11-1-2009

Ordinal Regression Analysis: Fitting the Proportional Odds Model Using Stata, SAS and SPSS

Xing Liu

Eastern Connecticut State University, liux@easternct.edu

STATISTICAL SOFTWARE APPLICATIONS & REVIEW
Ordinal Regression Analysis: Fitting the Proportional Odds Model Using Stata, SAS and SPSS

Xing Liu
Eastern Connecticut State University

Researchers have a variety of options when choosing statistical software packages that can perform ordinal logistic regression analyses. However, statistical software, such as Stata, SAS, and SPSS, may use different techniques to estimate the parameters. The purpose of this article is to (1) illustrate the use of Stata, SAS and SPSS to fit proportional odds models using educational data; and (2) compare the features and results for fitting the proportional odds model using Stata OLOGIT, SAS PROC LOGISTIC (ascending and descending), and SPSS PLUM. The assumption of the proportional odds was tested, and the results of the fitted models were interpreted.

Key words: Proportional Odds Models, Ordinal logistic regression, Stata, SAS, SPSS, Comparison.

Introduction

The proportional odds (PO) model, also called cumulative odds model (Agresti, 1996, 2002; Armstrong & Sloan, 1989; Long, 1997, Long & Freese, 2006; McCullagh, 1980; McCullagh & Nelder, 1989; Powers & Xie, 2000; O'Connell, 2006), is a commonly used model for the analysis of ordinal categorical data and comes from the class of generalized linear models. It is a generalization of a binary logistic regression model when the response variable has more than two ordinal categories. The proportional odds model is used to estimate the odds of being at or below a particular level of the response variable. For example, if there are j levels of ordinal outcomes, the model makes J-1 predictions, each estimating the cumulative probabilities at or below the jth level of the outcome variable. This model can estimate the odds of being at or beyond a particular level of the response variable as well, because below and beyond a

Xing Liu is an Assistant Professor at Eastern Connecticut State University. Email: liux@easternct.edu. This article was presented at the 2007 Annual Conference of the American Educational Research Association (AERA) in Chicago, IL.

particular category are just two complementary directions.

Researchers currently have a variety of options when choosing statistical software packages that can perform ordinal logistic regression models. For example, some general purpose statistical packages, such as Stata, SAS and SPSS, all provide the options of analyzing proportional odds models. However, these statistical packages may use different techniques to estimate the ordinal logistic models. Long and Freese (2006) noted that Stata estimates cut-points in the ordinal logistic model while setting the intercept to be 0; other statistical software packages might estimate intercepts rather than cut-points. Agresti (2002) introduced both the proportional odds model and the latent variable model, and stated that parameterization in SAS (Proc Logistic) followed the formulation of the proportional odds model rather than the latent variable model. Hosmer and Lemeshow (2000) used a formulation which was consistent with Stata's expression to define the ordinal regression model by negating the logit coefficients.

Because statistical packages may estimate parameters in the ordinal regression model differently following different equations, the outputs they produce may not be the same, and thus they seem confusing to applied

statisticians and researchers. Researchers are more likely to make mistakes in interpreting the results if ignoring the differences in parameter estimations using different software packages.

It is the aim of the article to clarify the misunderstanding and confusion when fitting ordinal regression models. To date, no study has been conducted to demonstrate fitting the proportional odds model using three general-purpose statistical packages, comparing differences and identifying similarities among them. Thus, this article seeks to fill this gap by: (1) demonstrating the use of Stata, SAS and SPSS to fit the proportional odds model; and (2) comparing the features and results for fitting the proportional odds model using Stata OLOGIT, SAS PROC LOGISTIC (ascending and descending), and SPSS PLUM. Data from a survey instrument TPGP (Teachers' Perceptions of Grading Practices) is used to demonstrate the PO analysis.

Theoretical Framework

In an ordinal logistic regression model, the outcome variable is ordered, and has more than two levels. For example, students' SES is ordered from low to high; childrens' proficiency in early reading is scored from level 0 to 5; and a response scale of a survey instrument is ordered from strongly disagree to strongly agree. One appealing way of creating the ordinal variable is via categorization of an underlying continuous variable (Hosmer & Lemeshow, 2000).

In this article, the ordinal outcome variable is teachers' teaching experience level, which is coded as 1, 2, or 3 (1 = low; 2 = medium; and 3 = high) and is categorized based on a continuous variable, teaching years. Teachers with less than five years of experience are categorized in the low teaching experience level; those with between 6 and 15 years are categorized in the medium level; and teachers with 15 years or more are categorized in the high level. The distribution of teaching years is highly positively skewed. The violation of the assumption of normality makes the use of Multiple Regression inappropriate. Therefore, the ordinal logistic regression is the most appropriate model for analyzing the ordinal outcome variable in this case.

A Latent-Variable Model

The ordinal logistic regression model can be expressed as a latent variable model (Agresti, 2002; Greene, 2003; Long, 1997, Long & Freese, 2006; Powers & Xie, 2000; Wooldridge & Jeffrey, 2001). Assuming a latent variable, Y^* exists, $Y^* = \mathbf{x}\boldsymbol{\beta} + \varepsilon$, can be defined where $\mathbf{x}$ is a row vector (1* k) containing no constant, $\boldsymbol{\beta}$ is a column vector (k*1) of structural coefficients, and ε is random error with standard normal distribution: $\varepsilon \sim N(0, 1)$.

Let Y^* be divided by some cut points (thresholds): $\alpha_1, \alpha_2, \alpha_3 \ldots \alpha_j$, and $\alpha_1 < \alpha_2 < \alpha_3 \ldots < \alpha_j$. Considering the observed teaching experience level is the ordinal outcome, y, ranging from 1 to 3, where 1= low, 2 = medium and 3 = high, define:

$$Y = \begin{cases} 1 & if\ y^* \leq \alpha_1 \\ 2 & if\ \alpha_1 < y^* \leq \alpha_2 \\ 3 & if\ \alpha_2 < y^* \leq \infty \end{cases}$$

Therefore, the probability of a teacher at each experience level can be computed. For example,

$$P(y = 1) = P(y^* \leq \alpha_1)$$
$$= P(\mathbf{x}\boldsymbol{\beta} + \varepsilon \leq \alpha_1)$$
$$= F(\alpha_1 - \mathbf{x}\boldsymbol{\beta});$$

$$P(y = 2) = P(\alpha_1 < y^* \leq \alpha_2)$$
$$= F(\alpha_2 - \mathbf{x}\boldsymbol{\beta}) - F(\alpha_1 - \mathbf{x}\boldsymbol{\beta});$$

$$P(y = 3) = P(\alpha_2 < y^* \leq \infty)$$
$$= 1 - F(\alpha_2 - \mathbf{x}\boldsymbol{\beta});$$

The cumulative probabilities can also be computed using the form:

$$P(Y \leq j) = F(\alpha_j - \mathbf{x}\boldsymbol{\beta}), \text{ where } j = 1, 2, \ldots J\text{-}1. \quad (1)$$

General Logistic Regression Model

In a binary logistic regression model, the response variable has two levels, with 1 = success of the events, and 0 = failure of the events. The probability of success is predicted on a set of predictors. The logistic regression model can be expressed as:

$$\ln(Y') = \text{logit}\,[\pi(\underline{x})]$$

$$= \ln\left(\frac{\pi(\underline{x})}{1-\pi(\underline{x})}\right)$$

$$= \alpha + \beta_1 X_1 + \beta_2 X_2 + \ldots \beta_p X_p. \qquad (2)$$

In Stata, the ordinal logistic regression model is expressed in logit form as follows:

$$\ln(Y_j') = \text{logit}\,[\pi(x)]$$

$$= \ln\left(\frac{\pi_j(x)}{1-\pi_j(x)}\right)$$

$$= \alpha_j + (-\beta_1 X_1 - \beta_2 X_2 - \ldots -\beta_p X_p), \qquad (3)$$

where $\pi_j(\underline{x}) = \pi(Y \le j | x_1, x_2, \ldots x_p)$, which is the probability of being at or below category j, given a set of predictors. j = 1, 2, … J -1. α_j are the cut points, and $\beta_1, \beta_2 \ldots \beta_p$ are logit coefficients. This is the form of a Proportional Odds (PO) model because the odds ratio of any predictor is assumed to be constant across all categories. Similar to logistic regression, in the proportional odds model we work with the logit, or the natural log of the odds. To estimate the ln (odds) of being at or below the jth category, the PO model can be rewritten as:

$$\text{logit}\,[\pi(Y \le j | x1, x2, \ldots x_p)]$$

$$= \ln\left(\frac{\pi(Y \le j | x_1, x_2, \ldots x_p)}{\pi(Y > j | x_1, x_2, \ldots x_p)}\right)$$

$$= \alpha_j + (-\beta_1 X_1 - \beta_2 X_2 - \ldots -\beta_p X_p) \qquad (4)$$

Thus, this model predicts cumulative logits across J -1 response categories. By transforming the cumulative logits, we can obtain the estimated cumulative odds as well as the cumulative probabilities being at or below the jth category.

SAS uses a different ordinal logit model for estimating the parameters from Stata. For SAS PROC LOGISTIC (the ascending option), the ordinal logit model has the following form:

$$\text{logit}\,[\pi(Y \le j | x_1, x_2, \ldots x_p)]$$

$$= \ln\left(\frac{\pi(Y \le j | x_1, x_2, \ldots x_p)}{\pi(Y > j | x_1, x_2, \ldots x_p)}\right)$$

$$= \alpha_j + \beta_1 X_1 + \beta_2 X_2 + \ldots +\beta_p X_p; \qquad (5)$$

Using SAS with the descending option, the ordinal logit model can be expressed as:

$$\text{logit}\,[\pi(Y \ge j | x_1, x_2, \ldots x_p)]$$

$$= \ln\left(\frac{\pi(Y \ge j | x_1, x_2, \ldots x_p)}{\pi(Y < j | x_1, x_2, \ldots x_p)}\right)$$

$$= \alpha_j + \beta_1 X_1 + \beta_2 X_2 + \ldots +\beta_p X_p, \qquad (6)$$

where in both equations α_j are the intercepts, and $\beta_1, \beta_2 \ldots \beta_p$ are logit coefficients.

SPSS PLUM (Polytomous Universal Model) is an extension of the generalized linear model for ordinal response data. It can provide five types of link functions including logit, probit, complementary log-log, cauchit and negative log-log. Just as Stata, the ordinal logit model is also based on the latent continuous outcome variable for SPSS PLUM, it takes the same form as follows:

$$\text{logit}\,[\pi(Y \le j | x_1, x_2, \ldots x_p)]$$

$$= \ln\left(\frac{\pi(Y \le j | x_1, x_2, \ldots x_p)}{\pi(Y > j | x_1, x_2, \ldots x_p)}\right)$$

$$= \alpha_j + (-\beta_1 X_1 - \beta_2 X_2 - \ldots -\beta_p X_p), \qquad (7)$$

where α_j's are the thresholds, and $\beta_1, \beta_2 \ldots \beta_p$ are logit coefficients; j = 1, 2…J-1.

Compared to both Stata and SPSS, SAS (ascending and descending) does not negate the signs before the logit coefficients in the equations, because SAS Logistic procedure (Proc Logistic) is used to model both the dichotomous and ordinal categorical dependent variables, and the signs before the coefficients in the ordinal logit model are kept consistent with those in the binary logistic regression model.

Although the signs in the equations are positive, SAS internally changes the signs of the estimated intercepts and coefficients according to different ordering of the dependent variable (with the ascending or descending option).

Methodology

Sample

The data were collected from teachers at three middle schools and a teacher's training school in Taizhou City, Jiangsu Province, China, using a survey instrument named Teachers' Perceptions of Grading Practices (TPGP) (Liu, 2004; Liu, O'Connell & McCoach, 2006). A total of 147 teachers responded to the survey with the response rate of 73.5%. The outcome variable of interest is teachers' teaching experiences, which is an ordinal categorical variable with 1 = low, 2 = medium and 3 = high.

Explanatory variables included gender (female = 1; male = 2) and a set of scale scores from the TPGP survey instrument The instrument included five scales measuring the importance of grading, the usefulness of grading, student effort influencing grading,

student ability influencing grading, and teachers' grading habits. Composite scale scores were created by taking a mean of all the items for each scale. Table 1 displays the descriptive statistics for these independent variables.

The proportional odds model was first fitted with a single explanatory variable using Stata (V. 9.2) OLOGIT. Afterwards, the full-model was fitted with all six explanatory variables. The assumption of proportional odds for both models was examined using the Brant test. Additional Stata subcommands demonstrated here included FITSTAT and LISTCEOF of Stata SPost (Long & Freese, 2006) used for the analysis of post-estimations for the models. The results of fit statistics, cut points, logit coefficients and cumulative odds of the independent variables for both models were interpreted and discussed. The same model was fit using SAS (V. 9.1.3) (ascending and descending), and SPSS (V. 13.0), and the similarities and differences of the results using all three programs were compared.

Table 1: Descriptive Statistics for All Variables, n = 147

Variable	Teaching Experience Level			
	1 n = 70 47.6%	2 n = 45 30.6%	3 n = 32 21.8%	Total n = 147 100%
% Gender (Female)	74.3%	66.7%	50%	66.7%
Importance	3.33 (.60)	3.31 (.63)	3.55 (.79)	3.37 (.66)
Usefulness	3.71 (.61)	3.38 (.82)	3.70 (.66)	3.60 (.70)
Effort	3.77 (.50)	3.79 (.46)	3.80 (.68)	3.78 (.53)
Ability	3.74 (.40)	3.75 (.54)	3.87 (.51)	3.77 (.47)
Habits	3.38 (.66)	3.57 (.66)	3.49 (.60)	3.46 (.65)

Results

Proportional Odds Model with a Single Explanatory Variable

OLOGIT is the Stata program estimating ordinal logistic regression models of ordinal outcome variable on the independent variables. In this example, the outcome variable, teaching was followed immediately by the independent variable, gender. Figure 1 displays the Stata output for the one-predictor proportional odds model.

The log likelihood ratio Chi-Square test with 1 degree of freedom, LR $\chi^2_{(1)}$ = 5.29, p = .0215, indicated that the logit regression coefficient of the predictor, gender was statistically different from 0, so the full model with one predictor provided a better fit than the null

model with no independent variables in predicting cumulative probability for teaching experience level. The likelihood ratio R^2_L = .0172, which is the Pseudo R^2, and is also called McFadden's R^2, suggested that the relationship between the response variable, teaching experience, and the predictor, gender was small. More measures of fit were obtained using SPost subcommand fitstat (Long & Freese, 2006). In addition to the deviance statistic and McFadden's R^2, several other types of R^2 statistics were reported (Figure2). The information measures, AIC and BIC, were used to compare either nested or non-nested models. Smaller AIC and BIC statistics indicate the better fitting model.

Figure 1: Stata Proportional Odds Model Example: Gender

```
. ologit teaching gender

Iteration 0:   log likelihood = -153.99556
Iteration 1:   log likelihood = -151.35669
Iteration 2:   log likelihood = -151.35194

Ordered logistic regression                    Number of obs    =        147
                                               LR chi2(1)       =       5.29
                                               Prob > chi2      =     0.0215
Log likelihood = -151.35194                    Pseudo R2        =     0.0172


------------------------------------------------------------------------------
    teaching |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      gender |   .7587563   .3310069     2.29   0.022     .1099947    1.407518
-------------+----------------------------------------------------------------
       /cut1 |   .9043487   .4678928                     -.0127044    1.821402
       /cut2 |   2.320024   .5037074                      1.332775    3.307272
------------------------------------------------------------------------------
```

Figure 2: Measures of Fit Statistics

```
. fitstat

Measures of Fit for ologit of teaching

Log-Lik Intercept Only:       -153.996   Log-Lik Full Model:           -151.352
D(144):                        302.704   LR(1):                           5.287
                                         Prob > LR:                       0.021
McFadden's R2:                   0.017   McFadden's Adj R2:              -0.002
ML (Cox-Snell) R2:               0.035   Cragg-Uhler(Nagelkerke) R2:      0.040
McKelvey & Zavoina's R2:         0.038
Variance of y*:                  3.419   Variance of error:               3.290
Count R2:                        0.476   Adj Count R2:                    0.000
AIC:                             2.100   AIC*n:                         308.704
BIC:                          -415.918   BIC':                           -0.297
BIC used by Stata:             317.675   AIC used by Stata:             308.704
```

636

The estimated logit regression coefficient, β = .7588. z = 2.29, p = .022, indicating that gender had a significant effect on teacher's teaching experience level. Substituting the value of the coefficient into the formula (4), logit $[\pi(Y \leq j \mid gender)] = \alpha_j + (-\beta_1 X_1)$, we calculated logit $[\pi(Y \leq j \mid gender)] = \alpha_j - .7588$ (gender). OR $= e^{(-.7588)} = .468$, indicating that male teachers were .468 times the odds for female teachers of being at or below at any category, i.e., female teachers were more likely than male teachers to be at or below a particular category, because males were coded as 2 and girls as 1.

The results table reports two cut-points: _cut1 and_cut2. These are the estimated cut-points on the latent variable, Y*, used to differentiate the adjacent levels of categories of teaching experiences. When the response category is 1, the latent variable falls at or below the first cut point, α_1. When the response category is 2, the latent variable falls between the first cut point α_1 and the second cut point α_2, and when the response category reaches 3 if the latent variable is at or beyond the second cut point α_2.

To estimate the cumulative odds being at or below a certain category, j for gender, the logit form of proportional odds model was used, logit $[\pi(Y \leq j \mid gender)] = \alpha_j - .7588$ (gender). For example, when $Y \leq 1$, α_1, .9043 is the first cut point for the model. Substituting it into the formula (4) results in logit $[\pi(Y \leq j \mid gender)] = .9043 - .7588$ (gender). For girls (x = 1), logit

$[\pi(Y \leq 1 \mid gender)] = .1455$. By exponentiating the logit, the odds for female teachers of being at or below experience category 1 is calculated, $e^{.1455} = 1.157$. For male teachers (x = 2), logit $[\pi(Y \leq 1 \mid gender)] = .9043 - .7588*2 = -.6133$, so the odds for male teachers being at or below teaching experience category 1, $e^{-.6133} = .542$. Odds ratio of male teachers versus female teachers = .542/1.157 = .468. Transforming the cumulative odds, results in the cumulative probabilities by using p = odds/(1+odds).

The Stata program brant was used to test the proportional odds assumption. Brant (1990) proposed a test of proportional odds assumption for the ordinal logistic model by examining the separate fits to the underlying binary logistic models. A non-significant omnibus test indicates that the proportional odds assumption is not violated. It also provides tests for each individual independent variable. When only one independent variable exists in the model, the results of the omnibus test and individual test are the same. The Brant test of parallel regression assumption yields $\chi^2_1 = .40$ (p > .527), indicating that the proportional odds assumptions for the full-model was upheld. This suggests that the effect of gender, the explanatory variable, was constant across separate binary models fit to the cumulative cut points. Figure 3 also shows the estimated coefficient from j-1 binary logistic regression models. Each logistic regression model estimates the probability of being at or beyond teaching experience level *j*.

Figure 3: Brant Test of Parallel Regression (Proportional Odds) Assumption

```
. brant, detail
Estimated coefficients from j-1 binary regressions


               y>1          y>2
gender   .66621777    .91021169
 _cons  -.78882009   -2.5443422


Brant Test of Parallel Regression Assumption

    Variable |      chi2   p>chi2    df
------------+-------------------------
        All |      0.40    0.527     1
------------+-------------------------
     gender |      0.40    0.527     1
---------------------------------------


A significant test statistic provides evidence that the parallel regression
assumption has been violated.
```

The Proportional Odds Model can also estimate the ln(odds) of being at or beyond category j, given a set of predictors. Again, these ln(odds) can be transformed into the cumulative odds, and cumulative probabilities as well. For example, the cumulative probability of a teacher's teaching experience can be estimated at or beyond category 3, $P(Y \geq 3)$, which is the complementary probability when $Y \leq 2$, at or beyond category 2, $P(Y \geq 2)$, and $P(Y \geq 1)$, which equals 1.

In Stata, when estimating the odds of being beyond category j, or at or beyond j+1, the sign of the cut points needs to be reversed and their magnitude remain unchanged because the cut points were estimated from the right to the left of the latent variable, Y*, that is, from the direction when Y = 3 approaches Y = 1. Therefore, two cut points from right to left turn to -2.32 and - .904. When the predictor is dichotomous, a positive sign of the logit coefficient indicates that it is more likely for the group (x = 1) to be at or beyond a particular category than for the relative group (x = 0). When the predictor is continuous, a positive coefficient indicates that when the value of the predictor variable increases, the probability of being at or beyond a particular category increases.

Using Stata syntax listcoef, the odds of being at or beyond a particular category at 2.136 can be obtained, which was constant across all cumulative categories. It also indicated that male teachers were 2.136 times the odds for female teachers of being at or beyond any category, i.e., male teachers were more likely than female teachers to be at or beyond a particular category. Figure 4 displays the results of Stata listcoef. Adding option percent after listcoef, the result of percentage change in odds of being at or beyond a particular category can be obtained when the predictor, gender, goes from males (x = 2) to females (x = 1).

Proportional Odds Model with Six Explanatory Variables

Next, a proportional odds model was fit with eight explanatory variables, which is referred to as the Full Model. Figure 5 displays the results for the fitting of the full model with six explanatory variables.

Before interpreting the results of the full model, the assumption of proportional odds was first examined. The Stata brant command provides the results of the Brant test of parallel regression (Proportional Odds) assumption for the full model with six predictors and tests for each independent variable. It also provides the estimated coefficient from j-1 binary logistic regression models results of two separate binary logistic regression models. The data are dichotomized according to the cumulative probability pattern so that each logistic

Figure 4: Results of Stata listcoef

```
. listcoef, help

ologit (N=147): Factor Change in Odds

  Odds of: >m vs <=m


  --------------------------------------------------------------------
    teaching |      b        z      P>|z|     e^b     e^bStdX     SDofX
  -------------+------------------------------------------------------
      gender |  0.75876    2.292    0.022    2.1356    1.4318    0.4730
  --------------------------------------------------------------------
        b = raw coefficient
        z = z-score for test of b=0
    P>|z| = p-value for z-test
      e^b = exp(b) = factor change in odds for unit increase in X
  e^bStdX = exp(b*SD of X) = change in odds for SD increase in X
    SDofX = standard deviation of X
```

regression model estimates the probability of being at or beyond teaching experience level *j*. For the omnibus Brant test, $\chi^2_6 = 8.10$, p = .230, indicating that the proportional odds assumptions for the full-model was upheld. Examining the Brant tests for each individual independent variable indicated that the Brant test of the assumption of parallel regression (proportional odds) were upheld for gender, importance, effort, ability and habits. For usefulness, the Brant test, $\chi^2_1 = 4.03$, p = .045, which is very close to .05, therefore, it may also be concluded that the PO assumption for this variable is nearly upheld. Checking the estimated coefficients for each independent variable across two binary logistic regression models shows that the logit coefficients for all

the variables were similar across two binary logistic models, supporting the results of the Brant test of proportional odds assumption.

The log likelihood ratio Chi-Square test, LR $\chi^2_{(6)}$ = 13.738, p = .033, indicating that the full model with six predictor provided a better fit than the null model with no independent variables in predicting cumulative probability for teaching experience. The likelihood ratio R^2_L = .045, much larger than that of the gender-only model, but still small, suggesting that the relationship between the response variable, teaching experience, and six predictors, was still small. Compared with the gender-only model, all R^2 statistics of the full-model shows improvement (see Figure 7).

Figure 6: Brant Test of Parallel Regression (Proportional Odds) Assumption

```
. brant, detail

Estimated coefficients from j-1 binary regressions

                    y>1          y>2
    gender    .74115294    .86086025
importance    .64416122    .46874536
usefulness   -.94566294   -.19259753
    effort    .09533898   -.03621639
   ability    .26862373    .68349765
    habits    .48959286   -.02795948
     _cons   -2.7097459   -5.7522624

Brant Test of Parallel Regression Assumption

    Variable |      chi2    p>chi2     df
-------------+-------------------------------
         All |      8.10     0.231      6
-------------+-------------------------------
      gender |      0.08     0.772      1
  importance |      0.24     0.622      1
  usefulness |      4.03     0.045      1
      effort |      0.10     0.746      1
     ability |      0.66     0.418      1
      habits |      2.15     0.142      1
---------------------------------------------

A significant test statistic provides evidence that the parallel
regression assumption has been violated.
```

Figure 7: Measure of Fit Statistics for Full-Model

```
. fitstat

Measures of Fit for ologit of teaching

Log-Lik Intercept Only:        -153.996   Log-Lik Full Model:            -147.127
D(139):                         294.253    LR(6):                          13.738
                                           Prob > LR:                       0.033
McFadden's R2:                    0.045    McFadden's Adj R2:              -0.007
ML (Cox-Snell) R2:                0.089    Cragg-Uhler(Nagelkerke) R2:      0.102
McKelvey & Zavoina's R2:          0.098
Variance of y*:                   3.646    Variance of error:               3.290
Count R2:                         0.429    Adj Count R2:                   -0.091
AIC:                              2.111    AIC*n:                         310.253
BIC:                           -399.417    BIC':                           16.205
BIC used by Stata:              334.177    AIC used by Stata:             310.253
```

The Stata listcoef command (Figure 8) produced more detailed results of logit coefficients and cumulative odds (exponentiated coefficients). For the proportional odds model, interpretation of cumulative odds is independent on the ancillary parameters (cut points) because they are constant across all levels of the response variable.

The effects of the independent variables can be interpreted in several ways, including how they contribute to the odds and their probabilities of being at or beyond a particular category. They can also be interpreted as how these variables contribute to the odds of being at or below a particular category, if the sign is reversed before the estimated logit coefficients and corresponding cumulative odds are computed. In terms of odds ratios, male teachers were 2.241 times the odds for female teachers to be at or beyond a particular category (OR=2.241), after controlling the effects of other predictors in the model. The usefulness of grading with a corresponding OR significantly less than 1.0 has significant negative effects in the model. These cumulative odds are associated with a teacher being in lower teaching experience categories rather than in higher categories. For a one unit increase in the usefulness of grading, the odds ratio of being in higher teaching experience categories versus lower categories was .53 times lower, after controlling for the effects of other variables. However, variables whose corresponding ORs are significantly greater than 1.0 have significant

positive effects on the response variable in the model. For example, the importance of grading (OR=1.778) had a positive effect on teachers being in higher teaching experience categories. For a one unit increase in the importance of grading, the odds ratio of being in higher teaching experience categories versus lower categories was 1.778 times greater, given the effects of other predictors are held constant. Variables such as student ability and teacher's grading habits, whose corresponding ORs were greater than 1.0, but were not statistically significant, had positive effects on the response variable, but these effects may be due to chance and need further investigation. Independent variables with ORs close to 1.0 have no effect on the response variable. For example, student effort influencing grading was not associated with teaching experience in this model (OR=1.0266, p=.946).

Comparison of Results of a Single-Variable PO Model Using Stata, SAS, and SPSS

Table 2 shows a comparison of the results for Stata OLOGIT with results from SAS PROC LOGISTIC with the ascending and descending options, and SPSS PLUM. The similarities and differences between these results should be noted, otherwise, it could be misleading to interpret the results in the same way, disregarding their different parameterizations. In estimating proportional odds models, Stata sets the intercept to 0, and estimates the cut points, while SAS ascending

Figure 8: Results of Logit Coefficient, Cumulative Odds, and Percentage Change in Odds

```
. listcoef, help

ologit (N=147): Factor Change in Odds

  Odds of: >m vs <=m

  -------------------------------------------------------------------
   teaching |      b         z      P>|z|     e^b     e^bStdX    SDofX
  ----------+--------------------------------------------------------
     gender |  0.80695     2.318    0.020    2.2411    1.4648    0.4730
 importance |  0.57547     1.897    0.058    1.7780    1.4601    0.6578
 usefulness | -0.63454    -2.322    0.020    0.5302    0.6402    0.7029
     effort |  0.02625     0.068    0.946    1.0266    1.0140    0.5283
    ability |  0.34300     0.825    0.409    1.4092    1.1752    0.4707
     habits |  0.31787     1.088    0.277    1.3742    1.2282    0.6466
  -------------------------------------------------------------------
       b = raw coefficient
       z = z-score for test of b=0
   P>|z| = p-value for z-test
     e^b = exp(b) = factor change in odds for unit increase in X
 e^bStdX = exp(b*SD of X) = change in odds for SD increase in X
   SDofX = standard deviation of X
```

estimates the intercepts and set the cut points to 0. Comparing Stata with SAS (ascending), the different choice of parameterization does not influence the magnitude of cut points (or intercepts) and coefficients. However, it does determine the sign before these estimates.

When estimating the odds of being at or below a response category, the estimates for the cut points using Stata are the same as the intercepts using SAS ascending in both sign and magnitude. The first cut point, α_1 in Stata estimation is the same as the first intercept α_1 in SAS ascending estimation, because there is no first intercept α_1 in Stata estimation. Using Stata and SAS (the ascending option), the estimated logit coefficients are the same in magnitude but are opposite in sign. Using Stata, the estimated logit coefficient $\beta = .759$. Substituting it into the logit form (4), we get logit $[\pi(Y \leq j \mid gender)] = \alpha_j - (.759)*(gender) = \alpha_j - .759*(gender)$. OR = $e^{(-.759)} = .468$, indicating that male teachers were .468 times the odds for female teachers of being at or below at any category, that is, female teachers were more likely than male teachers to be at or below a particular teaching experience level. Using SAS ascending, the estimated logit

coefficient, $\beta = -.759$. Substituting it into its corresponding logit form (5) results in the same equation: logit $[\pi(Y \leq j \mid gender)] = \alpha_j - .759*(gender)$. Therefore, the same results of estimated cumulative odds and cumulative probability were obtained using Stata and SAS ascending.

Comparing the results of the proportional odds model using Stata and SAS with the descending option, it was found that estimated cut points for Stata and the estimated intercepts for SAS descending are the same in magnitude but are opposite in sign. Using Stat and SAS descending, the estimated logit coefficients are the same in both magnitude and sign. To estimate the odds of being at or beyond a particular teaching experience level using Stata, it is only necessary to reverse the sign of the estimated cut points. The estimated logit coefficient is $\beta = .759$. Exponentiating this results in $e^{(.759)} = 2.136$, indicating male teachers are 2.136 times greater than female teachers to be at or beyond a particular category. In other words, female teachers are less likely than male teachers to be at or beyond a certain category.

Table 2: Results of Proportional Odds Model with a Single Variable Using Stata, SAS
(Ascending and descending) and SPSS: A Comparison

Model Estimates	STATA	SAS (Ascending)	SAS (Descending)	SPSS
	$P(Y \leq j)$	$P(Y \leq j)$	$P(Y \geq j)$	$P(Y \leq j)$
Cutpoints (Stata)/ Intercept (SAS)/ Threshold (SPSS)	_cut1(α_1) = .904	α_1 = .904	α_3 = -2.32	α_1 = -.613
	_cut2(α_2) = 2.32	α_2 = 2.32	α_2 = -.904	α_2 = .803
Gender (Male = 2)	.759	-.759	.759	0
Gender (Female = 1)				-.759
LR R^2	.017	.017	.017	.017
Brant Test (Omnibus Test)[a]	χ^2_1 = .40 (p > .527)			
Score Test[b]		χ^2_1 = .4026 (p = .5258)	χ^2_1 = .4026 (p = .5258)	χ^2_1 = .392 (p > .530)
Model Fit	LR $\chi^2_{(1)}$ = 5.29, p = .0215	LR $\chi^2_{(1)}$ = 5.29, p = .0215	LR $\chi^2_{(1)}$ = 5.29, p = .0215	LR $\chi^2_{(1)}$ = 5.287, p = .021

a. Brant test for proportional odds assumption.
b. Score test for proportional odds assumption.

Using Stata and SPSS, when estimating the effects of predictors on the log odds of being at or below a certain category of the outcome variable, the sign before the coefficients are both minus rather than plus. In other words, the effects of predictors are subtracted from the cut points or thresholds. SPSS PLUM labels the estimated logits for the predicator variables LOCATION. When the predicator variable is continuous, the estimated logit coefficients are the same as those estimated by Stata OLOGIT in both magnitude and sign. However, SPSS PLUM is different from Stata OLOGIT in this aspect: when the predictor variable is categorical, for example gender, with 1 = female and 2 = male, the estimated coefficient is only displayed for the category with smaller value, i.e., when gender = 1. The category with larger value, gender = 2, is the reference category, and has an estimate of 0. If gender is coded with 1 = female and 0 = male, the estimated coefficient is displayed for the case when gender = 0, and the estimated coefficient for female (gender = 1) is 0. Using SPSS PLUM, the estimated logit coefficient, β = -.759 for the case when female = 1, and β = 0 for the case when male = 2. Substituting it into the logit form (7) results in logit $[\pi(Y \leq j \mid gender)] = \alpha_j - (-.759)*(gender) = \alpha_j + .759*(gender)$. By exponentiating, OR = e $^{(.759)}$ = 2.136, indicating that female teachers are

2.136 times the odds for male teachers of being at or below at a particular teaching experience level. This result is equivalent to that of Stata.

The parameter estimation for the categorical predictor in SPSS PLUM makes the threshold values in the ordinal logit model different from those estimated by Stata OLOGIT. These differences can be observed in the results of the proportional odds model using Stata, SAS (ascending and descending), and SPSS (Table 2). In SPSS PLUM, the threshold estimates are for the case when gender = 2 (male teachers), while in Stata and SAS, the cut points or intercepts are for case when gender = 1 (female students).

Equivalent results of estimated logit can be obtained using different estimates of cutpoints (thresholds) and logit coefficients fitted by Stata and SPSS. For example, using SPSS, the predicted logit for male teachers (gender = 2) of being at or below teaching experience level 1, logit $[\pi(Y \leq 1 \mid gender)] = \alpha_1 - 0*(gender) = -.613 + 0*(2) = -.613$; the predicted logit for female teachers (gender = 1) of being at or below teaching experience level 1, logit $[\pi(Y \leq 1 \mid gender)] = \alpha_1 - (-.759)*(gender) = -.613 + .759*1 = .146$. Using Stata, the predicted logit for male teachers (gender = 2) of being at or below teaching experience level 1, logit $[\pi(Y \leq 1 \mid gender)] = \alpha_1 - (.759)*(gender) = .904 - .759*2 = -.614$; the predicted logit for female teachers (gender =1) of being at or below teaching experience level 1, logit $[\pi(Y \leq 1 \mid gender)] = \alpha_1 - (.759)*(gender) = .904 - .759*1 = .145$.

To test the proportional odds assumption, Stata uses the Brant test of parallel regression assumption with the result $\chi^2_1 = .40$ p > .527; SAS uses ascending and descending score test and has the same results $\chi^2_1 = .4026$, p = .5258; SPSS uses a test of parallel lines with the result $\chi^2_1 = .392$, p > .530. All tests produce similar results in that the proportional odds model assumption is upheld. Across the models, the omnibus likelihood ratio tests produce the same results, indicating the proportional odds model with one variable (gender) has better fit than the null model. Features of the ordinal logistic regression analysis using Stata, SAS and SPSS are shown and compared in Table 3.

Conclusion

In this article, the use of proportional odds models was illustrated to predict teachers' teaching experience level from a set of measures of teachers' perceptions of grading practices. A single independent variable model and a full-model with six independent variables were fitted and compared. The assumptions of proportional odds for both models were examined. It was found that the assumption of proportional odds for both the single-variable model and the full-model was upheld.

Results from the proportional odds model revealed that the usefulness of grading had a negative effect on the prediction of teaching experience level (OR = .53), while the importance of grading practices had a positive effect on the experience level (OR = 1.78), after controlling for the effects of other variables. Although student effort influencing teachers' grading practices (OR = 1.41) and teachers' grading habits (OR = 1.37) had positive effects on teaching experience level, these effects were not found to be significant. Compared to male teachers, female teachers were more likely to be at or below a particular category, or in other words, males were more likely to be at or beyond an experience level. Student effort influencing grading was not associated with teachers' teaching experience level in the model.

These findings suggest that teachers with longer teaching experience tended to feel the grading practices are more important than the teachers with fewer years of teaching. However, teachers with longer teaching experiences tended to doubt the usefulness of grading in their teaching; this may be due in part to their requirement of conducting test-oriented teaching in China. In addition, the gender difference suggests that female teachers were more easily categorized as inexperienced teachers; this may be due to greater numbers of female students receiving the opportunities of higher education in recent years and their choosing teaching as their profession. The frequencies of new female teachers are currently greater than those of new male teachers in China.

Comparing the results using Stata and SAS, it was found that both packages produced the same or similar results in model fit statistics,

Table 3: Feature Comparisons of the Ordinal Logistic Regression Analysis Using Stata, SAS and SPSS

	STATA	SAS	SPSS
Model Specification			
Cutpoints/Thresholds	✓		✓
Intercept		✓	
Test Hypotheses of Logit Coefficients	✓	✓	✓
Maximum Likelihood Estimates			
Odds Ratio	✓	✓	
z-statistic or Wald Test for Parameter Estimate	✓		✓
Chi-square Statistic for Parameter Estimate		✓	
Confidence Interval for Parameter Estimate	✓		✓
Fit Statistics			
Loglikelihood	✓	✓	✓
Goodness-of-Fit Test	✓	✓	✓
Pseudo R-Square	✓	✓	✓
Test of PO Assumption			
Omnibus Test of Assumption of Proportional Odds	✓	✓	✓
Test of Assumption of Proportional Odds for Individual Variables	✓		
Association of Predicted Probabilities and Observed Responses		✓	

and the test of proportional odds assumption. The estimated coefficients and cut points (thresholds) were the same in magnitude but may be reversed in sign. Comparing the results using Stata and SPSS, it was found that although the ordinal logit models are based on latent continuous response variables for both packages, SPSS PLUM estimated the logit coefficient for the category with smaller value when the predictor variable was categorical, and thus the estimated thresholds were different from those estimated by Stata. Researchers should understand the differences of parameterization of ordinal logistic models using Stata and other statistical packages. Researchers should pay attention to the sign before the estimated logit coefficients and the cut points in the model, and

exercise caution in interpreting the results.

In educational research, ordinal categorical data is frequently used and researchers need to understand and be familiar with the ordinal logistic regression models dealing with the internally ordinal outcome variables. In some situations, Ordinary Least Squares (OLS) techniques may be used for preliminary analysis of such data by treating the ordinal scale variable as continuous. However, ignoring the discrete ordinal nature of the variable would cause the analysis lose some useful information and could lead to misleading results. Therefore, it is crucial for researchers to use the most appropriate models to analyze ordinal categorical dependent variables. In addition, the role of any statistical software

package is a tool for researchers. The choice of software is the preference of researchers; it is therefore not the purpose of the study to suggest which one is the best for ordinal logistic regression analysis. This demonstration clarifies some of the issues that researchers must consider in using different statistical packages when analyzing ordinal data.

References

Agresti, A. (2002). *Categorical data analysis* (*2nd Ed.*). NY: John Wiley & Sons.

Agresti, A. (1996). *An introduction to categorical data analysis*. NY: John Wiley & Sons.

Allison, P. D. (1999). *Logistic regression using the SAS system: Theory and application*. Cary, NC: SAS Institute, Inc.

Ananth, C. V., & Kleinbaum, D. G. (1997). Regression models for ordinal responses: A review of methods and applications. *International Journal of Epidemiology*, *26*, 1323-1333.

Armstrong, B. B., & Sloan, M. (1989). Ordinal regression models for epidemiological data. *American Journal of Epidemiology*, *129*(*1*), 191-204.

Bender, R., & Benner, A. (2000). Calculating ordinal regression models in SAS and S-Plus. *Biometrical Journal*, *42*(*6*), 677-699.

Brant, R. (1990). Assessing proportionality in the proportional odds model for ordinal logistic regression. *Biometrics*, *46*, 1171-1178.

Clogg, C. C., & Shihadeh, E. S. (1994). *Statistical models for ordinal variables*. Thousand Oaks, CA: Sage.

Greene, W. H. (2003). *Econometric analysis* (*5th Ed.*). Upper Saddle River, NJ: Prentice Hall.

Hosmer, D. W., & Lemeshow, S. (2000). *Applied logistic regression* (*2nd Ed.*). NY: John Wiley & Sons.

Liu, X. (2004). *A validation of teachers' perceptions of grading practices*. Paper presented at the October Annual Conference of the Northeastern Educational Research Association (NERA), Kerhonkson, NY.

Liu, X., O'Connell, A. A., & McCoach D. B. (2006). *The initial validation of teachers' perceptions of grading practices*. Paper presented at the April 2006 Annual Conference of the American Educational Research Association (AERA), San Francisco, CA.

Long, J. S. (1997). *Regression models for categorical and limited dependent variables*. Thousand Oaks, CA: Sage.

Long, J. S., & Freese, J. (2006). *Regression models for categorical dependent variables using Stata* (*2nd Ed.*). Texas: Stata Press.

McCullagh, P. (1980). Regression models for ordinal data (with discussion). *Journal of the Royal Statistical Society*, *B*, *42*, 109-142.

McCullagh, P., & Nelder, J. A. (1989). *Generalized linear models* (*2nd Ed.*). London: Chapman and Hall.

Menard, S. (1995). *Applied logistic regression analysis*. Thousand Oaks, CA: Sage.

O'Connell, A. A., (2000). Methods for modeling ordinal outcome variables. *Measurement and Evaluation in Counseling and Development*, *33*(*3*), 170-193.

O'Connell, A. A. (2006). *Logistic regression models for ordinal response variables*. Thousand Oaks, CA: SAGE.

O'Connell, A. A., Liu, X., Zhao, J., & Goldstein, J. (2006). *Model Diagnostics for proportional and partial proportional odds models*. Paper presented at the April 2006 Annual Conference of the American Educational Research Association (AERA), San Francisco, CA.

Powers, D. A., & Xie, Y. (2000). *Statistical models for categorical data analysis*. San Diego, CA: Academic Press.

Wooldridge, J. M. (2001). *Econometric analysis of cross section and panel data*. Cambridge, MA: The MIT Press.