds of an individual having the 'most desirable morale available' is suitable, i.e. simultaneously modelling the log odds of higher morale versus neither high nor low or lower, and neither high nor low morale versus lower, the corresponding binary logistic models for the dichotomies of the response in that way must be examined, and corresponding parameter estimates compared.

Several continuation odds models were fitted for both officers and civil staff. Using the corresponding binary logistic models to gauge the continuation odds assumption, no satisfactory models were found. In all cases the parameters estimates, pertaining to more than one highly significant explanatory variable in the continuation odds model, were very different, either in magnitude, direction or significance. Therefore using the continuation model, to describe the relationships between the explanatory variables and morale, is not considered appropriate.

The application of the continuation odds model to the South Yorkshire Police data does not produce a reasonable model, for the description of the behaviour of respondent's own morale, whereas the proportional odds satisfactorily modelled the response variable. A simple explanation for the violation of the continuation odds assumption can be offered, in that the nature of the response is perhaps not appropriate for modelling using this technique. The proportional odds model is successfully applied to the data, simultaneously modelling the log odds of the two possible dichotomies of the morale variable, in both cases comparing higher morale with lower morale in some context. The continuation odds model, however, simultaneously models logits that compare higher morale with lower (and neither high nor low), and neither high nor low morale versus low. The continuation odds model is modelling the log odds of membership in the most favourable morale group available, and when higher morale is taken out, this group is neither high nor low morale. The characteristics of respondents who feel their morale is neither high nor low, compared with those of respondents with lower morale, are different to the characteristics of respondents with higher morale compared to those of individuals with neither high nor low or lower morale. For the continuation odds assumption to be satisfied, the characteristics of respondents with neither high nor low morale must be very complex. Firstly (not denoting any chronological ordering), the neutral morale group are combined with the lower morale group, and the odds of having higher morale rather than neither high nor low or lower morale are estimated. Secondly, the higher morale group is excluded from the analysis, and neither high nor low morale is now the most desirable, and the odds of having this (neutral) level of morale rather than lower morale are estimated. The continuation odds assumption assumes that these sets of odds are equivalent, whereas it seems, conceptually, that this is very unlikely.

4.6: Discussion of the proportional odds and continuation odds assumptions

The proportional odds assumption is reasonable, in that for the scenario where an individual who has certain characteristics is, say, most likely to have higher morale, it is expected that he/she will be less likely to have neither high nor low morale than higher, but more likely to have neither high nor low morale than lower. Therefore, the proportional odds assumption is assuming the logits for the dichotomies of morale are equivalent. This is feasible as both dichotomies are comparing odds of higher morale with lower morale. In both dichotomies, the neither high nor low morale group could be seen to be 'diluting' the differences between the higher and lower morale groups. Theoretically, the feasibility of the proportional odds assumption can be illustrated numerically, by using scores assigned to the levels of morale. If we assign integer scores 1, 2 and 3 to depict higher, neutral and lower morale respectively, when we simultaneously dichotomise these levels, we get 1 versus 2 + 3, and 1 + 2 versus 3. Averaging the levels in the groups where two levels are combined then gives us 1 versus 2.5, and 1.5 versus 3, in both cases a difference of 1.5, assuming the 'distance' between morale levels is equal. This also relies on the relationship between the explanatory variables and morale to be as expected, i.e. positive characteristics influence higher morale, negative characteristics influence lower morale, and where applicable, neutral characteristics influence neutral morale. Similarly, the same numerical illustration for the continuation odds assumption becomes 1 versus 2.5, and 2 versus 3. Although this illustration vastly simplifies the techniques, the concepts are comparable.

Chapter 5: Conclusion, discussion and further research

5.1: Conclusion and discussion

Techniques for analysing ordinal data are in their relative infancy. This investigation has illustrated and applied some of the existing methodology. The proportional odds model is perhaps one of the more widely used methods for analysing ordered categorical data. The assumption that the odds, for a 3 level ordinal response variable, say, of responding in category 1, as opposed to 2 or 3, should be equivalent to the odds of responding in 1 or 2 as opposed to 3, is comprehensible, especially if the middle category is neutral, or a distinct level no more similar to either of the extreme levels than the other. The application of the proportional odds model proved successful to the South Yorkshire Police dataset, enabling implications about the effect of certain factors on the morale of police staff to be made. The implications from the models constructed should not be taken out of the context of the data used, but the results in themselves represent original findings. The difference in behaviour of morale between officers and civil staff represents a difference in the states of minds of the two sets of individuals. The factors which influence morale in general terms are job satisfaction, length of service, relationships with management/superiors, how the respondent perceives the public view of the police force, possibly representing their own feeling about the force, and promotion issues, according to the implications of the proportional odds models fitted.

This information is useful to the South Yorkshire Police, if an improvement in morale is an objective. The findings are also relevant to other areas of research, possibly more in a social psychology context than any other, as the work backs up some theories discussed in Chapter 3. The introduction of the concept 'relative morale' (section 3.14) represents original research. Relative morale is an extremely useful and interesting descriptive tool, especially where an absolute level of morale may not be appropriate.

The application of the continuation odds model did not produce a meaningful model for the data. The idea that the continuation odds assumption should be satisfied for any set of data seems almost unfeasible. The concept of modelling the log odds of membership of the most desirable response category available, is fair motivation for the development of the model. However, the assumption, for a 3 level ordinal response variable, say, that the odds of responding in category 1 as opposed to 2 or 3 should be equal to the odds of responding in category 2 as opposed to 3, seems unlikely to be satisfied. The latter contrast, in the design of the model, uses only a subset of the data, as respondents in category 1 of the response are ignored, possibly further confusing the conditions under which the assumption may be valid. Discussion on the proportional odds and continuation odds assumption is given in section 4.6.

The goodness-of-fit test introduced by Lipsitz, Fitzmaurice and Molenberghs (1996), extended from the methods by Hosmer and Lemeshow (1980, 1989), represents a useful diagnostic tool, that has perhaps been absent from ordinal regression models. The method adapted for ordinal response variables is very new, as the published date suggests, and therefore the application of the technique is probably fairly limited at the present time. Criticism/development of the method is therefore scarce due to its newness. Section 2.5.3 represents original work, in the form of a modification of the guidelines to applying the goodness-of-fit test proposed by Lipsitz et al. The data is partitioned, and goodness-of-fit statistics produced for the partitions. When using explanatory variables that are categorical, discrete or simply have relatively few possible permutations of values, the partitioning must be made according to certain criteria discussed in section 2.5.3. The comments may be applied generally, to any setting that utilises this method for assessing the fit of an ordinal regression model. The application of the Lipsitz et al goodness-of-fit test using the SAS statistical package also represents original work, in terms of code written. The application of the goodness-of-fit test to the continuation odds model is particularly complex.When using Proc LOGISTIC to fit the continuation odds model, the individual response probabilities are not output directly, as data manipulation enables the model to be fitted using the binary logistic model. Therefore the use of the goodness-of-test in this case is not so straightforward. (The code created to apply the Lipsitz, Fitzmaurice and Molenberghs method, for the proportional odds and continuation odds models, is given within the skeleton SAS programs contained in Appendices 3a to 3d).

The exploratory analysis of data using CHAID is not a widely used technique. The application of the CHAID analysis is more descriptive than anything else. The package allows much more specific descriptive statements to be made, than would be possible without much complex examination of specific crosstabulations. CHAID's aid to model building is fairly limited to suggestions for variables that may be useful in a model, possible interaction terms that may be useful, and levels of explanatory variables that may benefit from collapsing.

The investigation has achieved an insight into the treatment of ordinal data. Many descriptive approaches are applied, though the research is centered around the application and discussion/criticism of the proportional odds and continuation odds models, including methods of assessing their goodness-of-fit. The subject matter of the analysis, i.e. morale within the South Yorkshire Police, is also a domain that may generate interest, and therefore the application of the methods to this data is a pertinent area of research. The data set can be seen to be fairly complex in its behaviour. The approaches and processes used to overcome problems within the structure of the data, also give insight into the philosophy of analysing ordinal data.

Evidence that the development of methods for analysing ordered categorical data is in the relative minority is given by the inability to easily apply most of the techniques that actually have been developed. With continuous data, on a vast proportion of occasions, one can obtain valid analyses using a wide range of methods in different statistical areas with great ease. For ordinal data, and to an extent nominal data, a lot more consideration of how to treat the data, and even what form to use it in, must be given, before even thinking about applying a particular technique.

For both the proportional and continuation odds models described in Chapter 4, there are drawbacks to their use and application. For an ordinal response with k categories, both methods simultaneously fit k-1 sub-models, dichotomising the response to do so. The models use the logit transformation of the dichotomies to model, say, the log odds of category 1 vs categories 2 to k. This latter category, the merging of levels of the response, represents a possible problem of interpretation when, as found in this investigation, the levels combined have contrasting meanings. Referring to the SYP data, respondent's own morale is reduced to 3 categories because the dichotomy of the original response results in the collapsing of levels with virtually opposite interpretations. The dichotomy of very high morale vs the rest of the morale levels, merges high morale with low and very low morale. Therefore in order to use the models and ensure clear and more simple interpretation, the response is collapsed. The scenario described above does not represent any technical problem, as the variable is assumed to be a continuum, however, the fact that this data reduction is desirable represents a possible problem with the methodology.

The dichotomy of very high morale versus not very high morale is not a desirable one due to the different levels combined, and the way the response is structured, however, if the SYP survey had asked the question :-

Q: Please rate your morale at this time (1=highest, 5=lowest)

1 2 3 4 5

then the dichotomy of morale level 1 versus morale levels 2 to 5 may be less cause for concern. Therefore if this is the case, one must consider seriously the method of analysis of the information to be collected, in the planning stages of an investigation. If the data is to be ordinal in nature, it should be collected in a form such that it requires minimum or no manipulation before analysis.

5.2: Ideas for further research

The different approaches adopted for the SYP data, i.e. using dummy variables and scored levels for ordinal variables, highlights a deficiency in the development of methods that account for ordinal explanatory variables. The proportional odds and continuation odds models utilise the cardinality of an ordered categorical response variable, thus not requiring the assignment of scores, however, no such technique exists for the independent variables in a model.

The continuation odds model is discussed in section 4.5. The motivation for modelling the log odds of membership in the most desirable response category available is plain. The conditions under which the continuation odds assumption would be satisfied, however, seem unlikely to be met. Rather than the continuation odds model, a model that describes the odds of membership in the most desirable of any 2 adjacent response categories may be more useful, or applicable. Therefore, for a 3 level ordinal response variable, say, the model would simultaneously describe the odds of responding in level 1 as opposed to 2, and the odds of responding in level 2 as opposed 3, assuming these odds were equal. This assumption seems more tenable than that for continuation odds. The concept of this method uses only a subset of the data in any of the odds comparisons, and therefore the validity of such a technique may be questionable.

Diagnostics for ordinal regression models are an area where relatively little definitive literature can be found. The Lipsitz et al (1996) goodness-of-fit is very new, and goes some way to fill a gap in ordinal analyses. The newness of the technique, however, could be seen to indicate that the robustness of the method is yet to be fully examined.

A more obvious point is the lack of software available, specifically for the purpose of analysing ordered categorical data. A recent version of SPSS has an option to fit the continuation odds model, which represents a significant development in the methods. Software that applies the Lipsitz et al (1996) goodness-of-fit test automatically would be useful.

References

Agresti, A. (1984) Analysis of Ordinal Categorical Data. Wiley

Agresti, A. (1990) Categorical Data Analysis. Wiley

Anderson, J. (1984) Regression and ordered categorical variables. JRSS, series B, **46**, p1-30

Ashby, D. and West, C. R. (1989) The ordered logistic regression model in psychiatry: Rising prevalence of dementia in old peoples' homes. Statistics in Medicine, **8**, p1317-1326

Berridge, D. M. and Whitehead. J. (1991) Analysis of failure time data with ordinal categories of response. Statistics in Medicine, **10**, p1703-1710

Burt, C. (1924) The mental differences between individuals. Journal Nat. Inst. of Ind. Psychology, **11** (2), p67-74

Carroll, K. J. (1993) The application and comparison of logit regression models in hte analysis of ordered categorical data. SHU Thesis

Collet, D. (1991) Modelling binary data. Chapman and Hall

Cox, C. and Chaung, C. (1984) A comparison of chi-square partitioning and two logit analyses of ordinal pain data from a pharmaceutical study. Statistics in Medicine, **3**, p273-285

Cox, C. (1988) Multinomial regression models based on continuation ratios. Statistics in medicine, **7**, p435-441

Cox, D. R. (1972) Regression models and life tables. JRSS, series B, **34**, p187-220

Feldman, H. (1937) Problems in labour relations. Macmillan

Fienberg, S. and Mason, W. (1979) Identification and estimation of age-period-cohort models in the analysis of discrete archival data. Sociological Methodology, p1-67

Fienberg, S. (1980) The analysis of cross-classified categorical data. MIT Press.

Freeman, D. (1987) Applied Categorical Data Analysis. Dekker

Goodman, L. and Kruskal, W. (1954) Measures of association for cross classifications JASA, **49**, p732-764

Greenland, S. (1994) Alternative models for ordinal logistic regression. Statistics in Medicine, **13**, p1665-1677

Haberman, S. J. (1974) loglinear models for frequency tables with ordered classifications. Biometrics, **30**, p589-600

Hastie, T. J., Botha, J. L. and Schnitzler, C. M. (1989) Regression with an ordered categorical response. Statistics in Medicine, **8**, p785-794

Hollway, W. (1991) Work Psychology and Organisational Behaviour. Sage Publications

Home Office (1990) The Hillsborough stadium disaster, 15 April, 1989: Inquiry by the Rt Hon Lord Justice Taylor. Final Report. Home Office

Hosmer, D. and Lemeshow, S. (1980) A goodness-of-fit test for the multiple logistic regression model. Communs Statist. Theory methodology, **10**, p1043-1069

Hosmer, D. and Lemeshow S. (1989) Applied Logistic Regression. Wiley

Hull, R. L. (1939) Measuring employee attitudes - A proving ground for personnel policy and practices. Conf. Brd. Mgmt. Record, **Nov**, p165-172

Hutchison, D. (1985) Ordinal variable regression using the McCullagh (proportional odds) model. GLIM newsletter, **9**, p9-17

Iyer, R. (1985) Continuation odds models in ordinal variable regression. GLIM newsletter, **10**, p4-8

Kalbfleisch, J. and Prentice, R. (1980) The statistical analysis of failure time data. Wiley

Kass, G.V. (1980) An exploratory technique for investigating large quantities of categorical data. Applied Statistics, **29**, No. 2, p119-127

Kendall, M. G. (1938) A new measure of rank correlation. Biometrika, **30**, p81-93

Koch, G., Landis, J., Freeman, J., Freeman, D. and Lehnen, R. (1977) a general methodology for the analysis of experiments with repeated measurement of categorical data. Biometrics, **33**, p133-158

Landwehr, J. M., Pregibon, D. and Shoemaker, A. C. (1984) Graphical methods for assessing logistic regression models. JASA., **79**, p61-71

Lazarsfeld, P. F., Pasanella, A. K. and Rosenberg, M. (1972) The language of social research. Free Press

Lindsey, J.K. (1989) The analysis of categorical data using GLIM. Springer-Verlag

Lipsitz, S., Fitzmaurice, G. and Molenberghs, G. (1996) Goodness of fit tests for ordinal regression models. Applied Statistics, **45**, No. 2, p175-190

Magidson, J. (1992) Chi-squared analysis of a scalable dependent variable. Notes from 1992 Annual Meeting of A. S. A., Educational Statistics section.

McCullagh, P. (1980) Regression models for ordinal data. JRSS, series B, **42**, p109-142

McCullagh, P. and Nelder, J. (1989) Generalised Linear Models. Chapman and Hall

Nelder, J. and Wedderburn, W. (1972) Generalised Linear Models. JRSS, series A, **135**, p761-768

Pearson, K. (1904) Mathematical contributions to the theory of evolution XIII: On the theory of contingency and it's relation to association and normal correlation. Draper's co. research memoirs. Biometrics, **1**. (Found in E. S. Pearson (1948) Karl Pearson's early papers. Cambridge University Press)

Rao C. R. (1973) Linear Statistical Inference and it's Applications. Wiley

Rosenberg, M. (1968) The logic of survey analysis. Basic Books

SAS Institute (1989) SAS/STAT user guide, version 6.4, volume 2. SAS Institute

Somers, R. (1962) A new asymmetric measure of association for ordinal variables. American Sociological Review, **27**, p799-811

SPSS (1993) SPSS for windows, CHAID. Release 6.0. Magidson/SPSS inc.

SSRC (1994) The South Yorkshire Police: Justice with Courage. Customer surveys 1994. Survey and Statistical Research Centre, Sheffield Hallam University

Thomas, G. and Kiwanga, S. (1993) Use of ranking and scoring methods in the analysis of ordered categorical data from factorial experiments. The statistician, **42**, p55-67

Tsiatis, A. A. (1980) A note on the goodness-of-fit for logistic regression models Biometrika, **67**, p250-251

University of Minnesota (1951) Triple audit of industrial relations. Industrial Relations Bulletin, **11**. (Uni. of Minnesota)

Viteles, M. S. (1954) Motivation and Morale in Industry. Staples Press Ltd.

Wald, A. (1943) Tests of statistical hypotheses concerning several parameters when the number of observations is large. Trans. Amer. Math. Assoc., **54**, p426-482

Appendices

Appendix 1: South Yorkshire Police Staff Survey

Appendix 2: Crosstabulation of Commsen and Omor, controlling for Promearn (Civil Staff)

Appendix 3: Skeleton SAS programs

- Appendix 3a: Proportional Odds model using dummy variables
- Appendix 3b: Proportional Odds model using scored categories
- Appendix 3c: Continuation Odds model using dummy variables
- Appendix 3d: Continuation Odds model using scored categories

Appendix 4: SAS output for models fitted

- Appendix 4a: Proportional Odds models using dummy variables
- Appendix 4b: Proportional Odds models using scored categories
- Appendix 4c: Continuation Odds models using dummy variables
- Appendix 4d: Continuation Odds models using scored categories

Appendix 1

South Yorkshire Police Staff Survey

Sheffield Hallam University

Appendix 3

February 1994

SOUTH YORKSHIRE POLICE - WHAT DO YOU THINK?

I am writing to ask for your help with a very important survey we are undertaking on behalf of the Chief Constable and the Police Authority. The main aims of the survey are to find out your views on the management, and the organisation of the South Yorkshire Police Service.

You may remember completing a very similar questionnaire in October 1992 we have been asked to conduct a follow-up survey to see if your views have changed over the last two years.

All members of the South Yorkshire Police Service work force have been sent a questionnaire. This questionnaire has been designed to collect information on many important issues including the structure of the organisation, career development, morale, staffing and resources.

We are looking for your honest opinion and all the information you give will be **STRICTLY CONFIDENTIAL**. The analysis is conducted in such a way that NO individual can be identified in any way, and especially not by rank, gender or place of work.

Please could you return the completed questionnaire in the envelope provided by the end of February. If you have any queries concerning the survey, please feel free to ring myself or Anne Kirby on Sheffield 533791.

Yours sincerely

Roma Eastwood

Roma Eastwood
The Survey and Statistical Research Centre

School of Computing and Management Sciences
Survey and Statistical Research Centre
Hallamshire Business Park 100 Napier Street Sheffield S11 8HD Telephone 0742 720911
Direct line 0742 533121/533791 Fax 0742 533161
Director Dr S J Wisher

Confidential

South Yorkshire

POLICE

JUSTICE *with* COURAGE

What do YOU think?

his is a staff survey, designed to evaluate opinion of the South Yorkshire olice Service.

t shouldn't take long to complete - most questions just require you to tick box.

CONFIDENTIAL

All information you give will be completely confidential. No individuals will be identifiable during any of the analysis.

Public Opinion

1. What view do you think **the public** have of South Yorkshire Police?

Very positive	☐
Positive	☐
Neither positive nor negative	☐
Negative	☐
Very negative	☐

2. Do you think this view has changed over the last 2 years?

No, not really	☐	→ *Please go to Question 5 on the next page*
Yes, it has	☐	

3. *If Yes,* Do you think the public's view of South Yorkshire Police has got better or worse over the last two years?

Better	☐
Worse	☐

4. Why do you think this is?

....

....

....

Quality of Service

5. Overall, how satisfied are you with the level of service provided to the public by South Yorkshire Police?

Very satisfied	☐
Satisfied	☐
Neither good nor bad	☐
Dissatisfied	☐
Very dissatisfied	☐

6. Does anything prevent you or your colleagues from delivering the level of service you would like to?

No	☐	→ *Please go to Question 8*
Yes	☐	

7. *If Yes,* What would you say is the **main** thing that prevents you from delivering the level of service you would like to?

Please give details of the one factor that has the most influence

....

....

....

The Structure of South Yorkshire Police

8. Which one of the following statements do you think best describes the working relationship between the sub-divisions and force headquarters?

"It's an us and them situation"	☐
"There is a reasonable working relationship but it could be improved"	☐
"There is a close link between force headquarters and the sub-divisions"	☐
"I'm not sure"	☐

9. How would you describe the following types of communication within South Yorkshire Police? Also for each of these, please give a brief comment indicating why you say this:

		Very Good	Good	Neither good nor bad	Bad	Very bad
A	Communication between force headquarters and the sub-divisions					

A ....

		Very Good	Good	Neither good nor bad	Bad	Very bad
B	Communication between police officers and civil support staff					

B ....

		Very Good	Good	Neither good nor bad	Bad	Very bad
C	Communication with the public					

C ....

10. Please describe the communication between the following groups of staff:

		Very Good	Good	Neither good nor bad	Bad	Very bad	
A	Between you and your **immediate** supervisors						
B	Between you and your **more senior** managers/officers						
C	Between you and the people you supervise						*Tick here if this doesn't apply to you*

A ....

B ....

C ....

11. How would you describe the working relationship between you and your line manager?

Very satisfactory	
Satisfactory	
Neither satisfactory nor unsatisfactory	
Unsatisfactory	
Very unsatisfactory	

12. Thinking about the overall strategic planning of South Yorkshire Police please indicate if you agree or disagree with the following statements relating to the Force Business Plan and Local Priority Setting.

The Force Business Plan

	Agree	Disagree	Unsure
I was involved in			
I was consulted about			
I was informed about			

Local Priority Setting

	Agree	Disagree	Unsure
I was involved in			
I was consulted about			
I was informed about			

Career Development

13. Please tell us whether you agree or disagree with the following statements about promotions within South Yorkshire Police. *We are interested in hearing about the promotion system that relates to you. (Please tick a box for each statement)*

	Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree
People are promoted fairly					
Promotions are influenced by undisclosed information					
Promotions are given to people who have earned them					
Women are less likely to be promoted than men					
Ethnic minority employees are less likely to be promoted than other employees					
It's not what you know, it's who you know					
The present system for civilian staff is adequate					

14. How could the promotion system be improved?

....

....

....

15. Have you been promoted in the last 2 years?

Yes	☐
No	☐

16. How would you describe the amount of training you have received over the last two years?

More than adequate	☐	→ *Please go to Question 18*
Adequate	☐	→ *Please go to Question 18*
Inadequate	☐	

17. *If inadequate,* In what way?

....

....

Staffing and Resources

18. What is your perception of the staffing levels in the following areas? (*Please tick a box for each category*)

	Too many	About Right	Not Enough	Not sure
PC's in uniform patrol				
Operational police units at force HQ				
Special squads/units at division/ sub-division				
Civilian staff at force HQ				
Civilian staff in sub-divisions				
CID at sub-divisions				
Non-operational police staff at force HQ				

19. How would you describe the provision of the following?

	More than adequate	Adequate	Inadequate	Don't know / not sure
vehicles				
personal radios				
computers/word processing equipment				
Police buildings				

Morale

20. How would you describe morale at the moment?

	Your own morale	Amongst your colleagues
Very high		
High		
Neither high nor low		
Low		
Very low		

21. Has **your** morale changed over the last 2 years?

Yes it has improved	
Yes it has lowered	
No it has stayed the same	

22. If you had to select two things that you thought affected morale for the **better**, what would they be?

1st thing ....

....

2nd thing ....

....

23. If you had to select two things which you thought affected morale for the **worse**, what would they be?

1st thing ....

....

2nd thing ....

....

24. Summarise how you feel at the moment by ticking the appropriate boxes below:

(Please tick one box for each category)

	YES	NO	UNSURE
Satisfied with my job			
Satisfied with the criminal justice system			
Proud to be part of SYP			
Undervalued by others			
Treated unfairly			
Overworked			
Kept in the dark			
Paid a fair wage for the job			

25. Considering the above, please comment briefly on the two items which you feel most strongly about.

1 ....

2 ....

About Yourself

In order for us to understand a little about who is answering these questions, we need to know a few details about yourself.

26. Are you:

- Male ☐
- Female ☐

27. Would you describe yourself as

- White ☐
- Black ☐
- Asian *(Indian subcontinent)* ☐
- Other *(Please specify)* ☐

. . . .

28. How many years have you worked in the Police Service?

- Less than 2 years ☐
- 2-5 years ☐
- 6-10 years ☐
- 11-20 years ☐
- 21-30 years ☐
- Over 30 years ☐

29. Do you work shifts?

- Yes ☐
- No ☐

30. Are you:

- Civilian staff ☐
- Police officer ☐

31. Which one of the following describes your current duties?

- Uniformed patrol ☐
- Community constable ☐
- Control room ☐
- CID ☐
- Operational Support Units ☐
- Specialist role ☐
- Senior Management/ Management/supervision ☐
- Clerical ☐
- Administration ☐
- Manual support ☐
- Other ☐ *Please give details* ....

32. Where are you permanently based? (*Please tick just one of these boxes*)

- Doncaster A DHQ ☐
- Doncaster A1 sub-division ☐
- Doncaster A2 sub-division ☐
- Doncaster A3 sub-division ☐
- Rotherham C DHQ ☐
- Rotherham C1 sub-division ☐
- Rotherham C2 sub-division ☐
- Rotherham C3 sub-division ☐
- Barnsley B DHQ ☐
- Barnsley B1 sub-division ☐
- Barnsley B2 sub-division ☐
- Sheffield Road Traffic ☐
- Rotherham Road Traffic ☐
- Doncaster Road Traffic ☐
- Barnsley Road Traffic ☐
- Sheffield North F DHQ ☐
- Sheffield North F1 sub-division ☐
- Sheffield North F2 sub-division ☐
- Sheffield North F3 sub-division ☐
- Sheffield South E DHQ ☐
- Sheffield South E1 sub-division ☐
- Sheffield South E2 sub-division ☐
- Sheffield South E3 sub-division ☐
- Headquarters Buildings (inc Heeley, Richfield Hse, Castle Green, Escafeld Hse, R.C.S., Training) ☐
- Operations ☐

33. Are you a:

Police Constable	☐	Principal Officer/Senior Officer	☐
Police Sergeant	☐	Scale 4 - 6	☐
Inspector	☐	Scale 1 - 3	☐
Chief Inspector/Superintendent	☐	Hourly Paid Work member of staff	☐
More senior than Superintendent	☐	Traffic Warden	☐

YOUR COMMENTS

If you have any further comments you would like to make, please use the space be (add additional sheets if necessary).

....

....

....

....

....

....

....

....

....

....

....

....

....

THANK YOU FOR HELPING WITH THIS SURVEY

Please return the questionnaire in the envelope provided.

Appendix 2

OMORALE respondents own morale (3 cats)
by NEWCOM Communication with more senior officers/
Controlling for..
NEWPROM Promotions are given to those who earn t Value = 1.00 Agree

NEWCOM — Page 1 of 1

OMORALE	Count / Row Pct / Col Pct	Good 1.00	Neither good nor 2.00	Bad 3.00	Row Total
higher	1.00	29	10	7	46
		63.0	21.7	15.2	40.7
		39.7	37.0	53.8	
neither nor low	2.00	33	12	4	49
		67.3	24.5	8.2	43.4
		45.2	44.4	30.8	
lower	3.00	11	5	2	18
		61.1	27.8	11.1	15.9
		15.1	18.5	15.4	
	Column Total	73	27	13	113
		64.6	23.9	11.5	100.0

Chi-Square	Value	DF	Significance
Pearson	1.34783	4	.85321
Likelihood Ratio	1.34779	4	.85322
Mantel-Haenszel test for linear association	.12640	1	.72219

Minimum Expected Frequency - 2.071
Cells with Expected Frequency < 5 - 2 OF 9 (22.2%)

OMORALE respondents own morale (3 cats)
by NEWCOM Communication with more senior officers/
Controlling for..
NEWPROM Promotions are given to those who earn Value = 2.00 Neither agree nor dis

NEWCOM Page 1 of 1

OMORALE	Count / Row Pct / Col Pct	Good 1.00	Neither good nor 2.00	Bad 3.00	Row Total
higher	1.00	33	10	4	47
		70.2	21.3	8.5	33.1
		41.3	23.3	21.1	
neither nor low	2.00	28	20	3	51
		54.9	39.2	5.9	35.9
		35.0	46.5	15.8	
lower	3.00	19	13	12	44
		43.2	29.5	27.3	31.0
		23.8	30.2	63.2	
	Column Total	80	43	19	142
		56.3	30.3	13.4	100.0

Chi-Square	Value	DF	Significance
Pearson	14.90967	4	.00489
Likelihood Ratio	14.16105	4	.00680
Mantel-Haenszel test for linear association	9.13666	1	.00251

Minimum Expected Frequency - 5.887

OMORALE respondents own morale (3 cats)
by NEWCOM Communication with more senior officers/
Controlling for..
NEWPROM Promotions are given to those who earn t Value = 3.00 Disagree

NEWCOM — Page 1 of 1

OMORALE	Count / Row Pct / Col Pct	Good 1.00	Neither good nor 2.00	Bad 3.00	Row Total
higher	1.00	21	8	3	32
		65.6	25.0	9.4	20.4
		28.4	17.4	8.1	
neither nor low	2.00	27	17	8	52
		51.9	32.7	15.4	33.1
		36.5	37.0	21.6	
lower	3.00	26	21	26	73
		35.6	28.8	35.6	46.5
		35.1	45.7	70.3	
	Column Total	74	46	37	157
		47.1	29.3	23.6	100.0

Chi-Square	Value	DF	Significance
Pearson	13.75161	4	.00813
Likelihood Ratio	14.09326	4	.00700
Mantel-Haenszel test for linear association	12.37752	1	.00043

Minimum Expected Frequency - 7.541

Number of Missing Observations: 55

Appendix 3a

```
/***********************************************************
 * John Gretton - SSRC / CMS: Sheffield Hallam University  *
 *                                                         *
 * SAS programs - MPhil in Ordinal Data Analysis           *
 *                                                         *
 * Proportional odds using dummy variables                 *
 *                                                         *
 * Main skeleton incl. code for Lipsitz g-o-f              *
 ***********************************************************/

data syp;

options nocenter ls=80 pagesize=80;

infile 'c:\syp\final.dat' missover;

input pubview 1 service 2 commimm 3 commsen 4 promearn 5 whouknow 6 omor 7
      cmor 8 gender 9 lenserv 10 officer 11 borw 12 linemgr 13 relmoral 14
      jobsat 15 proud 16 ;

label
pubview = 'perceived public view of SYP'
commsen ='communication with senior mgrs/officers'
promearn = 'promotions given to those who earn them'
omor = 'respondents own morale'
lenserv  ='years in the police service'
officer  ='civilian staff or police officer'
jobsat ='satisfied with job yes/no'

if omor=. then delete;
if jobsat=. then delete;
if commsen=. then delete;
if pubview=. then delete;
if promearn=. then delete;
if lenserv=. then delete;
if officer=. then delete;

run;

proc format;

value pubfmt 1='Very Positive'
             2='Positive'
             3='Neither'
             4='Negative'
             5='Very Negative';

value servfmt 1='V Satisfied'
              2='Satisfied'
              3='Neither'
              4='Dissatisfied'
              5='V Dissatisfied';

value comfmt 1='Very Good'
             2='Good'
             3='Neither'
             4='Bad'
             5='Very Bad';

value morfmt 1='Very High'
             2='High'
             3='Neither'
             4='Low'
             5='Very Low';

value promfmt 1='Strongly Agree'
              2='Agree'
              3='Neither'
              4='Disagree'
              5='Strongly Disagree';
```

```
value jobfmt 1='Yes'
             2='No';

value lenfmt 1='less than 2 yrs'
             2='2 - 5 yrs'
             3='6 - 10 yrs'
             4='11 - 20 yrs'
             5='21 + yrs';

value offfmt 1='Civilian Staff'
             2='Police Officer';

run;

data prop;
set syp;

/* Recode omor to 3 categories */

if omor=1 or omor=2 then newomor=1;
if omor=3 then newomor=2;
if omor=4 or omor=5 then newomor=3;

/* dichotomise response to fit bin. logistic model
   if p.o. assumption fails */

if newomor=1 then binomor1=1;
if newomor>1 then binomor1=2;

if newomor<3 then binomor2=1;
if newomor=3 then binomor2=2;


/* Recode covariates into dummy variables */

if commsen ne . then com1=0;if commsen ne . then com2=0;
if commsen ne . then com3=0;if commsen ne . then com4=0;

if lenserv ne . then len1=0;if lenserv ne . then len2=0;
if lenserv ne . then len3=0;if lenserv ne . then len4=0;

if promearn ne . then prom1=0;if promearn ne . then prom2=0;
if promearn ne . then prom3=0;if promearn ne . then prom4=0;

if pubview ne . then pub1=0;if pubview ne . then pub2=0;
if pubview ne . then pub3=0;if pubview ne . then pub4=0;

if jobsat ne . then jobnew=0;

if officer ne . then offnew=0;

if lenserv=1 then len1=1;
if lenserv=2 then len2=1;
if lenserv=3 then len3=1;
if lenserv=4 then len4=1;

if commsen=1 then com1=1;
if commsen=2 then com2=1;
if commsen=3 then com3=1;
if commsen=4 then com4=1;

if promearn=1 then prom1=1;
if promearn=2 then prom2=1;
if promearn=3 then prom3=1;
if promearn=4 then prom4=1;

if pubview=1 then pub1=1;
if pubview=2 then pub2=1;
if pubview=3 then pub3=1;
if pubview=4 then pub4=1;
```

```
if jobsat=1 then jobnew=1;

if officer=1 then offnew=1;

/* Reduce covariate categories where necessary */

newpub1=0;newpub2=0;
If pubview=1 or pubview=2 then newpub1=1;
if pubview=3 then newpub2=1;


newcom1=0;newcom2=0;
if commsen=1 or commsen=2 then newcom1=1;
if commsen=3 then newcom2=1;

newprom1=0;newprom2=0;
if promearn=1 or promearn=2 then newprom1=1;
if promearn=3 then newprom2=1;

/* recode lenserv to binary */

if lenserv=1 then lendum=1;
if lenserv>1 then lendum=0;

run;


/* proportional odds model: newomor vs dummy vars */

title1 'newomor vs dummies';

proc logistic data=prop;
model newomor=jobnew com1 com2 com3 com4 pub1 pub2 pub3 pub4 prom1 prom2
              prom3 prom4 lendum offnew;

output out=lipsitz p=cump;
run;
quit;

/* *** Goodness of Fit *** */

/* Lipsitz mean scores */

data fitzmaur;

retain tot 0 cump1 cump2 ;

set lipsitz;

tot=tot+cump;

if _level_=1 then cump1=cump;
if _level_=2 then cump2=cump;

if mod(_n_,2)=0 then do;  /* if _level_=2 then */

p1=cump1;
p2=(cump2-cump1);
p3=(1-cump2);

score=3-tot;

output;tot=0;end;

run;

/* order data by mean score */
proc sort data=fitzmaur;
by score;
run;
```

```
/* list scores to partition data without separating ties */

proc freq data=fitzmaur;
tables score;
run;

/* data partitioning into g groups */

data molenbrg;
set fitzmaur;

no=_n_;

i1=0;i2=0;i3=0;i4=0;i5=0;i6=0;i7=0;i8=0;i9=0;i10=0;
y1=0;y2=0;y3=0;

/* N from first model used to calculate desired size of groups */

if no<154 then i1=1;
if no>153 and no<267 then i2=1;
if no>266 and no<427 then i3=1;
if no>426 and no<548 then i4=1;
if no>547 and no<723 then i5=1;
if no>722 and no<879 then i6=1;
if no>878 and no<1018 then i7=1;
if no>1017 and no<1154 then i8=1;
if no>1153 and no<1295 then i9=1;
if no>1294 then i10=1;
run;

/* Refit model with group indicators */

title1 'newomor vs dummies';

proc logistic data=molenbrg;
model newomor=jobnew com1 com2 com3 com4 pub1 pub2 pub3 pub4 prom1 prom2
              prom3 prom4 lendum offnew i1-i9;

run;


/* calculate observed and expected frequencies
   within each group and response level*/

data lip2;
set molenbrg;

y1=(newomor=1);
y2=(newomor=2);
y3=(newomor=3);

if i1=1 then g=1;
if i2=1 then g=2;
if i3=1 then g=3;
if i4=1 then g=4;
if i5=1 then g=5;
if i6=1 then g=6;
if i7=1 then g=7;
if i8=1 then g=8;
if i9=1 then g=9;
if i10=1 then g=10;

p1=cump1;
p2=(cump2-cump1);
p3=(1-cump2);

run;

data lip3;
set lip2;

e11=p1*i1;
```

```
e21=p2*i1;
e31=p3*i1;
e12=p1*i2;
e22=p2*i2;
e32=p3*i2;
e13=p1*i3;
e23=p2*i3;
e33=p3*i3;
e14=p1*i4;
e24=p2*i4;
e34=p3*i4;
e15=p1*i5;
e25=p2*i5;
e35=p3*i5;
e16=p1*i6;
e26=p2*i6;
e36=p3*i6;
e17=p1*i7;
e27=p2*i7;
e37=p3*i7;
e18=p1*i8;
e28=p2*i8;
e38=p3*i8;
e19=p1*i9;
e29=p2*i9;
e39=p3*i9;
e110=p1*i10;
e210=p2*i10;
e310=p3*i10;

o11=y1*i1;
o21=y2*i1;
o31=y3*i1;
o12=y1*i2;
o22=y2*i2;
o32=y3*i2;
o13=y1*i3;
o23=y2*i3;
o33=y3*i3;
o14=y1*i4;
o24=y2*i4;
o34=y3*i4;
o15=y1*i5;
o25=y2*i5;
o35=y3*i5;
o16=y1*i6;
o26=y2*i6;
o36=y3*i6;
o17=y1*i7;
o27=y2*i7;
o37=y3*i7;
o18=y1*i8;
o28=y2*i8;
o38=y3*i8;
o19=y1*i9;
o29=y2*i9;
o39=y3*i9;
o110=y1*i10;
o210=y2*i10;
o310=y3*i10;

etot1=(e11+e12+e13+e14+e15+e16+e17+e18+e19+e110);
otot1=(o11+o12+o13+o14+o15+o16+o17+o18+o19+o110);
etot2=(e21+e22+e23+e24+e25+e26+e27+e28+e29+e210);
otot2=(o21+o22+o23+o24+o25+o26+o27+o28+o29+o210);
etot3=(e31+e32+e33+e34+e35+e36+e37+e38+e39+e310);
otot3=(o31+o32+o33+o34+o35+o36+o37+o38+o39+o310);

run;

proc sort data=lip3;
by g;
```

```
run;

proc summary data=lip3 print n mean sum;
var etot1-etot3 otot1-otot3;
by g;

run;

/* Diagnostic plot of scores v response */

title1 'Model main effects dummies:';
title3 'Observed response vs Predicted mean score';

proc plot data=molenbrg;
plot newomor*score;
run;
quit;

/* Phew - hope it all works! */
```

Appendix 3b

```
/**********************************************************
 * John Gretton - SSRC / CMS: Sheffield Hallam University *
 *                                                        *
 * SAS programs - MPhil in Ordinal Data Analysis          *
 *                                                        *
 * Proportional odds using scored variables               *
 *                                                        *
 * Main skeleton incl. code for Lipsitz g-o-f             *
 **********************************************************/

data syp;

options nocenter ls=80 pagesize=80;

infile 'c:\syp\final.dat' missover;

input pubview 1 service 2 commimm 3 commsen 4 promearn 5 whouknow 6 omor 7
      cmor 8 gender 9 lenserv 10 officer 11 borw 12 linemgr 13 relmoral 14
      jobsat 15 proud 16 ;

label
pubview = 'perceived public view of SYP'
commsen ='communication with senior mgrs/officers'
promearn = 'promotions given to those who earn them'
omor = 'respondents own morale'
lenserv  ='years in the police service'
officer  ='civilian staff or police officer'
jobsat ='satisfied with job yes/no'

if omor=. then delete;
if jobsat=. then delete;
if commsen=. then delete;
if pubview=. then delete;
if promearn=. then delete;
if lenserv=. then delete;
if officer=. then delete;

run;

proc format;

value pubfmt 1='Very Positive'
             2='Positive'
             3='Neither'
             4='Negative'
             5='Very Negative';

value servfmt 1='V Satisfied'
              2='Satisfied'
              3='Neither'
              4='Dissatisfied'
              5='V Dissatisfied';

value comfmt 1='Very Good'
             2='Good'
             3='Neither'
             4='Bad'
             5='Very Bad';

value morfmt 1='Very High'
             2='High'
             3='Neither'
             4='Low'
             5='Very Low';

value promfmt 1='Strongly Agree'
              2='Agree'
              3='Neither'
              4='Disagree'
              5='Strongly Disagree';
```

```
value jobfmt 1='Yes'
             2='No';

value lenfmt 1='less than 2 yrs'
             2='2 - 5 yrs'
             3='6 - 10 yrs'
             4='11 - 20 yrs'
             5='21 + yrs';

value offfmt 1='Civilian Staff'
             2='Police Officer';

run;

data prop;
set syp;

/* Recode omor to 3 categories */

if omor=1 or omor=2 then newomor=1;
if omor=3 then newomor=2;
if omor=4 or omor=5 then newomor=3;

/* dichotomise response to fit bin. logistic model
   if p.o. assumption fails */

if newomor=1 then binomor1=1;
if newomor>1 then binomor1=2;

if newomor<3 then binomor2=1;
if newomor=3 then binomor2=2;

/* Recode jobsat and officer */

if jobsat ne . then jobnew=0;

if officer ne . then offnew=0;

if jobsat=1 then jobnew=1;

if officer=1 then offnew=1;

/* Recode covariates com, prom, pub into CHAID scored cats */

if commsen=1 then comchd=0;
if commsen=2 then comchd=23.61;
if commsen=3 then comchd=40.49;
if commsen=4 then comchd=62.63;
if commsen=5 then comchd=100;

if promearn=1 then promchd=0;
if promearn=2 then promchd=47.07;
if promearn=3 then promchd=63.12;
if promearn=4 then promchd=86.49;
if promearn=5 then promchd=100;

if pubview=1 then pubchd=0;
if pubview=2 then pubchd=13;
if pubview=3 then pubchd=18.89;
if pubview=4 then pubchd=22.13;
if pubview=5 then pubchd=100;

/* Recode covariates to 3 cats with CHAID scores */

pub3chd=0;
If pubview=1 or pubview=2 then pub3chd=1;
if pubview=3 then newpub2=62.1;
If pubview=4 or pubview=5 then pub3chd=100;

com3chd=0;
```

```
if commsen=1 or commsen=2 then com3chd=1;
if commsen=3 then com3chd=45.85;
if commsen=4 or commsen=5 then com3chd=100;

prom3chd=0;
if promearn=1 or promearn=2 then prom3chd=1;
if promearn=3 then prom3chd=40.87;
if promearn=4 or promearn=5 then prom3chd=100;

/* collapse predictors to 3 cats (integer scores) */

if commsen=1 or commsen=2 then newcom=1;
if commsen=3 then newcom=2;
if commsen=4 or commsen=5 then newcom=3;

if promearn=1 or promearn=2 then newprom=1;
if promearn=3 then newprom=2;
if promearn=4 or promearn=5 then newprom=3;

if pubview=1 or pubview=2 then newpub=1;
if pubview=3 then newpub=2;
if pubview=4 or pubview=5 then newpub=3;

/* recode lenserv to binary */

if lenserv=1 then lendum=1;
if lenserv>1 then lendum=0;

run;


/* proportional odds model: newomor vs scored vars */

title1 'newomor vs scored vars (3 cat integers) ';

proc logistic data=prop;
model newomor=jobnew newcom newpub newprom lendum offnew;

output out=lipsitz p=cump;
run;
quit;


/* *** Goodness of Fit *** */

/* Lipsitz mean scores */

data fitzmaur;

retain tot 0 cump1 cump2 ;

set lipsitz;

tot=tot+cump;

if _level_=1 then cump1=cump;
if _level_=2 then cump2=cump;

if mod(_n_,2)=0 then do;  /* if _level_=2 then */

p1=cump1;
p2=(cump2-cump1);
p3=(1-cump2);

score=3-tot;

output;tot=0;end;

run;

/* order data by mean score */
proc sort data=fitzmaur;
```

```
by score;
run;

/* list scores to partition data without separating ties */

proc freq data=fitzmaur;
tables score;
run;

/* data partitioning into g groups */

data molenbrg;
set fitzmaur;

no=_n_;

i1=0;i2=0;i3=0;i4=0;i5=0;i6=0;i7=0;i8=0;i9=0;i10=0;
y1=0;y2=0;y3=0;

/* N from first model used to calculate desired size of groups */

if no<154 then i1=1;
if no>153 and no<267 then i2=1;
if no>266 and no<427 then i3=1;
if no>426 and no<548 then i4=1;
if no>547 and no<723 then i5=1;
if no>722 and no<879 then i6=1;
if no>878 and no<1018 then i7=1;
if no>1017 and no<1154 then i8=1;
if no>1153 and no<1295 then i9=1;
if no>1294 then i10=1;
run;

/* Refit model with group indicators */

title1 'newomor vs scored vars (3 cat integers) + g-o-f';

proc logistic data=prop;
model newomor=jobnew newcom newpub newprom lendum offnew i1-i9;

run;
quit;


/* calculate observed and expected frequencies
   within each group and response level */

data lip2;
set molenbrg;

y1=(newomor=1);
y2=(newomor=2);
y3=(newomor=3);

if i1=1 then g=1;
if i2=1 then g=2;
if i3=1 then g=3;
if i4=1 then g=4;
if i5=1 then g=5;
if i6=1 then g=6;
if i7=1 then g=7;
if i8=1 then g=8;
if i9=1 then g=9;
if i10=1 then g=10;

p1=cump1;
p2=(cump2-cump1);
p3=(1-cump2);

run;

data lip3;
```

```
set lip2;

e11=p1*i1;
e21=p2*i1;
e31=p3*i1;
e12=p1*i2;
e22=p2*i2;
e32=p3*i2;
e13=p1*i3;
e23=p2*i3;
e33=p3*i3;
e14=p1*i4;
e24=p2*i4;
e34=p3*i4;
e15=p1*i5;
e25=p2*i5;
e35=p3*i5;
e16=p1*i6;
e26=p2*i6;
e36=p3*i6;
e17=p1*i7;
e27=p2*i7;
e37=p3*i7;
e18=p1*i8;
e28=p2*i8;
e38=p3*i8;
e19=p1*i9;
e29=p2*i9;
e39=p3*i9;
e110=p1*i10;
e210=p2*i10;
e310=p3*i10;

o11=y1*i1;
o21=y2*i1;
o31=y3*i1;
o12=y1*i2;
o22=y2*i2;
o32=y3*i2;
o13=y1*i3;
o23=y2*i3;
o33=y3*i3;
o14=y1*i4;
o24=y2*i4;
o34=y3*i4;
o15=y1*i5;
o25=y2*i5;
o35=y3*i5;
o16=y1*i6;
o26=y2*i6;
o36=y3*i6;
o17=y1*i7;
o27=y2*i7;
o37=y3*i7;
o18=y1*i8;
o28=y2*i8;
o38=y3*i8;
o19=y1*i9;
o29=y2*i9;
o39=y3*i9;
o110=y1*i10;
o210=y2*i10;
o310=y3*i10;

etot1=(e11+e12+e13+e14+e15+e16+e17+e18+e19+e110);
otot1=(o11+o12+o13+o14+o15+o16+o17+o18+o19+o110);
etot2=(e21+e22+e23+e24+e25+e26+e27+e28+e29+e210);
otot2=(o21+o22+o23+o24+o25+o26+o27+o28+o29+o210);
etot3=(e31+e32+e33+e34+e35+e36+e37+e38+e39+e310);
otot3=(o31+o32+o33+o34+o35+o36+o37+o38+o39+o310);

run;
```

```
proc sort data=lip3;
by g;
run;

proc summary data=lip3 print n mean sum;
var etot1-etot3 otot1-otot3;
by g;

run;

/* Diagnostic plot of scores v response */

title1 'Model main effects scores';
title3 'Observed response vs Predicted mean score';

proc plot data=molenbrg;
plot newomor*score;
run;
quit;

/* Phew - hope it all works! */
```

Appendix 3c

```
/*********************************************************
 * John Gretton - SSRC / CMS: Sheffield Hallam University  *
 *                                                        *
 * SAS programs - MPhil in Ordinal Data Analysis           *
 *                                                        *
 * Continuation odds using dummy variables                 *
 *                                                        *
 * Main skeleton incl. code for Lipsitz g-o-f              *
 *********************************************************/

data syp;

options nocenter ls=80 pagesize=80;

infile 'c:\syp\final.dat' missover;

input pubview 1 service 2 commimm 3 commsen 4 promearn 5 whouknow 6 omor 7
      cmor 8 gender 9 lenserv 10 officer 11 borw 12 linemgr 13 relmoral 14
      jobsat 15 proud 16 ;

label
pubview = 'perceived public view of SYP'
commsen ='communication with senior mgrs/officers'
promearn = 'promotions given to those who earn them'
omor = 'respondents own morale'
lenserv  ='years in the police service'
officer  ='civilian staff or police officer'
jobsat ='satisfied with job yes/no'

if omor=. then delete;
if jobsat=. then delete;
if commsen=. then delete;
if pubview=. then delete;
if promearn=. then delete;
if lenserv=. then delete;
if officer=. then delete;

run;

proc format;

value pubfmt 1='Very Positive'
             2='Positive'
             3='Neither'
             4='Negative'
             5='Very Negative';

value servfmt 1='V Satisfied'
              2='Satisfied'
              3='Neither'
              4='Dissatisfied'
              5='V Dissatisfied';

value comfmt 1='Very Good'
             2='Good'
             3='Neither'
             4='Bad'
             5='Very Bad';

value morfmt 1='Very High'
             2='High'
             3='Neither'
             4='Low'
             5='Very Low';

value promfmt 1='Strongly Agree'
              2='Agree'
              3='Neither'
              4='Disagree'
              5='Strongly Disagree';
```

```
value jobfmt 1='Yes'
             2='No';

value lenfmt 1='less than 2 yrs'
             2='2 - 5 yrs'
             3='6 - 10 yrs'
             4='11 - 20 yrs'
             5='21 + yrs';

value offfmt 1='Civilian Staff'
             2='Police Officer';

run;

data cont;
set syp;

/* Recode omor to 3 categories */

if omor=1 or omor=2 then newomor=1;
if omor=3 then newomor=2;
if omor=4 or omor=5 then newomor=3;

/* dichotomise response to fit bin. logistic model
   if p.o. assumption fails */

if newomor=1 then binomor1=1;
if newomor>1 then binomor1=2;

if newomor<3 then binomor2=1;
if newomor=3 then binomor2=2;


/* Recode covariates into dummy variables */

if commsen ne . then com1=0;if commsen ne . then com2=0;
if commsen ne . then com3=0;if commsen ne . then com4=0;

if lenserv ne . then len1=0;if lenserv ne . then len2=0;
if lenserv ne . then len3=0;if lenserv ne . then len4=0;

if promearn ne . then prom1=0;if promearn ne . then prom2=0;
if promearn ne . then prom3=0;if promearn ne . then prom4=0;

if pubview ne . then pub1=0;if pubview ne . then pub2=0;
if pubview ne . then pub3=0;if pubview ne . then pub4=0;

if jobsat ne . then jobnew=0;

if officer ne . then offnew=0;

if lenserv=1 then len1=1;
if lenserv=2 then len2=1;
if lenserv=3 then len3=1;
if lenserv=4 then len4=1;

if commsen=1 then com1=1;
if commsen=2 then com2=1;
if commsen=3 then com3=1;
if commsen=4 then com4=1;

if promearn=1 then prom1=1;
if promearn=2 then prom2=1;
if promearn=3 then prom3=1;
if promearn=4 then prom4=1;

if pubview=1 then pub1=1;
if pubview=2 then pub2=1;
if pubview=3 then pub3=1;
if pubview=4 then pub4=1;
```

```
if jobsat=1 then jobnew=1;

if officer=1 then offnew=1;

/* Reduce covariate categories where necessary */

newpub1=0;newpub2=0;
If pubview=1 or pubview=2 then newpub1=1;
if pubview=3 then newpub2=1;


newcom1=0;newcom2=0;
if commsen=1 or commsen=2 then newcom1=1;
if commsen=3 then newcom2=1;

newprom1=0;newprom2=0;
if promearn=1 or promearn=2 then newprom1=1;
if promearn=3 then newprom2=1;

/* recode lenserv to binary */

if lenserv=1 then lendum=1;
if lenserv>1 then lendum=0;

/* set up cutpt for cont odds */

l1=1;
l2=2;

run;
quit;

data cont2;
set cont;

array a(i) l1 l2;
do over a;
cutpt=a;
output;
end;

/* set up ind for cont odds */

data cont3;
set cont2;

if newomor=1 and cutpt=1 then ind=0;
if newomor=2 and cutpt=1 then ind=1;
if newomor=3 and cutpt=1 then ind=1;

if newomor=1 and cutpt=2 then delete;
if newomor=2 and cutpt=2 then ind=0;
if newomor=3 and cutpt=2 then ind=1;

run;
quit;

/* cont odds model using proc logistic
   newomor vs dummies*/

title1 'cont odds using dummy vars';

proc logistic data=cont3;
model ind=cutpt jobnew com1 com2 com3 com4 pub1 pub2 pub3 pub4
          prom1 prom2 prom3 prom4 lendum offnew;

run;

/* Binary logistic models to check global odds ratios assumption */

title1 'bin log 1 for cont odds: dummies';
```

```
proc logistic data=cont;
model binomor1=jobnew com1 com2 com3 com4 pub1 pub2 pub3 pub4
          prom1 prom2 prom3 prom4 lendum offnew;

run;

title1 'bin log 2 for cont odds: dummies';

proc logistic data=cont;
model binomor2=jobnew com1 com2 com3 com4 pub1 pub2 pub3 pub4
          prom1 prom2 prom3 prom4 lendum offnew;

run;

/* *** Goodness of Fit *** */

/* In order to get individual probabilities for response categories
   the model parameters must be fed back in to the dataset to create the
   probability of being in response cats 1, 2 and 3 for each respondent */


data probs;
set cont;

k1=-3.3449+(1.6987*jobnew)+(0.8453*newcom1)+(0.4441*newcom2)
   +(0.6959*newprom1)+(0.4068*newprom2)+(1.0153*newpub1)+(0.4409*newpub2)
   +(1.2351*lenbin2)+(-0.3585*offnew);

k2=-2.2725+(1.6987*jobnew)+(0.8453*newcom1)+(0.4441*newcom2)
   +(0.6959*newprom1)+(0.4068*newprom2)+(1.0153*newpub1)+(0.4409*newpub2)
   +(1.2351*lenbin2)+(-0.3585*offnew);

p1=exp(k1)/(1+exp(k1));
p3=1/((1+exp(k1))*(1+exp(k2)));
p2=(1-p1-p3);

score=p1+(2*p2)+(3*p3);
run;


/* order data by mean score */

proc sort data=probs;
by score;
run;

/* list scores to partition data without separating ties */

proc freq data=probs;
tables score;
run;

/* data partitioning into g groups */

data molenbrg;
set probs;

no=_n_;

i1=0;i2=0;i3=0;i4=0;i5=0;i6=0;i7=0;i8=0;i9=0;i10=0;
y1=0;y2=0;y3=0;

/* N from first model used to calculate desired size of groups */

if no<154 then i1=1;
if no>153 and no<267 then i2=1;
if no>266 and no<427 then i3=1;
if no>426 and no<548 then i4=1;
if no>547 and no<723 then i5=1;
if no>722 and no<879 then i6=1;
if no>878 and no<1018 then i7=1;
```

```
if no>1017 and no<1154 then i8=1;
if no>1153 and no<1295 then i9=1;
if no>1294 then i10=1;
run;

/* Refit model with group indicators */

title1 'cont odds using dummy vars + g-o-f';

proc logistic data=cont3;
model ind=cutpt jobnew com1 com2 com3 com4 pub1 pub2 pub3 pub4
          prom1 prom2 prom3 prom4 lendum offnew i1-i9;

run;

/* calculate observed and expected frequencies
   within each group and response level*/

data lip2;
set molenbrg;

y1=(newomor=1);
y2=(newomor=2);
y3=(newomor=3);

if i1=1 then g=1;
if i2=1 then g=2;
if i3=1 then g=3;
if i4=1 then g=4;
if i5=1 then g=5;
if i6=1 then g=6;
if i7=1 then g=7;
if i8=1 then g=8;
if i9=1 then g=9;
if i10=1 then g=10;

p1=cump1;
p2=(cump2-cump1);
p3=(1-cump2);

run;

data lip3;
set lip2;

e11=p1*i1;
e21=p2*i1;
e31=p3*i1;
e12=p1*i2;
e22=p2*i2;
e32=p3*i2;
e13=p1*i3;
e23=p2*i3;
e33=p3*i3;
e14=p1*i4;
e24=p2*i4;
e34=p3*i4;
e15=p1*i5;
e25=p2*i5;
e35=p3*i5;
e16=p1*i6;
e26=p2*i6;
e36=p3*i6;
e17=p1*i7;
e27=p2*i7;
e37=p3*i7;
e18=p1*i8;
e28=p2*i8;
e38=p3*i8;
e19=p1*i9;
e29=p2*i9;
e39=p3*i9;
```

```
e110=p1*i10;
e210=p2*i10;
e310=p3*i10;

o11=y1*i1;
o21=y2*i1;
o31=y3*i1;
o12=y1*i2;
o22=y2*i2;
o32=y3*i2;
o13=y1*i3;
o23=y2*i3;
o33=y3*i3;
o14=y1*i4;
o24=y2*i4;
o34=y3*i4;
o15=y1*i5;
o25=y2*i5;
o35=y3*i5;
o16=y1*i6;
o26=y2*i6;
o36=y3*i6;
o17=y1*i7;
o27=y2*i7;
o37=y3*i7;
o18=y1*i8;
o28=y2*i8;
o38=y3*i8;
o19=y1*i9;
o29=y2*i9;
o39=y3*i9;
o110=y1*i10;
o210=y2*i10;
o310=y3*i10;

etot1=(e11+e12+e13+e14+e15+e16+e17+e18+e19+e110);
otot1=(o11+o12+o13+o14+o15+o16+o17+o18+o19+o110);
etot2=(e21+e22+e23+e24+e25+e26+e27+e28+e29+e210);
otot2=(o21+o22+o23+o24+o25+o26+o27+o28+o29+o210);
etot3=(e31+e32+e33+e34+e35+e36+e37+e38+e39+e310);
otot3=(o31+o32+o33+o34+o35+o36+o37+o38+o39+o310);

run;

proc sort data=lip3;
by g;
run;

proc summary data=lip3 print n mean sum;
var etot1-etot3 otot1-otot3;
by g;

run;

/* Diagnostic plot of scores v response */

title1 'CO Model main effects dummies:';
title3 'Observed response vs Predicted mean score';

proc plot data=molenbrg;
plot newomor*score;
run;
quit;

/* Phew - hope it all works! */
```

Appendix 3d

```
/*********************************************************
 * John Gretton - SSRC / CMS: Sheffield Hallam University  *
 *                                                          *
 * SAS programs - MPhil in Ordinal Data Analysis            *
 *                                                          *
 * Continuation odds using dummy variables                  *
 *                                                          *
 * Main skeleton incl. code for Lipsitz g-o-f               *
 *********************************************************/

data syp;

options nocenter ls=80 pagesize=80;

infile 'c:\syp\final.dat' missover;

input pubview 1 service 2 commimm 3 commsen 4 promearn 5 whouknow 6 omor 7
      cmor 8 gender 9 lenserv 10 officer 11 borw 12 linemgr 13 relmoral 14
      jobsat 15 proud 16 ;

label
pubview = 'perceived public view of SYP'
commsen ='communication with senior mgrs/officers'
promearn = 'promotions given to those who earn them'
omor = 'respondents own morale'
lenserv  ='years in the police service'
officer  ='civilian staff or police officer'
jobsat ='satisfied with job yes/no'

if omor=. then delete;
if jobsat=. then delete;
if commsen=. then delete;
if pubview=. then delete;
if promearn=. then delete;
if lenserv=. then delete;
if officer=. then delete;

run;

proc format;

value pubfmt 1='Very Positive'
             2='Positive'
             3='Neither'
             4='Negative'
             5='Very Negative';

value servfmt 1='V Satisfied'
              2='Satisfied'
              3='Neither'
              4='Dissatisfied'
              5='V Dissatisfied';

value comfmt 1='Very Good'
             2='Good'
             3='Neither'
             4='Bad'
             5='Very Bad';

value morfmt 1='Very High'
             2='High'
             3='Neither'
             4='Low'
             5='Very Low';

value promfmt 1='Strongly Agree'
              2='Agree'
              3='Neither'
              4='Disagree'
              5='Strongly Disagree';
```

```
value jobfmt 1='Yes'
             2='No';

value lenfmt 1='less than 2 yrs'
             2='2 - 5 yrs'
             3='6 - 10 yrs'
             4='11 - 20 yrs'
             5='21 + yrs';

value offfmt 1='Civilian Staff'
             2='Police Officer';

run;

data cont;
set syp;

/* Recode omor to 3 categories */

if omor=1 or omor=2 then newomor=1;
if omor=3 then newomor=2;
if omor=4 or omor=5 then newomor=3;

/* dichotomise response to fit bin. logistic model
   if p.o. assumption fails */

if newomor=1 then binomor1=1;
if newomor>1 then binomor1=2;

if newomor<3 then binomor2=1;
if newomor=3 then binomor2=2;

/* Recode jobsat and officer */

if jobsat ne . then jobnew=0;

if officer ne . then offnew=0;

if jobsat=1 then jobnew=1;

if officer=1 then offnew=1;

/* Recode covariates com, prom, pub into CHAID scored cats */

if commsen=1 then comchd=0;
if commsen=2 then comchd=23.61;
if commsen=3 then comchd=40.49;
if commsen=4 then comchd=62.63;
if commsen=5 then comchd=100;

if promearn=1 then promchd=0;
if promearn=2 then promchd=47.07;
if promearn=3 then promchd=63.12;
if promearn=4 then promchd=86.49;
if promearn=5 then promchd=100;

if pubview=1 then pubchd=0;
if pubview=2 then pubchd=13;
if pubview=3 then pubchd=18.89;
if pubview=4 then pubchd=22.13;
if pubview=5 then pubchd=100;

/* Recode covariates to 3 cats with CHAID scores */

pub3chd=0;
If pubview=1 or pubview=2 then pub3chd=1;
if pubview=3 then newpub2=62.1;
If pubview=4 or pubview=5 then pub3chd=100;

com3chd=0;
```

```
if commsen=1 or commsen=2 then com3chd=1;
if commsen=3 then com3chd=45.85;
if commsen=4 or commsen=5 then com3chd=100;

prom3chd=0;
if promearn=1 or promearn=2 then prom3chd=1;
if promearn=3 then prom3chd=40.87;
if promearn=4 or promearn=5 then prom3chd=100;

/* collapse predictors to 3 cats (integer scores) */

if commsen=1 or commsen=2 then newcom=1;
if commsen=3 then newcom=2;
if commsen=4 or commsen=5 then newcom=3;

if promearn=1 or promearn=2 then newprom=1;
if promearn=3 then newprom=2;
if promearn=4 or promearn=5 then newprom=3;

if pubview=1 or pubview=2 then newpub=1;
if pubview=3 then newpub=2;
if pubview=4 or pubview=5 then newpub=3;

/* recode lenserv to binary */

if lenserv=1 then lendum=1;
if lenserv>1 then lendum=0;

/* set up cutpt for cont odds */

l1=1;
l2=2;

run;
quit;

data cont2;
set cont;

array a(i) l1 l2;
do over a;
cutpt=a;
output;
end;

/* set up ind for cont odds */

data cont3;
set cont2;

if newomor=1 and cutpt=1 then ind=0;
if newomor=2 and cutpt=1 then ind=1;
if newomor=3 and cutpt=1 then ind=1;

if newomor=1 and cutpt=2 then delete;
if newomor=2 and cutpt=2 then ind=0;
if newomor=3 and cutpt=2 then ind=1;

run;
quit;

/* cont odds model using proc logistic
   newomor vs scores*/

title1 'cont odds using scores';

proc logistic data=cont3;
model ind=cutpt jobnew newcom newpub newprom lendum offnew;

run;

/* Binary logistic models to check global odds ratios assumption */
```

```
title1 'bin log 1 for cont odds: scores';

proc logistic data=cont;
model binomor1= jobnew newcom newpub newprom lendum offnew;

run;

title1 'bin log 2 for cont odds: dummies';

proc logistic data=cont;
model binomor2= jobnew newcom newpub newprom lendum offnew;

run;

/* *** Goodness of Fit *** */

/* In order to get individual probabilities for response categories
   the model parameters must be fed back in to the dataset to create the
   probability of being in response cats 1, 2 and 3 for each respondent */


data probs;
set cont;

k1=-3.3449+(1.6987*jobnew)+(0.8453*newcom1)+(0.4441*newcom2)
   +(0.6959*newprom1)+(0.4068*newprom2)+(1.0153*newpub1)+(0.4409*newpub2)
   +(1.2351*lenbin2)+(-0.3585*offnew);

k2=-2.2725+(1.6987*jobnew)+(0.8453*newcom1)+(0.4441*newcom2)
   +(0.6959*newprom1)+(0.4068*newprom2)+(1.0153*newpub1)+(0.4409*newpub2)
   +(1.2351*lenbin2)+(-0.3585*offnew);

p1=exp(k1)/(1+exp(k1));
p3=1/((1+exp(k1))*(1+exp(k2)));
p2=(1-p1-p3);

score=p1+(2*p2)+(3*p3);
run;


/* order data by mean score */

proc sort data=probs;
by score;
run;

/* list scores to partition data without separating ties */

proc freq data=probs;
tables score;
run;

/* data partitioning into g groups */

data molenbrg;
set probs;

no=_n_;

i1=0;i2=0;i3=0;i4=0;i5=0;i6=0;i7=0;i8=0;i9=0;i10=0;
y1=0;y2=0;y3=0;

/* N from first model used to calculate desired size of groups */

if no<154 then i1=1;
if no>153 and no<267 then i2=1;
if no>266 and no<427 then i3=1;
if no>426 and no<548 then i4=1;
if no>547 and no<723 then i5=1;
if no>722 and no<879 then i6=1;
if no>878 and no<1018 then i7=1;
```

```
if no>1017 and no<1154 then i8=1;
if no>1153 and no<1295 then i9=1;
if no>1294 then i10=1;
run;

/* Refit model with group indicators */

title1 'cont odds using scored vars + g-o-f';

proc logistic data=cont3;
model ind=cutpt jobnew newcom newpub newprom lendum offnew i1-i9;

run;

/* calculate observed and expected frequencies
   within each group and response level*/

data lip2;
set molenbrg;

y1=(newomor=1);
y2=(newomor=2);
y3=(newomor=3);

if i1=1 then g=1;
if i2=1 then g=2;
if i3=1 then g=3;
if i4=1 then g=4;
if i5=1 then g=5;
if i6=1 then g=6;
if i7=1 then g=7;
if i8=1 then g=8;
if i9=1 then g=9;
if i10=1 then g=10;

p1=cump1;
p2=(cump2-cump1);
p3=(1-cump2);

run;

data lip3;
set lip2;

e11=p1*i1;
e21=p2*i1;
e31=p3*i1;
e12=p1*i2;
e22=p2*i2;
e32=p3*i2;
e13=p1*i3;
e23=p2*i3;
e33=p3*i3;
e14=p1*i4;
e24=p2*i4;
e34=p3*i4;
e15=p1*i5;
e25=p2*i5;
e35=p3*i5;
e16=p1*i6;
e26=p2*i6;
e36=p3*i6;
e17=p1*i7;
e27=p2*i7;
e37=p3*i7;
e18=p1*i8;
e28=p2*i8;
e38=p3*i8;
e19=p1*i9;
e29=p2*i9;
e39=p3*i9;
e110=p1*i10;
```

```
e210=p2*i10;
e310=p3*i10;

o11=y1*i1;
o21=y2*i1;
o31=y3*i1;
o12=y1*i2;
o22=y2*i2;
o32=y3*i2;
o13=y1*i3;
o23=y2*i3;
o33=y3*i3;
o14=y1*i4;
o24=y2*i4;
o34=y3*i4;
o15=y1*i5;
o25=y2*i5;
o35=y3*i5;
o16=y1*i6;
o26=y2*i6;
o36=y3*i6;
o17=y1*i7;
o27=y2*i7;
o37=y3*i7;
o18=y1*i8;
o28=y2*i8;
o38=y3*i8;
o19=y1*i9;
o29=y2*i9;
o39=y3*i9;
o110=y1*i10;
o210=y2*i10;
o310=y3*i10;

etot1=(e11+e12+e13+e14+e15+e16+e17+e18+e19+e110);
otot1=(o11+o12+o13+o14+o15+o16+o17+o18+o19+o110);
etot2=(e21+e22+e23+e24+e25+e26+e27+e28+e29+e210);
otot2=(o21+o22+o23+o24+o25+o26+o27+o28+o29+o210);
etot3=(e31+e32+e33+e34+e35+e36+e37+e38+e39+e310);
otot3=(o31+o32+o33+o34+o35+o36+o37+o38+o39+o310);

run;

proc sort data=lip3;
by g;
run;

proc summary data=lip3 print n mean sum;
var etot1-etot3 otot1-otot3;
by g;

run;

/* Diagnostic plot of scores v response */

title1 'CO Model main effects scores:';
title3 'Observed response vs Predicted mean score';

proc plot data=molenbrg;
plot newomor*score;
run;
quit;

/* Phew - hope it all works! */
```

Appendix 4a

newomor vs dummies (model 4.1)

The LOGISTIC Procedure

Data Set: WORK.PROP
Response Variable: NEWOMOR
Response Levels: 3
Number of Observations: 1837
Link Function: Logit

Response Profile

Ordered Value	NEWOMOR	Count
1	1	762
2	2	557
3	3	518

WARNING: 6 observation(s) were deleted due to missing values for the response or explanatory variables.

Score Test for the Proportional Odds Assumption

Chi-Square = 53.5780 with 15 DF (p=0.0001)

Model Fitting Information and Testing Global Null Hypothesis BETA=0

Criterion	Intercept Only	Intercept and Covariates	Chi-Square for Covariates
AIC	4002.124	3271.044	.
SC	4013.165	3364.888	.
-2 LOG L	3998.124	3237.044	761.080 with 15 DF (p=0.0001)
Score	.	.	633.009 with 15 DF (p=0.0001)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCP1	1	-6.1969	1.0544	34.5426	0.0001	.	.
INTERCP2	1	-4.3980	1.0494	17.5627	0.0001	.	.
JOBNEW	1	1.9561	0.1211	260.8720	0.0001	0.469376	7.072
COM1	1	1.4003	0.3360	17.3635	0.0001	0.277475	4.056
COM2	1	0.9330	0.3182	8.5985	0.0034	0.250215	2.542
COM3	1	0.5851	0.3193	3.3580	0.0669	0.148048	1.795
COM4	1	0.1328	0.3345	0.1578	0.6912	0.024731	1.142
PUB1	1	4.5461	1.2619	12.9781	0.0003	0.225133	94.262
PUB2	1	3.4430	1.0168	11.4663	0.0007	0.939822	31.281
PUB3	1	2.7887	1.0153	7.5447	0.0060	0.756328	16.260
PUB4	1	2.4417	1.0196	5.7352	0.0166	0.475201	11.492
PROM1	1	0.9675	0.4024	5.7824	0.0162	0.073801	2.631
PROM2	1	0.7900	0.1928	16.7939	0.0001	0.185367	2.203
PROM3	1	0.4330	0.1801	5.7792	0.0162	0.113771	1.542
PROM4	1	0.0684	0.1810	0.1428	0.7055	0.017294	1.071
LENDUM	1	1.3162	0.1867	49.7063	0.0001	0.216438	3.729
OFFNEW	1	-0.4755	0.1247	14.5408	0.0001	-0.108156	0.622

Association of Predicted Probabilities and Observed Responses

Concordant	= 77.7%	Somers' D	= 0.567
Discordant	= 21.0%	Gamma	= 0.574
Tied	= 1.3%	Tau-a	= 0.372
(1116986 pairs)		c	= 0.783

newomor vs dummies (3 cat) (model 4.2)

The LOGISTIC Procedure

Data Set: WORK.PROP
Response Variable: NEWOMOR
Response Levels: 3
Number of Observations: 1837
Link Function: Logit

Response Profile

Ordered Value	NEWOMOR	Count
1	1	762
2	2	557
3	3	518

WARNING: 6 observation(s) were deleted due to missing values for the response or explanatory variables.

Score Test for the Proportional Odds Assumption

Chi-Square = 44.6812 with 9 DF (p=0.0001)

Model Fitting Information and Testing Global Null Hypothesis BETA=0

Criterion	Intercept Only	Intercept and Covariates	Chi-Square for Covariates
AIC	4002.124	3281.586	.
SC	4013.165	3342.309	.
-2 LOG L	3998.124	3259.586	738.538 with 9 DF (p=0.0001)
Score	.	.	622.396 with 9 DF (p=0.0001)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCP1	1	-3.6870	0.2007	337.4077	0.0001	.	.
INTERCP2	1	-1.9009	0.1845	106.1609	0.0001	.	.
JOBNEW	1	1.9615	0.1205	264.8680	0.0001	0.470663	7.110
NEWCOM1	1	0.9427	0.1426	43.7255	0.0001	0.259252	2.567
NEWCOM2	1	0.4819	0.1500	10.3194	0.0013	0.121943	1.619
NEWPUB1	1	1.0958	0.1472	55.3757	0.0001	0.299764	2.991
NEWPUB2	1	0.4222	0.1431	8.7005	0.0032	0.114513	1.525
NEWPROM1	1	0.7822	0.1251	39.0859	0.0001	0.188465	2.186
NEWPROM2	1	0.4045	0.1096	13.6140	0.0002	0.106300	1.499
LENDUM	1	1.3192	0.1856	50.5186	0.0001	0.216925	3.740
OFFNEW	1	-0.4614	0.1234	13.9873	0.0002	-0.104947	0.630

Association of Predicted Probabilities and Observed Responses

Concordant	= 77.0%	Somers' D	= 0.563
Discordant	= 20.7%	Gamma	= 0.576
Tied	= 2.3%	Tau-a	= 0.369
(1116986 pairs)		c	= 0.781

newomor vs dummies (3 cat) bin log (model 4.3a)

The LOGISTIC Procedure

Data Set: WORK.PROP
Response Variable: BINOMOR1
Response Levels: 2
Number of Observations: 1837
Link Function: Logit

Response Profile

Ordered Value	BINOMOR1	Count
1	1	762
2	2	1075

WARNING: 6 observation(s) were deleted due to missing values for the response or explanatory variables.

Model Fitting Information and Testing Global Null Hypothesis BETA=0

Criterion	Intercept Only	Intercept and Covariates	Chi-Square for Covariates
AIC	2507.039	1989.492	.
SC	2512.559	2044.694	.
-2 LOG L	2505.039	1969.492	535.547 with 9 DF (p=0.0001)
Score	.	.	450.553 with 9 DF (p=0.0001)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCPT	1	-3.3136	0.2616	160.4057	0.0001	.	.
JOBNEW	1	2.1585	0.1805	142.9354	0.0001	0.517930	8.658
NEWCOM1	1	0.8145	0.1791	20.6710	0.0001	0.223996	2.258
NEWCOM2	1	0.2214	0.1927	1.3203	0.2505	0.056031	1.248
NEWPUB1	1	0.8489	0.1769	23.0360	0.0001	0.232231	2.337
NEWPUB2	1	0.1051	0.1776	0.3501	0.5541	0.028508	1.111
NEWPROM1	1	0.6537	0.1416	21.3063	0.0001	0.157510	1.923
NEWPROM2	1	0.2097	0.1293	2.6281	0.1050	0.055098	1.233
LENDUM	1	1.4972	0.2067	52.4672	0.0001	0.246200	4.469
OFFNEW	1	-0.7720	0.1539	25.1493	0.0001	-0.175595	0.462

Association of Predicted Probabilities and Observed Responses

Concordant	= 78.5%	Somers' D	= 0.590
Discordant	= 19.4%	Gamma	= 0.603
Tied	= 2.1%	Tau-a	= 0.287
(826826 pairs)		c	= 0.795

newomor vs dummies (3 cat) bin log (model 4.3b)

The LOGISTIC Procedure

Data Set: WORK.PROP
Response Variable: BINOMOR2
Response Levels: 2
Number of Observations: 1837
Link Function: Logit

Response Profile

Ordered Value	BINOMOR2	Count
1	1	1319
2	2	518

Model Fitting Information and Testing Global Null Hypothesis BETA=0

Criterion	Intercept Only	Intercept and Covariates	Chi-Square for Covariates
AIC	2200.354	1643.278	.
SC	2205.878	1698.513	.
-2 LOG L	2198.354	1623.278	575.076 with 9 DF (p=0.0001)
Score	.	.	556.561 with 9 DF (p=0.0001)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCPT	1	-2.4265	0.2195	122.1924	0.0001	.	.
JOBNEW	1	1.9651	0.1330	218.2716	0.0001	0.471032	7.136
NEWCOM1	1	0.9810	0.1684	33.9296	0.0001	0.269788	2.667
NEWCOM2	1	0.6501	0.1756	13.7033	0.0002	0.164514	1.916
NEWPUB1	1	1.3245	0.1816	53.2118	0.0001	0.362306	3.760
NEWPUB2	1	0.7711	0.1700	20.5834	0.0001	0.209188	2.162
NEWPROM1	1	0.9124	0.1718	28.2065	0.0001	0.219930	2.490
NEWPROM2	1	0.7052	0.1416	24.8051	0.0001	0.185318	2.024
LENDUM	1	1.3679	0.2916	22.0054	0.0001	0.225706	3.927
OFFNEW	1	-0.0968	0.1552	0.3891	0.5328	-0.022038	0.908

Association of Predicted Probabilities and Observed Responses

Concordant	= 82.2%	Somers' D	= 0.659
Discordant	= 16.3%	Gamma	= 0.669
Tied	= 1.4%	Tau-a	= 0.266
(692120 pairs)		c	= 0.830

newomor vs dummies (3 cat) officers (model 4.4o)

The LOGISTIC Procedure

Data Set: WORK.PROP
Response Variable: NEWOMOR
Response Levels: 3
Number of Observations: 1438
Link Function: Logit

Response Profile

Ordered Value	NEWOMOR	Count
1	1	640
2	2	411
3	3	387

WARNING: 4 observation(s) were deleted due to missing values for the response or explanatory variables.

Score Test for the Proportional Odds Assumption

Chi-Square = 36.3461 with 8 DF (p=0.0001)

Model Fitting Information and Testing Global Null Hypothesis BETA=0

Criterion	Intercept Only	Intercept and Covariates	Chi-Square for Covariates
AIC	3097.238	2548.363	.
SC	3107.789	2601.115	.
-2 LOG L	3093.238	2528.363	564.875 with 8 DF (p=0.0001)
Score	.	.	475.430 with 8 DF (p=0.0001)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCP1	1	-3.7593	0.2347	256.4820	0.0001	.	.
INTERCP2	1	-2.0645	0.2174	90.1774	0.0001	.	.
JOBNEW	1	2.0007	0.1406	202.3836	0.0001	0.464299	7.394
NEWCOM1	1	1.1032	0.1620	46.3494	0.0001	0.303675	3.014
NEWCOM2	1	0.4951	0.1694	8.5472	0.0035	0.126011	1.641
NEWPUB1	1	1.0299	0.1766	33.9980	0.0001	0.283841	2.801
NEWPUB2	1	0.3952	0.1765	5.0132	0.0252	0.106481	1.485
NEWPROM1	1	0.8200	0.1439	32.4559	0.0001	0.196005	2.270
NEWPROM2	1	0.4279	0.1237	11.9702	0.0005	0.112650	1.534
LENDUM	1	1.6474	0.2987	30.4232	0.0001	0.212667	5.193

Association of Predicted Probabilities and Observed Responses

Concordant	= 76.3%	Somers' D	= 0.557
Discordant	= 20.6%	Gamma	= 0.575
Tied	= 3.1%	Tau-a	= 0.361
(675056 pairs)		c	= 0.779

newomor vs dummies (3 cat) civil staff (model 4.4c)

The LOGISTIC Procedure

Data Set: WORK.PROP
Response Variable: NEWOMOR
Response Levels: 3
Number of Observations: 399
Link Function: Logit

Response Profile

Ordered Value	NEWOMOR	Count
1	1	122
2	2	146
3	3	131

WARNING: 2 observation(s) were deleted due to missing values for the response or explanatory variables.

Score Test for the Proportional Odds Assumption

Chi-Square = 11.4107 with 8 DF (p=0.1795)

Model Fitting Information and Testing Global Null Hypothesis BETA=0

Criterion	Intercept Only	Intercept and Covariates	Chi-Square for Covariates
AIC	883.087	727.764	.
SC	891.075	767.704	.
-2 LOG L	879.087	707.764	170.323 with 8 DF (p=0.0001)
Score	.	.	144.929 with 8 DF (p=0.0001)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCP1	1	-4.0522	0.3918	106.9884	0.0001	.	.
INTERCP2	1	-1.8853	0.3404	30.6783	0.0001	.	.
JOBNEW	1	1.9561	0.2415	65.5972	0.0001	0.511209	7.072
NEWCOM1	1	0.3508	0.3053	1.3206	0.2505	0.096157	1.420
NEWCOM2	1	0.4110	0.3259	1.5897	0.2074	0.101779	1.508
NEWPUB1	1	1.3971	0.2883	23.4777	0.0001	0.344088	4.043
NEWPUB2	1	0.5779	0.2523	5.2467	0.0220	0.159205	1.782
NEWPROM1	1	0.6728	0.2591	6.7408	0.0094	0.166632	1.960
NEWPROM2	1	0.3271	0.2435	1.8054	0.1791	0.085496	1.387
LENDUM	1	1.2595	0.2493	25.5248	0.0001	0.298780	3.524

Association of Predicted Probabilities and Observed Responses

Concordant	= 78.8%	Somers' D	= 0.593
Discordant	= 19.5%	Gamma	= 0.604
Tied	= 1.7%	Tau-a	= 0.395
(53466 pairs)		c	= 0.797

newomor vs dummies + promcom (model 4.41c)

The LOGISTIC Procedure

Data Set: WORK.PROP
Response Variable: NEWOMOR
Response Levels: 3
Number of Observations: 399
Link Function: Logit

Response Profile

Ordered Value	NEWOMOR	Count
1	1	122
2	2	146
3	3	131

Score Test for the Proportional Odds Assumption

Chi-Square = 13.4836 with 12 DF (p=0.3349)

Model Fitting Information and Testing Global Null Hypothesis BETA=0

Criterion	Intercept Only	Intercept and Covariates	Chi-Square for Covariates
AIC	878.495	717.092	.
SC	886.473	772.937	.
-2 LOG L	874.495	689.092	185.403 with 12 DF (p=0.0001)
Score	.	.	152.120 with 12 DF (p=0.0001)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCP1	1	-4.8336	0.5115	89.2853	0.0001	.	.
INTERCP2	1	-2.5995	0.4630	31.5180	0.0001	.	.
JOBNEW	1	2.0130	0.2469	66.4491	0.0001	0.525776	7.486
NEWCOM1	1	1.1866	0.4617	6.6059	0.0102	0.325246	3.276
NEWCOM2	1	1.1346	0.5042	5.0652	0.0244	0.280672	3.110
NEWPUB1	1	1.5189	0.2963	26.2811	0.0001	0.373614	4.567
NEWPUB2	1	0.6571	0.2589	6.4437	0.0111	0.181022	1.929
NEWPROM1	1	3.2042	0.7504	18.2318	0.0001	0.794792	24.636
NEWPROM2	1	0.6897	0.6529	1.1160	0.2908	0.180153	1.993
LENBIN2	1	1.3226	0.2521	27.5346	0.0001	0.314283	3.753
PRMCOM1	1	-3.0109	0.8216	13.4285	0.0002	-0.642616	0.049
PRMCOM2	1	-0.5798	0.7295	0.6317	0.4267	-0.126316	0.560
PRMCOM3	1	-2.7050	0.9034	8.9656	0.0028	-0.375065	0.067
PRMCOM4	1	-0.4467	0.7888	0.3206	0.5712	-0.074866	0.640

Association of Predicted Probabilities and Observed Responses

Concordant	= 79.2%	Somers' D	= 0.604
Discordant	= 18.8%	Gamma	= 0.617
Tied	= 2.1%	Tau-a	= 0.403
(52920 pairs)		c	= 0.802

newomor vs dummies + promcom (model 4.41c) + g-o-f

The LOGISTIC Procedure

Data Set: WORK.MOLENBRG
Response Variable: NEWOMOR
Response Levels: 3
Number of Observations: 399
Link Function: Logit

Response Profile

Ordered Value	NEWOMOR	Count
1	1	122
2	2	146
3	3	131

Score Test for the Proportional Odds Assumption

Chi-Square = 24.5561 with 17 DF (p=0.1051)

Model Fitting Information and Testing Global Null Hypothesis BETA=0

Criterion	Intercept Only	Intercept and Covariates	Chi-Square for Covariates
AIC	878.495	724.271	.
SC	886.473	800.061	.
-2 LOG L	874.495	686.271	188.225 with 17 DF (p=0.0001)
Score	.	.	156.150 with 17 DF (p=0.0001)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCP1	1	-4.5201	0.6601	46.8937	0.0001	.	.
INTERCP2	1	-2.2505	0.6326	12.6542	0.0004	.	.
JOBNEW	1	1.6095	0.8496	3.5884	0.0582	0.420375	5.000
NEWCOM1	1	0.9082	0.6162	2.1723	0.1405	0.248939	2.480
NEWCOM2	1	0.9084	0.5783	2.4669	0.1163	0.224701	2.480
NEWPUB1	1	1.3101	0.5490	5.6943	0.0170	0.322266	3.707
NEWPUB2	1	0.6529	0.3974	2.6988	0.1004	0.179863	1.921
NEWPROM1	1	2.6592	1.3808	3.7092	0.0541	0.659614	14.285
NEWPROM2	1	0.6541	0.6590	0.9852	0.3209	0.170835	1.923
LENBIN2	1	1.1779	0.6038	3.8059	0.0511	0.279883	3.247
PRMCOM1	1	-2.4370	1.3817	3.1110	0.0778	-0.520119	0.087
PRMCOM2	1	-0.4668	0.7328	0.4058	0.5241	-0.101687	0.627
PRMCOM3	1	-2.1575	1.2917	2.7899	0.0949	-0.299148	0.116
PRMCOM4	1	-0.3488	0.7854	0.1972	0.6570	-0.058457	0.706
I1	1	0.4454	1.7693	0.0634	0.8012	0.090799	1.561
I2	1	0.7036	1.3643	0.2660	0.6060	0.145189	2.021
I3	1	0.3049	1.1573	0.0694	0.7922	0.065073	1.356
I4	1	0.2456	0.9206	0.0712	0.7896	0.049444	1.278
I5	1	-0.2823	0.5704	0.2449	0.6207	-0.055703	0.754

Association of Predicted Probabilities and Observed Responses

Concordant = 79.3%	Somers' D = 0.606
Discordant = 18.7%	Gamma = 0.618
Tied = 2.0%	Tau-a = 0.404
(52920 pairs)	c = 0.803

Newomor vs dummies (3 cat) civil staff (model 4.5c)

The LOGISTIC Procedure

Data Set: WORK.PROP
Response Variable: NEWOMOR
Response Levels: 3
Number of Observations: 399
Link Function: Logit

Response Profile

Ordered Value	NEWOMOR	Count
1	1	122
2	2	146
3	3	131

Score Test for the Proportional Odds Assumption

Chi-Square = 3.7057 with 6 DF (p=0.7164)

Model Fitting Information and Testing Global Null Hypothesis BETA=0

Criterion	Intercept Only	Intercept and Covariates	Chi-Square for Covariates
AIC	878.495	721.775	.
SC	886.473	753.687	.
-2 LOG L	874.495	705.775	168.720 with 6 DF (p=0.0001)
Score	.	.	142.905 with 6 DF (p=0.0001)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCP1	1	-3.7784	0.3392	124.0711	0.0001	.	.
INTERCP2	1	-1.6064	0.2836	32.0860	0.0001	.	.
JOBNEW	1	1.9802	0.2338	71.7027	0.0001	0.517196	7.244
NEWPUB1	1	1.4408	0.2868	25.2315	0.0001	0.354416	4.224
NEWPUB2	1	0.5856	0.2500	5.4841	0.0192	0.161306	1.796
NEWPROM1	1	0.6924	0.2580	7.1995	0.0073	0.171736	1.998
NEWPROM2	1	0.3329	0.2426	1.8829	0.1700	0.086958	1.395
LENDUM	1	1.2568	0.2483	25.6227	0.0001	0.298639	3.514

Association of Predicted Probabilities and Observed Responses

Concordant	= 77.7%	Somers' D	= 0.590
Discordant	= 18.7%	Gamma	= 0.612
Tied	= 3.6%	Tau-a	= 0.393
(52920 pairs)		c	= 0.795

Newomor vs dummies (3 cat) civil staff + g.o.f. (model 4.6c)

The LOGISTIC Procedure

Data Set: WORK.MOLENBRG
Response Variable: NEWOMOR
Response Levels: 3
Number of Observations: 399
Link Function: Logit

Response Profile

Ordered Value	NEWOMOR	Count
1	1	122
2	2	146
3	3	131

Score Test for the Proportional Odds Assumption

Chi-Square = 13.1648 with 11 DF (p=0.2827)

Model Fitting Information and Testing Global Null Hypothesis BETA=0

Criterion	Intercept Only	Intercept and Covariates	Chi-Square for Covariates
AIC	878.495	721.800	.
SC	886.473	773.657	.
-2 LOG L	874.495	695.800	178.695 with 11 DF (p=0.0001)
Score	.	.	151.702 with 11 DF (p=0.0001)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCP1	1	-3.3370	0.4048	67.9497	0.0001	.	.
INTERCP2	1	-1.0807	0.3673	8.6579	0.0033	.	.
JOBNEW	1	0.6727	0.7272	0.8559	0.3549	0.175713	1.960
NEWPUB1	1	0.3880	0.5922	0.4292	0.5124	0.095443	1.474
NEWPUB2	1	0.1618	0.3577	0.2047	0.6510	0.044583	1.176
NEWPROM1	1	0.2451	0.3787	0.4190	0.5174	0.060800	1.278
NEWPROM2	1	0.0214	0.3358	0.0041	0.9492	0.005585	1.022
LENBIN2	1	0.6042	0.5610	1.1600	0.2815	0.143569	1.830
I1	1	2.3560	1.6644	2.0036	0.1569	0.455094	10.548
I2	1	2.4572	1.2836	3.6647	0.0556	0.524440	11.672
I3	1	1.6420	1.0566	2.4148	0.1202	0.321743	5.165
I4	1	1.0815	0.8366	1.6713	0.1961	0.230823	2.949
I5	1	0.1660	0.5586	0.0884	0.7663	0.035435	1.181

Association of Predicted Probabilities and Observed Responses

Concordant	= 77.6%	Somers' D	= 0.592
Discordant	= 18.4%	Gamma	= 0.617
Tied	= 4.0%	Tau-a	= 0.394
(52920 pairs)		c	= 0.796

Appendix 4b

newomor vs 3 CHAID scored cats officers

The LOGISTIC Procedure

Data Set: WORK.PROPOFF
Response Variable: NEWOMOR
Response Levels: 3
Number of Observations: 1438
Link Function: Logit

Response Profile

Ordered Value	NEWOMOR	Count
1	1	640
2	2	411
3	3	387

Score Test for the Proportional Odds Assumption

Chi-Square = 20.9811 with 5 DF (p=0.0008)

Model Fitting Information and Testing Global Null Hypothesis BETA=0

Criterion	Intercept Only	Intercept and Covariates	Chi-Square for Covariates
AIC	3085.637	2564.347	.
SC	3096.179	2601.244	.
-2 LOG L	3081.637	2550.347	531.290 with 5 DF (p=0.0001)
Score	.	.	451.934 with 5 DF (p=0.0001)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCP1	1	-1.0661	0.1626	43.0073	0.0001	.	.
INTERCP2	1	0.6012	0.1600	14.1195	0.0002	.	.
JOBNEW	1	2.0020	0.1399	204.8464	0.0001	0.465286	7.404
COM3CHD	1	-0.0124	0.00156	64.0201	0.0001	-0.249436	0.988
PUB3CHD	1	-0.00697	0.00168	17.2512	0.0001	-0.126170	0.993
PROM3CHD	1	-0.00876	0.00137	40.7839	0.0001	-0.193781	0.991
LENBIN2	1	1.6045	0.2971	29.1653	0.0001	0.207541	4.976

Association of Predicted Probabilities and Observed Responses

Concordant	= 74.0%	Somers' D	= 0.534
Discordant	= 20.6%	Gamma	= 0.564
Tied	= 5.4%	Tau-a	= 0.346
(669777 pairs)		c	= 0.767

newomor vs integer 3 cats officers (model 4.7o)

The LOGISTIC Procedure

Data Set: WORK.PROPOFF
Response Variable: NEWOMOR
Response Levels: 3
Number of Observations: 1438
Link Function: Logit

Response Profile

Ordered Value	NEWOMOR	Count
1	1	640
2	2	411
3	3	387

Score Test for the Proportional Odds Assumption

Chi-Square = 7.8406 with 5 DF (p=0.1652)

Model Fitting Information and Testing Global Null Hypothesis BETA=0

Criterion	Intercept Only	Intercept and Covariates	Chi-Square for Covariates
AIC	3085.637	2533.209	.
SC	3096.179	2570.106	.
-2 LOG L	3081.637	2519.209	562.428 with 5 DF (p=0.0001)
Score	.	.	472.080 with 5 DF (p=0.0001)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCP1	1	0.6990	0.2571	7.3932	0.0065	.	.
INTERCP2	1	2.3955	0.2644	82.0900	0.0001	.	.
JOBNEW	1	2.0000	0.1406	202.4739	0.0001	0.464816	7.389
NEWCOM	1	-0.5681	0.0759	55.9859	0.0001	-0.234142	0.567
NEWPUB	1	-0.5539	0.0808	47.0314	0.0001	-0.211263	0.575
NEWPROM	1	-0.4099	0.0709	33.4709	0.0001	-0.179138	0.664
LENBIN2	1	1.6536	0.2991	30.5623	0.0001	0.213881	5.226

Association of Predicted Probabilities and Observed Responses

Concordant	= 76.3%	Somers' D	= 0.557
Discordant	= 20.7%	Gamma	= 0.574
Tied	= 3.0%	Tau-a	= 0.361
(669777 pairs)		c	= 0.778

newomor vs integer 3 cats + g-o-f officers (model 4.8o)

The LOGISTIC Procedure

Data Set: WORK.MOLENOFF
Response Variable: NEWOMOR
Response Levels: 3
Number of Observations: 1438
Link Function: Logit

Response Profile

Ordered Value	NEWOMOR	Count
1	1	640
2	2	411
3	3	387

Score Test for the Proportional Odds Assumption

Chi-Square = 31.4177 with 14 DF (p=0.0048)

Model Fitting Information and Testing Global Null Hypothesis BETA=0

Criterion	Intercept Only	Intercept and Covariates	Chi-Square for Covariates
AIC	3085.637	2547.141	.
SC	3096.179	2631.477	.
-2 LOG L	3081.637	2515.141	566.496 with 14 DF (p=0.0001)
Score	.	.	481.850 with 14 DF (p=0.0001)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCP1	1	0.2988	1.0660	0.0786	0.7792	.	.
INTERCP2	1	2.0009	1.0665	3.5196	0.0606	.	.
JOBNEW	1	1.8731	0.5905	10.0613	0.0015	0.435331	6.509
NEWCOM	1	-0.4423	0.1756	6.3443	0.0118	-0.182308	0.643
NEWPUB	1	-0.4814	0.1563	9.4877	0.0021	-0.183593	0.618
NEWPROM	1	-0.3962	0.1476	7.2044	0.0073	-0.173118	0.673
LENBIN2	1	1.4611	0.4637	9.9270	0.0016	0.188981	4.310
I1	1	0.4300	1.2406	0.1202	0.7289	0.078300	1.537
I2	1	0.3201	1.0998	0.0847	0.7710	0.050795	1.377
I3	1	0.2244	1.0753	0.0435	0.8347	0.031771	1.252
I4	1	0.3972	0.9784	0.1648	0.6848	0.074178	1.488
I5	1	-0.00529	0.9420	0.0000	0.9955	-0.000762	0.995
I6	1	0.1108	0.8565	0.0167	0.8971	0.017996	1.117
I7	1	0.0435	0.7512	0.0033	0.9539	0.007759	1.044
I8	1	-0.0730	0.6119	0.0142	0.9050	-0.011632	0.930
I9	1	0.0495	0.3924	0.0159	0.8996	0.008145	1.051

Association of Predicted Probabilities and Observed Responses

Concordant	= 76.3%	Somers' D	= 0.556
Discordant	= 20.7%	Gamma	= 0.574
Tied	= 3.0%	Tau-a	= 0.361
(669777 pairs)		c	= 0.778

newomor vs 3 CHAID scored cats civil staff

The LOGISTIC Procedure

Data Set: WORK.PROPCIV
Response Variable: NEWOMOR
Response Levels: 3
Number of Observations: 399
Link Function: Logit

Response Profile

Ordered Value	NEWOMOR	Count
1	1	122
2	2	146
3	3	131

Score Test for the Proportional Odds Assumption

Chi-Square = 7.1975 with 5 DF (p=0.2064)

Model Fitting Information and Testing Global Null Hypothesis BETA=0

Criterion	Intercept Only	Intercept and Covariates	Chi-Square for Covariates
AIC	878.495	731.050	.
SC	886.473	758.973	.
-2 LOG L	874.495	717.050	157.445 with 5 DF (p=0.0001)
Score	.	.	135.568 with 5 DF (p=0.0001)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCP1	1	-1.9848	0.2893	47.0634	0.0001	.	.
INTERCP2	1	0.1338	0.2698	0.2460	0.6199	.	.
JOBNEW	1	1.9219	0.2379	65.2566	0.0001	0.501969	6.834
COM3CHD	1	-0.00360	0.00303	1.4096	0.2351	-0.069986	0.996
PUB3CHD	1	-0.00840	0.00237	12.5554	0.0004	-0.202283	0.992
PROM3CHD	1	-0.00767	0.00258	8.8211	0.0030	-0.168552	0.992
LENBIN2	1	1.2060	0.2432	24.5885	0.0001	0.286557	3.340

Association of Predicted Probabilities and Observed Responses

Concordant	= 76.7%	Somers' D	= 0.553
Discordant	= 21.4%	Gamma	= 0.564
Tied	= 1.9%	Tau-a	= 0.369
(52920 pairs)		c	= 0.777

The LOGISTIC Procedure

Data Set: WORK.PROPCIV
Response Variable: NEWOMOR
Response Levels: 3
Number of Observations: 399
Link Function: Logit

Response Profile

Ordered Value	NEWOMOR	Count
1	1	122
2	2	146
3	3	131

Score Test for the Proportional Odds Assumption

Chi-Square = 3.1045 with 4 DF (p=0.5405)

Model Fitting Information and Testing Global Null Hypothesis BETA=0

Criterion	Intercept Only	Intercept and Covariates	Chi-Square for Covariates
AIC	878.495	730.380	.
SC	886.473	754.313	.
-2 LOG L	874.495	718.380	156.116 with 4 DF (p=0.0001)
Score	.	.	134.670 with 4 DF (p=0.0001)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCP1	1	-2.0802	0.2784	55.8325	0.0001	.	.
INTERCP2	1	0.0328	0.2558	0.0164	0.8980	.	.
JOBNEW	1	1.9868	0.2321	73.2802	0.0001	0.518931	7.292
PUB3CHD	1	-0.00871	0.00235	13.7297	0.0002	-0.209620	0.991
PROM3CHD	1	-0.00802	0.00257	9.7546	0.0018	-0.176289	0.992
LENBIN2	1	1.2030	0.2428	24.5466	0.0001	0.285858	3.330

Association of Predicted Probabilities and Observed Responses

Concordant	= 75.0%	Somers' D	= 0.549
Discordant	= 20.1%	Gamma	= 0.577
Tied	= 4.8%	Tau-a	= 0.366
(52920 pairs)		c	= 0.775

newomor vs integer 3 cats civil staff include com

The LOGISTIC Procedure

Data Set: WORK.PROPCIV
Response Variable: NEWOMOR
Response Levels: 3
Number of Observations: 399
Link Function: Logit

Response Profile

Ordered Value	NEWOMOR	Count
1	1	122
2	2	146
3	3	131

Score Test for the Proportional Odds Assumption

Chi-Square = 4.4365 with 5 DF (p=0.4884)

Model Fitting Information and Testing Global Null Hypothesis BETA=0

Criterion	Intercept Only	Intercept and Covariates	Chi-Square for Covariates
AIC	878.495	719.668	.
SC	886.473	747.590	.
-2 LOG L	874.495	705.668	168.828 with 5 DF (p=0.0001)
Score	.	.	143.078 with 5 DF (p=0.0001)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCP1	1	-0.4800	0.4666	1.0580	0.3037	.	.
INTERCP2	1	1.6940	0.4762	12.6533	0.0004	.	.
JOBNEW	1	1.9410	0.2408	64.9656	0.0001	0.506971	6.966
NEWCOM	1	-0.1080	0.1425	0.5744	0.4485	-0.044827	0.898
NEWPUB	1	-0.7108	0.1443	24.2578	0.0001	-0.285967	0.491
NEWPROM	1	-0.3397	0.1289	6.9404	0.0084	-0.151356	0.712
LENBIN2	1	1.2393	0.2455	25.4775	0.0001	0.294474	3.453

Association of Predicted Probabilities and Observed Responses

Concordant	= 78.5%	Somers' D	= 0.587
Discordant	= 19.8%	Gamma	= 0.597
Tied	= 1.7%	Tau-a	= 0.391
(52920 pairs)		c	= 0.794

newomor vs 3 -integer scored cats exl com civil staff (mod 4.7c)

The LOGISTIC Procedure

Data Set: WORK.PROPCIV
Response Variable: NEWOMOR
Response Levels: 3
Number of Observations: 399
Link Function: Logit

Response Profile

Ordered Value	NEWOMOR	Count
1	1	122
2	2	146
3	3	131

Score Test for the Proportional Odds Assumption

Chi-Square = 2.7929 with 4 DF (p=0.5931)

Model Fitting Information and Testing Global Null Hypothesis BETA=0

Criterion	Intercept Only	Intercept and Covariates	Chi-Square for Covariates
AIC	878.495	718.221	.
SC	886.473	742.155	.
-2 LOG L	874.495	706.221	168.274 with 4 DF (p=0.0001)
Score	.	.	142.690 with 4 DF (p=0.0001)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCP1	1	-0.6347	0.4201	2.2828	0.1308	.	.
INTERCP2	1	1.5362	0.4280	12.8848	0.0003	.	.
JOBNEW	1	1.9851	0.2335	72.2898	0.0001	0.518486	7.280
NEWPUB	1	-0.7228	0.1433	25.4545	0.0001	-0.290764	0.485
NEWPROM	1	-0.3503	0.1283	7.4524	0.0063	-0.156075	0.705
LENBIN2	1	1.2437	0.2455	25.6729	0.0001	0.295515	3.468

Association of Predicted Probabilities and Observed Responses

Concordant	= 77.4%	Somers' D	= 0.585
Discordant	= 18.8%	Gamma	= 0.608
Tied	= 3.8%	Tau-a	= 0.390
(52920 pairs)		c	= 0.793

newomor vs 3 integer scored cats excl com + g-o-f civil staff (model 4.8c)

The LOGISTIC Procedure

Data Set: WORK.MOLENCIV
Response Variable: NEWOMOR
Response Levels: 3
Number of Observations: 399
Link Function: Logit

Response Profile

Ordered Value	NEWOMOR	Count
1	1	122
2	2	146
3	3	131

Score Test for the Proportional Odds Assumption

Chi-Square = 12.2645 with 9 DF (p=0.1988)

Model Fitting Information and Testing Global Null Hypothesis BETA=0

Criterion	Intercept Only	Intercept and Covariates	Chi-Square for Covariates
AIC	878.495	719.165	.
SC	886.473	763.044	.
-2 LOG L	874.495	697.165	177.330 with 9 DF (p=0.0001)
Score	.	.	150.121 with 9 DF (p=0.0001)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCP1	1	-1.8708	1.0106	3.4269	0.0641	.	.
INTERCP2	1	0.3703	1.0055	0.1356	0.7127	.	.
JOBNEW	1	0.9329	0.7094	1.7291	0.1885	0.243651	2.542
NEWPUB	1	-0.2608	0.3034	0.7389	0.3900	-0.104905	0.770
NEWPROM	1	-0.2610	0.1544	2.8579	0.0909	-0.116311	0.770
LENBIN2	1	0.8234	0.5246	2.4639	0.1165	0.195651	2.278
I1	1	1.6140	1.5110	1.1410	0.2854	0.311762	5.023
I2	1	1.9702	1.2094	2.6540	0.1033	0.420496	7.172
I3	1	1.0584	0.9358	1.2794	0.2580	0.238203	2.882
I4	1	0.7493	0.7445	1.0129	0.3142	0.134552	2.115
I5	1	-0.0238	0.5183	0.0021	0.9634	-0.005079	0.976

Association of Predicted Probabilities and Observed Responses

Concordant	= 77.4%	Somers' D	= 0.593
Discordant	= 18.1%	Gamma	= 0.621
Tied	= 4.5%	Tau-a	= 0.395
(52920 pairs)		c	= 0.796

Appendix 4c

Cont odds model for dummy vars (full sample)

The LOGISTIC Procedure

Data Set: WORK.CONT2
Response Variable: IND
Response Levels: 2
Number of Observations: 2912
Link Function: Logit

Response Profile

Ordered Value	IND	Count
1	0	1319
2	1	1593

Model Fitting Information and Testing Global Null Hypothesis BETA=0

Criterion	Intercept Only	Intercept and Covariates	Chi-Square for Covariates
AIC	4013.069	3279.491	.
SC	4019.046	3345.233	.
-2 LOG L	4011.069	3257.491	753.579 with 10 DF (p=0.0001)
Score	.	.	661.714 with 10 DF (p=0.0001)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCPT	1	-4.4173	0.2419	333.4749	0.0001	.	.
CUTPT	1	1.0724	0.0948	127.8799	0.0001	0.285378	2.922
JOBNEW	1	1.6987	0.1081	246.7352	0.0001	0.431922	5.467
NEWCOM1	1	0.8453	0.1293	42.7627	0.0001	0.233070	2.329
NEWCOM2	1	0.4441	0.1359	10.6710	0.0011	0.113899	1.559
NEWPROM1	1	0.6959	0.1128	38.0876	0.0001	0.161838	2.005
NEWPROM2	1	0.4068	0.0984	17.0749	0.0001	0.106684	1.502
NEWPUB1	1	1.0153	0.1343	57.1688	0.0001	0.273972	2.760
NEWPUB2	1	0.4409	0.1299	11.5161	0.0007	0.120508	1.554
LENBIN2	1	1.2351	0.1691	53.3445	0.0001	0.188354	3.439
OFFNEW	1	-0.3585	0.1110	10.4243	0.0012	-0.083472	0.699

Association of Predicted Probabilities and Observed Responses

Concordant	= 77.4%	Somers' D	= 0.559
Discordant	= 21.5%	Gamma	= 0.565
Tied	= 1.1%	Tau-a	= 0.277
(2101167 pairs)		c	= 0.780

bin log 1 for cont odds dummy (full sample)

The LOGISTIC Procedure

Data Set: WORK.PROP
Response Variable: BINOMOR1
Response Levels: 2
Number of Observations: 1837
Link Function: Logit

Response Profile

Ordered Value	BINOMOR1	Count
1	1	762
2	2	1075

Model Fitting Information and Testing Global Null Hypothesis BETA=0

Criterion	Intercept Only	Intercept and Covariates	Chi-Square for Covariates
AIC	2495.031	1979.777	.
SC	2500.547	2034.936	.
-2 LOG L	2493.031	1959.777	533.254 with 9 DF (p=0.0001)
Score	.	.	448.553 with 9 DF (p=0.0001)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCPT	1	-3.3046	0.2616	159.5578	0.0001	.	.
JOBNEW	1	2.1561	0.1806	142.4586	0.0001	0.517730	8.637
NEWCOM1	1	0.8130	0.1793	20.5704	0.0001	0.223645	2.255
NEWCOM2	1	0.2230	0.1928	1.3381	0.2474	0.056467	1.250
NEWPROM1	1	0.6486	0.1419	20.8973	0.0001	0.156395	1.913
NEWPROM2	1	0.2014	0.1297	2.4097	0.1206	0.052903	1.223
NEWPUB1	1	0.8486	0.1770	22.9759	0.0001	0.232025	2.336
NEWPUB2	1	0.0992	0.1777	0.3116	0.5767	0.026924	1.104
LENBIN2	1	1.5067	0.2070	52.9881	0.0001	0.248244	4.512
OFFNEW	1	-0.7845	0.1546	25.7593	0.0001	-0.178393	0.456

Association of Predicted Probabilities and Observed Responses

Concordant	= 78.5%	Somers' D	= 0.590
Discordant	= 19.4%	Gamma	= 0.603
Tied	= 2.1%	Tau-a	= 0.287
(819150 pairs)		c	= 0.795

bin log 2 for cont odds dummy (full sample)

The LOGISTIC Procedure

Data Set: WORK.PROP
Response Variable: BINOMOR2
Response Levels: 2
Number of Observations: 1075
Link Function: Logit

Response Profile

Ordered Value	BINOMOR2	Count
1	1	557
2	2	518

WARNING: 762 observation(s) were deleted due to missing values for the response or explanatory variables.

Model Fitting Information and Testing Global Null Hypothesis BETA=0

Criterion	Intercept Only	Intercept and Covariates	Chi-Square for Covariates
AIC	1490.851	1263.688	.
SC	1495.831	1313.489	.
-2 LOG L	1488.851	1243.688	245.163 with 9 DF (p=0.0001)
Score	.	.	224.430 with 9 DF (p=0.0001)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCPT	1	-2.6278	0.2477	112.5806	0.0001	.	.
JOBNEW	1	1.3904	0.1444	92.6846	0.0001	0.375131	4.017
NEWCOM1	1	0.7971	0.1892	17.7394	0.0001	0.218036	2.219
NEWCOM2	1	0.6753	0.1939	12.1318	0.0005	0.176621	1.965
NEWPROM1	1	0.6958	0.1929	13.0164	0.0003	0.149643	2.005
NEWPROM2	1	0.7037	0.1541	20.8394	0.0001	0.184168	2.021
NEWPUB1	1	1.1641	0.2097	30.8098	0.0001	0.302399	3.203
NEWPUB2	1	0.8850	0.1928	21.0768	0.0001	0.243727	2.423
LENBIN2	1	0.8991	0.3332	7.2803	0.0070	0.114737	2.457
OFFNEW	1	0.1739	0.1684	1.0658	0.3019	0.041950	1.190

Association of Predicted Probabilities and Observed Responses

Concordant	= 75.6%	Somers' D	= 0.526
Discordant	= 23.0%	Gamma	= 0.533
Tied	= 1.4%	Tau-a	= 0.263
(288526 pairs)		c	= 0.763

cont odds using dummy vars officers

The LOGISTIC Procedure

Data Set: WORK.CONTOFF2
Response Variable: IND
Response Levels: 2
Number of Observations: 2236
Link Function: Logit

Response Profile

Ordered Value	IND	Count
1	0	1051
2	1	1185

Model Fitting Information and Testing Global Null Hypothesis BETA=0

Criterion	Intercept Only	Intercept and Covariates	Chi-Square for Covariates
AIC	3093.719	2548.987	.
SC	3099.431	2606.111	.
-2 LOG L	3091.719	2528.987	562.732 with 9 DF (p=0.0001)
Score	.	.	494.893 with 9 DF (p=0.0001)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCPT	1	-4.3009	0.2787	238.1872	0.0001	.	.
CUTPT	1	0.9181	0.1074	73.0150	0.0001	0.242555	2.505
JOBNEW	1	1.7227	0.1256	188.0262	0.0001	0.428383	5.600
NEWCOM1	1	0.9681	0.1452	44.4620	0.0001	0.266829	2.633
NEWCOM2	1	0.4297	0.1518	8.0130	0.0046	0.111005	1.537
NEWPUB1	1	0.9756	0.1616	36.4432	0.0001	0.267283	2.653
NEWPUB2	1	0.4450	0.1613	7.6125	0.0058	0.121145	1.560
NEWPROM1	1	0.7274	0.1292	31.6865	0.0001	0.166651	2.070
NEWPROM2	1	0.4414	0.1104	15.9820	0.0001	0.116092	1.555
LENBIN2	1	1.5235	0.2685	32.1905	0.0001	0.174476	4.588

Association of Predicted Probabilities and Observed Responses

Concordant	= 76.6%	Somers' D	= 0.550
Discordant	= 21.6%	Gamma	= 0.561
Tied	= 1.9%	Tau-a	= 0.274
(1245435 pairs)		c	= 0.775

The LOGISTIC Procedure

Data Set: WORK.PROPOFF
Response Variable: BINOMOR1
Response Levels: 2
Number of Observations: 1438
Link Function: Logit

Response Profile

Ordered Value	BINOMOR1	Count
1	1	640
2	2	798

Model Fitting Information and Testing Global Null Hypothesis BETA=0

Criterion	Intercept Only	Intercept and Covariates	Chi-Square for Covariates
AIC	1978.096	1571.016	.
SC	1983.367	1618.455	.
-2 LOG L	1976.096	1553.016	423.080 with 8 DF (p=0.0001)
Score	.	.	355.072 with 8 DF (p=0.0001)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCPT	1	-3.5226	0.3095	129.5227	0.0001	.	.
JOBNEW	1	2.3055	0.2137	116.3548	0.0001	0.535807	10.029
NEWCOM1	1	1.0545	0.2028	27.0296	0.0001	0.290351	2.871
NEWCOM2	1	0.3486	0.2163	2.5978	0.1070	0.088812	1.417
NEWPUB1	1	0.6953	0.2086	11.1112	0.0009	0.191596	2.004
NEWPUB2	1	-0.0150	0.2127	0.0050	0.9437	-0.004047	0.985
NEWPROM1	1	0.7246	0.1608	20.3127	0.0001	0.173279	2.064
NEWPROM2	1	0.1900	0.1434	1.7549	0.1853	0.050004	1.209
LENBIN2	1	1.8041	0.3268	30.4824	0.0001	0.233355	6.075

Association of Predicted Probabilities and Observed Responses

Concordant	= 77.7%	Somers' D	= 0.587
Discordant	= 19.1%	Gamma	= 0.606
Tied	= 3.2%	Tau-a	= 0.290
(510720 pairs)		c	= 0.793

The LOGISTIC Procedure

Data Set: WORK.PROPOFF
Response Variable: BINOMOR2
Response Levels: 2
Number of Observations: 798
Link Function: Logit

Response Profile

Ordered Value	BINOMOR2	Count
1	1	411
2	2	387

WARNING: 640 observation(s) were deleted due to missing values for the response or explanatory variables.

Model Fitting Information and Testing Global Null Hypothesis BETA=0

Criterion	Intercept Only	Intercept and Covariates	Chi-Square for Covariates
AIC	1107.541	948.459	.
SC	1112.223	990.598	.
-2 LOG L	1105.541	930.459	175.082 with 8 DF (p=0.0001)
Score	.	.	160.217 with 8 DF (p=0.0001)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCPT	1	-2.6895	0.2958	82.6668	0.0001	.	.
JOBNEW	1	1.2848	0.1673	58.9473	0.0001	0.343977	3.614
NEWCOM1	1	0.8275	0.2153	14.7745	0.0001	0.224500	2.288
NEWCOM2	1	0.5649	0.2180	6.7144	0.0096	0.149084	1.759
NEWPUB1	1	1.2891	0.2581	24.9408	0.0001	0.344699	3.630
NEWPUB2	1	1.0486	0.2500	17.5932	0.0001	0.288624	2.854
NEWPROM1	1	0.6126	0.2251	7.4030	0.0065	0.126695	1.845
NEWPROM2	1	0.8414	0.1774	22.4891	0.0001	0.221149	2.320
LENBIN2	1	0.7312	0.5584	1.7147	0.1904	0.058244	2.078

Association of Predicted Probabilities and Observed Responses

Concordant	= 74.6%	Somers' D	= 0.516
Discordant	= 23.0%	Gamma	= 0.529
Tied	= 2.4%	Tau-a	= 0.258
(159057 pairs)		c	= 0.758

The LOGISTIC Procedure

Data Set: WORK.CONTCIV2
Response Variable: IND
Response Levels: 2
Number of Observations: 676
Link Function: Logit

Response Profile

Ordered Value	IND	Count
1	0	268
2	1	408

Model Fitting Information and Testing Global Null Hypothesis BETA=0

Criterion	Intercept Only	Intercept and Covariates	Chi-Square for Covariates
AIC	909.930	721.534	.
SC	914.446	766.696	.
-2 LOG L	907.930	701.534	206.396 with 9 DF (p=0.0001)
Score	.	.	179.263 with 9 DF (p=0.0001)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCPT	1	-5.4427	0.5122	112.9159	0.0001	.	.
CUTPT	1	1.6736	0.2072	65.2253	0.0001	0.454120	5.331
JOBNEW	1	1.7596	0.2234	62.0413	0.0001	0.471613	5.810
NEWCOM1	1	0.3580	0.2895	1.5292	0.2162	0.098453	1.431
NEWCOM2	1	0.4874	0.3075	2.5126	0.1129	0.121835	1.628
NEWPUB1	1	1.2904	0.2671	23.3358	0.0001	0.305809	3.634
NEWPUB2	1	0.5301	0.2307	5.2796	0.0216	0.146099	1.699
NEWPROM1	1	0.5889	0.2379	6.1279	0.0133	0.143114	1.802
NEWPROM2	1	0.2814	0.2249	1.5663	0.2107	0.073171	1.325
LENBIN2	1	1.2011	0.2316	26.8873	0.0001	0.269957	3.324

Association of Predicted Probabilities and Observed Responses

Concordant	= 80.3%	Somers' D	= 0.615
Discordant	= 18.9%	Gamma	= 0.620
Tied	= 0.8%	Tau-a	= 0.295
(109344 pairs)		c	= 0.807

bin log 1 for cont odds dummy civil staff

The LOGISTIC Procedure

Data Set: WORK.PROPCIV
Response Variable: BINOMOR1
Response Levels: 2
Number of Observations: 399
Link Function: Logit

Response Profile

Ordered Value	BINOMOR1	Count
1	1	122
2	2	277

Model Fitting Information and Testing Global Null Hypothesis BETA=0

Criterion	Intercept Only	Intercept and Covariates	Chi-Square for Covariates
AIC	493.304	408.128	.
SC	497.293	444.029	.
-2 LOG L	491.304	390.128	101.176 with 8 DF (p=0.0001)
Score	.	.	90.882 with 8 DF (p=0.0001)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCPT	1	-3.3382	0.4871	46.9583	0.0001	.	.
JOBNEW	1	1.8701	0.3531	28.0499	0.0001	0.488437	6.489
NEWCOM1	1	-0.1635	0.3958	0.1706	0.6795	-0.044813	0.849
NEWCOM2	1	-0.2525	0.4371	0.3337	0.5635	-0.062459	0.777
NEWPUB1	1	1.4117	0.3582	15.5355	0.0001	0.347252	4.103
NEWPUB2	1	0.4897	0.3383	2.0953	0.1478	0.134906	1.632
NEWPROM1	1	0.3945	0.3167	1.5514	0.2129	0.097857	1.484
NEWPROM2	1	0.3019	0.3142	0.9231	0.3367	0.078840	1.352
LENBIN2	1	1.2817	0.2785	21.1788	0.0001	0.304557	3.603

Association of Predicted Probabilities and Observed Responses

Concordant	= 78.9%	Somers' D	= 0.597
Discordant	= 19.2%	Gamma	= 0.609
Tied	= 1.9%	Tau-a	= 0.254
(33794 pairs)		c	= 0.799

bin log 2 for cont odds dummy civil staff

The LOGISTIC Procedure

Data Set: WORK.PROPCIV
Response Variable: BINOMOR2
Response Levels: 2
Number of Observations: 277
Link Function: Logit

Response Profile

Ordered Value	BINOMOR2	Count
1	1	146
2	2	131

WARNING: 122 observation(s) were deleted due to missing values for the response or explanatory variables.

Model Fitting Information and Testing Global Null Hypothesis BETA=0

Criterion	Intercept Only	Intercept and Covariates	Chi-Square for Covariates
AIC	385.191	321.269	.
SC	388.815	353.885	.
-2 LOG L	383.191	303.269	79.922 with 8 DF (p=0.0001)
Score	.	.	71.382 with 8 DF (p=0.0001)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCPT	1	-2.5351	0.4561	30.8881	0.0001	.	.
JOBNEW	1	1.7066	0.2980	32.7895	0.0001	0.468335	5.510
NEWCOM1	1	0.7991	0.4165	3.6808	0.0550	0.220604	2.223
NEWCOM2	1	1.1560	0.4363	7.0209	0.0081	0.293494	3.177
NEWPUB1	1	1.0609	0.4166	6.4840	0.0109	0.235331	2.889
NEWPUB2	1	0.6001	0.3231	3.4494	0.0633	0.165709	1.822
NEWPROM1	1	0.9039	0.3821	5.5953	0.0180	0.212680	2.469
NEWPROM2	1	0.2306	0.3226	0.5108	0.4748	0.059642	1.259
LENBIN2	1	1.1117	0.4251	6.8390	0.0089	0.224445	3.040

Association of Predicted Probabilities and Observed Responses

Concordant	= 79.1%	Somers' D	= 0.600
Discordant	= 19.2%	Gamma	= 0.610
Tied	= 1.7%	Tau-a	= 0.300
(19126 pairs)		c	= 0.800

Appendix 4d

cont odds using integers officers

The LOGISTIC Procedure

Data Set: WORK.CONTOFF2
Response Variable: IND
Response Levels: 2
Number of Observations: 2236
Link Function: Logit

Response Profile

Ordered Value	IND	Count
1	0	1051
2	1	1185

Model Fitting Information and Testing Global Null Hypothesis BETA=0

Criterion	Intercept Only	Intercept and Covariates	Chi-Square for Covariates
AIC	3093.719	2543.903	.
SC	3099.431	2583.890	.
-2 LOG L	3091.719	2529.903	561.816 with 6 DF (p=0.0001)
Score	.	.	493.081 with 6 DF (p=0.0001)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCPT	1	-0.2412	0.2627	0.8431	0.3585	.	.
CUTPT	1	0.9165	0.1073	72.9471	0.0001	0.242143	2.501
JOBNEW	1	1.7234	0.1254	188.9950	0.0001	0.428555	5.604
NEWCOM	1	-0.4971	0.0679	53.5795	0.0001	-0.210014	0.608
NEWPUB	1	-0.4990	0.0731	46.6411	0.0001	-0.191487	0.607
NEWPROM	1	-0.3726	0.0635	34.4001	0.0001	-0.160150	0.689
LENBIN2	1	1.5254	0.2690	32.1634	0.0001	0.174691	4.597

Association of Predicted Probabilities and Observed Responses

Concordant	= 76.4%	Somers' D	= 0.549
Discordant	= 21.5%	Gamma	= 0.561
Tied	= 2.1%	Tau-a	= 0.274
(1245435 pairs)		c	= 0.774

bin log 1 for- cont odds integers officers

The LOGISTIC Procedure

Data Set: WORK.PROPOFF
Response Variable: BINOMOR1
Response Levels: 2
Number of Observations: 1438
Link Function: Logit

Response Profile

Ordered Value	BINOMOR1	Count
1	1	640
2	2	798

Model Fitting Information and Testing Global Null Hypothesis BETA=0

Criterion	Intercept Only	Intercept and Covariates	Chi-Square for Covariates
AIC	1978.096	1575.220	.
SC	1983.367	1606.847	.
-2 LOG L	1976.096	1563.220	412.875 with 5 DF (p=0.0001)
Score	.	.	343.564 with 5 DF (p=0.0001)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCPT	1	0.2230	0.3206	0.4840	0.4866	.	.
JOBNEW	1	2.2696	0.2125	114.0808	0.0001	0.527472	9.676
NEWCOM	1	-0.5867	0.0909	41.6955	0.0001	-0.241789	0.556
NEWPUB	1	-0.4694	0.0936	25.1744	0.0001	-0.179044	0.625
NEWPROM	1	-0.3516	0.0793	19.6749	0.0001	-0.153665	0.704
LENBIN2	1	1.8098	0.3277	30.5009	0.0001	0.234092	6.109

Association of Predicted Probabilities and Observed Responses

Concordant	= 77.3%	Somers' D	= 0.580
Discordant	= 19.3%	Gamma	= 0.600
Tied	= 3.4%	Tau-a	= 0.287
(510720 pairs)		c	= 0.790

bin log 2 for cont odds officers

The LOGISTIC Procedure

Data Set: WORK.PROPOFF
Response Variable: BINOMOR2
Response Levels: 2
Number of Observations: 798
Link Function: Logit

Response Profile

Ordered Value	BINOMOR2	Count
1	1	411
2	2	387

WARNING: 640 observation(s) were deleted due to missing values for the response or explanatory variables.

Model Fitting Information and Testing Global Null Hypothesis BETA=0

Criterion	Intercept Only	Intercept and Covariates	Chi-Square for Covariates
AIC	1107.541	959.353	.
SC	1112.223	987.445	.
-2 LOG L	1105.541	947.353	158.188 with 5 DF (p=0.0001)
Score	.	.	147.020 with 5 DF (p=0.0001)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCPT	1	1.7866	0.3768	22.4850	0.0001	.	.
JOBNEW	1	1.3119	0.1658	62.5715	0.0001	0.351229	3.713
NEWCOM	1	-0.3671	0.1032	12.6468	0.0004	-0.158827	0.693
NEWPUB	1	-0.5382	0.1171	21.1140	0.0001	-0.206446	0.584
NEWPROM	1	-0.4044	0.1081	13.9977	0.0002	-0.165978	0.667
LENBIN2	1	0.6588	0.5488	1.4410	0.2300	0.052482	1.933

Association of Predicted Probabilities and Observed Responses

Concordant = 73.4%	Somers' D = 0.495
Discordant = 23.9%	Gamma = 0.509
Tied = 2.7%	Tau-a = 0.248
(159057 pairs)	c = 0.748

cont odds using integers civil staff

The LOGISTIC Procedure

Data Set: WORK.CONTCIV2
Response Variable: IND
Response Levels: 2
Number of Observations: 676
Link Function: Logit

Response Profile

Ordered Value	IND	Count
1	0	268
2	1	408

Model Fitting Information and Testing Global Null Hypothesis BETA=0

Criterion	Intercept Only	Intercept and Covariates	Chi-Square for Covariates
AIC	909.930	717.767	.
SC	914.446	749.381	.
-2 LOG L	907.930	703.767	204.163 with 6 DF (p=0.0001)
Score	.	.	177.672 with 6 DF (p=0.0001)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCPT	1	-2.1321	0.5062	17.7431	0.0001	.	.
CUTPT	1	1.6709	0.2071	65.0692	0.0001	0.453392	5.317
JOBNEW	1	1.7541	0.2227	62.0260	0.0001	0.470162	5.778
NEWCOM	1	-0.1047	0.1323	0.6255	0.4290	-0.043823	0.901
NEWPUB	1	-0.6466	0.1335	23.4748	0.0001	-0.257520	0.524
NEWPROM	1	-0.3065	0.1182	6.7259	0.0095	-0.136054	0.736
LENBIN2	1	1.1610	0.2279	25.9599	0.0001	0.260931	3.193

Association of Predicted Probabilities and Observed Responses

Concordant	= 80.1%	Somers' D	= 0.611
Discordant	= 19.0%	Gamma	= 0.616
Tied	= 0.9%	Tau-a	= 0.293
(109344 pairs)		c	= 0.806

bin log 1 for cont odds integers civil staff

The LOGISTIC Procedure

Data Set: WORK.PROPCIV
Response Variable: BINOMOR1
Response Levels: 2
Number of Observations: 399
Link Function: Logit

Response Profile

Ordered Value	BINOMOR1	Count
1	1	122
2	2	277

Model Fitting Information and Testing Global Null Hypothesis BETA=0

Criterion	Intercept Only	Intercept and Covariates	Chi-Square for Covariates
AIC	493.304	403.363	.
SC	497.293	427.297	.
-2 LOG L	491.304	391.363	99.942 with 5 DF (p=0.0001)
Score	.	.	89.558 with 5 DF (p=0.0001)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCPT	1	-0.8817	0.6021	2.1442	0.1431	.	.
JOBNEW	1	1.8676	0.3513	28.2606	0.0001	0.487789	6.473
NEWCOM	1	0.0488	0.1831	0.0709	0.7900	0.020230	1.050
NEWPUB	1	-0.7279	0.1787	16.5890	0.0001	-0.292843	0.483
NEWPROM	1	-0.1997	0.1569	1.6189	0.2032	-0.088974	0.819
LENBIN2	1	1.2780	0.2729	21.9273	0.0001	0.303683	3.590

Association of Predicted Probabilities and Observed Responses

Concordant = 79.1%	Somers' D = 0.598
Discordant = 19.2%	Gamma = 0.609
Tied = 1.7%	Tau-a = 0.255
(33794 pairs)	c = 0.799

bin log 2 for cont odds integers civil staff

The LOGISTIC Procedure

Data Set: WORK.PROPCIV
Response Variable: BINOMOR2
Response Levels: 2
Number of Observations: 277
Link Function: Logit

Response Profile

Ordered Value	BINOMOR2	Count
1	1	146
2	2	131

WARNING: 122 observation(s) were deleted due to missing values for the response or explanatory variables.

Model Fitting Information and Testing Global Null Hypothesis BETA=0

Criterion	Intercept Only	Intercept and Covariates	Chi-Square for Covariates
AIC	385.191	321.239	.
SC	388.815	342.984	.
-2 LOG L	383.191	309.239	73.951 with 5 DF (p=0.0001)
Score	.	.	67.062 with 5 DF (p=0.0001)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCPT	1	1.6429	0.6452	6.4829	0.0109	.	.
JOBNEW	1	1.6406	0.2908	31.8264	0.0001	0.450214	5.158
NEWCOM	1	-0.2578	0.1897	1.8469	0.1741	-0.109255	0.773
NEWPUB	1	-0.5493	0.2033	7.2969	0.0069	-0.214117	0.577
NEWPROM	1	-0.4374	0.1818	5.7872	0.0161	-0.192599	0.646
LENBIN2	1	0.9428	0.4146	5.1705	0.0230	0.190342	2.567

Association of Predicted Probabilities and Observed Responses

Concordant	= 78.1%	Somers' D	= 0.578
Discordant	= 20.3%	Gamma	= 0.588
Tied	= 1.7%	Tau-a	= 0.289
(19126 pairs)		c	= 0.789