Ordinal Regression Analysis: Predicting Mathematics Proficiency Using the Continuation Ratio Model

Xing Liu
Eastern Connecticut State University, liux@easternct.edu

Ann A. O'Connell
The Ohio State University, aoconnell@ehe.osu.edu

Hari Koirala
Eastern Connecticut State University, koiralah@easternct.edu

Ordinal Regression Analysis: Predicting Mathematics Proficiency Using the Continuation Ratio Model

Xing Liu
Eastern Connecticut State University,
Willimantic, CT

Ann A. O'Connell
The Ohio State University,
Columbus OH

Hari Koirala
Eastern Connecticut State University,
Willimantic, CT

One commonly used model to analyze ordinal response data is the proportional odds (PO) model. However, if research interest is focused on a particular category and if an individual must pass through lower categories before achieving a higher level, the continuation ratio (CR) model is a more appropriate choice than the PO model. In addition, statistical software, such as Stata and SAS, may use different techniques to estimate the parameters. The CR model is used to illustrate the analysis of ordinal data in education using Stata and SAS and compares the results of fitting the CR model between these two packages.

Key words: Continuation ratio models, proportional odds models, ordinal regression analysis, mathematics proficiency, Stata, SAS, comparison.

Introduction

Ordinal data are abundantly collected in educational research. For example, it is common for data on student's SES to be ordered from low to high, responses to a survey item scaled from strongly disagree to strongly agree, children's reading proficiency scored from level 0 to 5 or students' educational proficiency levels in a state test ranging from fail to pass to proficient. One commonly used model to analyze ordinal data is the proportional odds (PO), or cumulative odds, model (Agresti, 1996, 2002, 2007; Armstrong & Sloan, 1989; Hilbe, 2009; Liu; 2009; Long, 1997, Long & Freese, 2006; McCullagh, 1980; McCullagh & Nelder, 1989;

Xing Liu is an Associate Professor of Research and Assessment in the Education Department. Email him at: liux@easternct.edu. Ann A. O'Connell is a Professor in the Program in Quantitative Research, Evaluation, and Measurement (QREM) in the School of Educational Policy and Leadership within the College of Education and Human Ecology. Email her at: aoconnell@ehe.osu.edu. Hari Koirala is a Professor in the Education Department. Email: koiralah@easternct.edu.

O'Connell, 2000, 2006; O'Connell & Liu, 2011; Powers & Xie, 2000).

The PO model is used to estimate the cumulative probability of being at or below a particular level of a response variable, or being beyond a particular level, which is the complementary direction. However, when research is focused on a particular category, rather than at or below that category, given that an individual has achieved a higher level, the continuation ratio (CR) model (Fienberg, 1980; Hardin & Hilbe, 2007; Long & Freese, 2006) is a more appropriate choice than the PO model. In particular, the CR model is more appealing than other models when analyzing educational attainment data (Allison, 1999). The CR model is very useful in analyzing data such as student academic proficiency levels that are measured annually or frequently using a mastery test as under the No Child Left Behind Act (NCLB).

In a CR model, the ordinal categories represent successive stages, or proficiency levels, through which an individual can progress; for example, faculty ranks from assistant professor to associate professor to full professor, or educational attainment from high school diploma to Bachelor's degree, Master's degree and to doctorate degree. In both of these examples, individuals must pass through lower stages or levels in order to reach higher stages or

levels. A CR model estimates the odds of being in a certain category relative to being beyond that category. In terms of probability, this model estimates the probability of being in a category, given that an individual has been in that category or beyond. In addition, because these two conditional probabilities are complementary, the model estimates the conditional probability of being beyond a category given a person has attained that particular category.

Although the PO model is commonly used, the CR model seems to be overlooked. In addition, not all general-purpose statistical software packages have developed procedures to directly estimate a CR model, and for those packages which are capable of conducting a CR analysis, they may use different parameterizations to estimate the model. However, no study has been conducted to identify these differences and clarify misunderstandings.

Ignoring these differences may result in erroneous interpretations of results. Therefore, it is critical for researchers to understand this model and apply it correctly. To fill this gap, this study was conducted to demonstrate the use of the continuation ratio (CR) model to predict the mathematics proficiency of high school students using Stata and SAS, and to compare the results of fitting the continuation ratio model between these two packages. Ordinal regression analyses were based on the data from the Educational Longitudinal Study of 2002 (ELS:2002) in which the ordinal outcome of students' mathematics proficiency was predicted from a set of students' classroom activities, such as, reviewing work from the previous day in math class, listening to teachers' lectures, copying notes from the board, using books besides textbooks, doing problem solving in class, using general and graphing calculators, using computers, explaining work orally and participating in student-led discussions.

Theoretical Framework: General Logistic Regression Model and the Proportional Odds Model

The binary logistic regression model predicts an outcome variable with two categories, with 1 = experiencing the event, and categories, with 1 = experiencing the event, and 0 = not experiencing the event. This model estimates the log odds of the outcome, and thus the probability of success on a set of predictors. The logistic regression model has the following form:

$$
\begin{aligned}
\ln(Y') &= \text{logit}\,[\pi(x)] \\
&= \ln\left(\frac{\pi(\underline{x})}{1-\pi(\underline{x})}\right) \\
&= \alpha + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p
\end{aligned}
\tag{1}
$$

An ordinal logistic regression model is a generalization of a binary logistic regression model when the outcome variable has more than two ordinal levels. It estimates the probability of being at or below a specific outcome level, conditional on a collection of explanatory variables. 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^* can be defined as a function of a set of predictor variables and a random error. Let Y^* be divided by some cut points (thresholds): $\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_j$, and $\alpha_1 < \alpha_2 < \alpha_3 \ldots < \alpha_j$. The values of the observed ordinal variable, Y, fall within the regions divided by these cut points (thresholds). For example, $Y = 0$, if $Y^* \leq \alpha_1$. The observed mathematics proficiency level is the ordinal outcome, y, ranging from 0 to 5, is defined as follows:

$$
y = \begin{cases}
0 & \text{if } y^* \leq \alpha_1 \\
1 & \text{if } \alpha_1 < y^* \leq \alpha_2 \\
2 & \text{if } \alpha_2 < y^* \leq \alpha_3 \\
3 & \text{if } \alpha_3 < y^* \leq \alpha_4 \\
4 & \text{if } \alpha_4 < y^* \leq \alpha_5 \\
5 & \text{if } \alpha_5 < y^* \leq \infty
\end{cases}
\tag{2}
$$

Therefore, the probability of a student achieving each proficiency level and the cumulative probabilities as can both be predicted by: $P(Y \leq j) = F(\alpha_j - x\beta)$, where $j = 1, 2, \ldots, J-1$.

Because different software packages utilize different parameterizations in estimating logit coefficients, the ordinal logistic regression model can be expressed in different forms (Liu, 2009). In Stata, it is expressed in logit form as follows:

$$\ln(Y_j') = \text{logit}\,[\pi(x)]$$
$$= \ln\left(\frac{\pi_j(x)}{1-\pi_j(x)}\right)$$
$$= \alpha_j + \left(-\beta_1 X_1 - \beta_2 X_2 - \ldots - \beta_p X_p\right),$$
$$(3)$$

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; $j = 1, 2, \ldots, J-1$. α_j are the cut points, and $\beta_1, \beta_2, \ldots, \beta_p$ are logit coefficients. To estimate the ln (odds) of being at or below the jth category, the PO model can be rewritten as:

$$\text{logit}\,[\pi(Y \le j | x_1, x_2, \ldots, x_p)]$$
$$= \ln\left(\frac{\pi(Y \le j | x_1, x_2, \ldots, x_p)}{\pi\left(Y > j | x_1, x_2, \ldots, x_p\right)}\right)$$
$$= \alpha_j + \left(-\beta_1 X_1 - \beta_2 X_2 - \ldots - \beta_p X_p\right).$$
$$(4)$$

This is the form of the proportional odds (PO) model because it assumes that the logit coefficients of any predictor are identical across all comparisons; this equal logit slope assumption can be assessed by the Brant test (Brant, 1990). Similar to the binary logistic regression, the PO model estimates the logit, or the log of the odds of being at or below a particular category versus being beyond that category. Thus, this model predicts cumulative logits across J−1 response categories. Methods of model diagnostics for the ordinal logistic regression models are provided by O'Connell and Liu (2011).

Just as Stata, the ordinal logit model is also based on the latent continuous outcome variable for SPSS PLUM, and it takes the same form. However, 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:

$$\text{logit}\,[\pi(Y \le j | x_1, x_2, \ldots, x_p)]$$
$$= \ln\left(\frac{\pi(Y \le j | x_1, x_2, \ldots, x_p)}{\pi\left(Y > j | x_1, x_2, \ldots, x_p\right)}\right)$$
$$= \alpha_j + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p.$$
$$(5)$$

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

$$\text{logit}\,[\pi(Y \ge j | x_1, x_2, \ldots, x_p)]$$
$$= \ln\left(\frac{\pi(Y \ge j | x_1, x_2, \ldots, x_p)}{\pi\left(Y < j | x_1, x_2, \ldots, x_p\right)}\right)$$
$$= \alpha_j + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p.$$
$$(6)$$

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

Theoretical Framework: The Continuation Ratio Model

As notes, statistical software packages, such as Stata, SAS and SPSS, use different techniques to estimate the parameters in the proportional odds (PO) models (Liu, 2009). This is also true for the continuation ratio (CR) model: they use different formulations, estimate parameters differently, and produce different output results. When estimating the conditional probability of being beyond a category, given that individual has attained that particular category (e,g., $\pi(Y > j \mid Y \ge j$ |), the CR model can be expressed as (Allison, 1999; O'Connell, 2006):

$$\ln\left(\frac{\pi\left(Y > j | x_1, x_2, \ldots x_p\right)}{\pi\left(Y = j | x_1, x_2, \ldots x_p\right)}\right)$$
$$= \alpha_j + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p,$$
$$(7)$$

515

where $\pi(Y > j \,|x_1,x_2, \dots, x_p)$ is the conditional probability of being beyond a category j, conditional on being in that category, given a set of predictors. $j = 1, 2, \dots, J-1$ and where α_j are the cut points and $\beta_1, \beta_2, \beta_p$ are logit coefficients. SAS follows this form in estimating the continuation ratio model with the PROC LOGISTIC command. Before the model is fitted, the data set must be restructured following a series of steps (Allison, 1999; O'Connell, 2006).

First, separate sub-data set must be constructed with the binary outcome variable being beyond a category coded as 1 and 0 otherwise. Individuals who have not advanced to a particular proficiency level are dropped at each stage. If the ordinal dependent variable has j categories, $J-1$ sub-data sets should be created, these data sets are then combined into one data set with a new binary outcome variable with 1 = beyond a particular category. Finally, the CR model is fitted using the SAS PROC LOGISTICS with the descending option.

The CR models also estimates the odds of being in a particular category j relative to being beyond that category. In this situation, the CR model can be formulated as (Ananth & Kleinbaum, 1997; Armstrong & Sloan, 1989; Fienberg, 1980; Long & Freese, 2006):

$$\ln\left(\frac{\pi\left(Y = j|x_1, x_2, \dots x_p\right)}{\pi\left(Y > j|x_1, x_2, \dots x_p\right)}\right)$$
$$= \alpha_j + \left(-\beta_1 X_1 - \beta_2 X_2 - \dots - \beta_p X_p\right)$$
$$(8)$$

where $\pi(Y = j \,|x_1,x_2, \dots, x_p)$ is the conditional probability of being in category j, conditional on being that category or beyond, given a set of predictors, and $j = 1, 2, \dots, J-1$, α_j are the cut points, and $\beta_1, \beta_2 \dots \beta_p$ are logit coefficients. Different from SAS, Stata follows this form to fit the CR model, which is known as the forward CR model (Bender & Bender, 2000). Another distinctive difference is that Stata does not require data restructuring before model fitting; this makes data analysis of the CR model much easier. The following analyses demonstrate how to fit a CR model using Stata; results of model

fitting between Stata and SAS are also compared.

Methodology

Sample

Data were from the Educational Longitudinal Study (ELS, 2002). The ELS:2002 study was conducted by the National Center for Educational Statistics (NCES) and was designed to provide longitudinal data regarding the transitions of high school sophomores in 2002 to postsecondary school education and their future careers. In the 2002 base year of the study, more than 15,000 high school sophomores from a national sample of 752 public and private high schools participated in the study by taking cognitive tests and responding to surveys.

The outcome variable of interest was students' mathematics proficiency levels in high school, which was an ordinal categorical variable with five levels (1 = students can do simple arithmetical operations on whole numbers; 2 = students can do simple operations with decimals, fractions, powers and root; 3 = students can do simple problem solving; 4 = students can understand intermediate-level mathematical concepts and/or find multi-step solutions to word problems; and 5 = students can solve complex multiple-step word problems and/or understand advanced mathematical material) (Ingels, Pratt, Roger, Siegel & Stutts, 2004, 2005). The five proficiency domains were hierarchically structured: mastery of higher proficiency level indicated mastery of all previous levels. Students had to pass through the first four levels of proficiency before achieving the final fifth level; those students who failed to pass through level 1 were assigned to level 0. Table 1 shows the frequency of the six mathematics proficiency levels.

Data Analysis

The continuation ratio model is first fitted with a single explanatory variable using the Stata *ocratio* command (Wolfe, 1998) with the link functions of logit and CLOG-LOG, a proportional odds (PO) model was fitted next, and finally, a full-model with all 11 explanatory variables was fitted. The *eform* option was used to estimate the odds ratios and corresponding standard errors and the confidence intervals. The

ologit command in Stata was used to fit the proportional odds models. The results from both the CR models and the PO models were compared and interpreted. For comparison, the same model was fitted using SAS (V. 9.1.3).

Model fit statistics in the CR model, such as likelihood ratio test and Pseudo R^2, were reported. Other fit statistics, such as Hosmer-Lemeshow GoF test, and Pulkstenis-Robinson (2004) modification, are currently unavailable in the CR model. Following a suggestion by Hilbe (2009), the Stata AIC command was also used to compare model fit.

The log likelihood ratio Chi-Square test with 1 degree of freedom, LR $\chi^2_{(1)} = 38.90$, p < 0.001, indicated that the logit regression coefficient of the predictor, gender was statistically different from 0, therefore, the model with one predictor provides a better fit than the null model with no independent variables in predicting conditional probabilities for mathematics proficiency level. The Pseudo $R^2 = .0008$, which is the likelihood ratio R^2_L, suggested that the relationship between the response variable, mathematics proficiency and the predictor (gender) was small: the AIC statistic was 0.922.

Results

Continuation Ratio Model with a Single Explanatory Variable

A continuation ratio model with a single predictor, gender, was fitted first. The Stata *ocratio* command with the logit function as default was used. Figure 1 displays the Stata output for the single predictor continuation ratio model.

Table 1: Proficiency Categories and Frequencies (Proportions) for the Study Sample, ELS 2002 ($\underline{N} = 15,976$)

Proficiency Category	Description	Frequency
0	Did not pass level 1	842 (5.27%)
1	Can do simple arithmetical operations on whole numbers	3882 (24.30%)
2	Can do simple operations with decimals, fractions, powers, and root	3422 (21.42%)
3	Can do simple problem solving	4521 (28.30%)
4	Can understand intermediate-level mathematical concepts and/or find multi-step solutions to word problems	3196 (20.01%)
5	Can solve complex multiple-step word problems and/or understand advanced mathematical material	113 (0.71%)

Figure 1: Stata Continuation Ratio Model with Logit Link: Single Predictor, Gender

```
. ocratio Profmath BYGENDER, link (logit)

Continuation-ratio logit Estimates                    Number of obs =   51353
                                                      chi2(1)       =   38.90
                                                      Prob > chi2   =  0.0000
Log Likelihood =  -23683.4                            Pseudo R2     =  0.0008


------------------------------------------------------------------------------
     Profmath |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     BYGENDER |   .1416361   .0227235     6.23   0.000     .0970989    .1861732
------------------------------------------------------------------------------
 _cut1   |   -2.790613   .0372137              (Ancillary parameters)
 _cut2   |   -.9961043   .0219305
 _cut3   |   -.7736138   .0238228
 _cut4   |     .368887    .026111
 _cut5   |    3.392331   .0966743
------------------------------------------------------------------------------




. ocratio Profmath BYGENDER, link (logit) eform

Continuation-ratio logit Estimates                    Number of obs =   51353
                                                      chi2(1)       =   38.90
                                                      Prob > chi2   =  0.0000
Log Likelihood =  -23683.4                            Pseudo R2     =  0.0008


------------------------------------------------------------------------------
     Profmath | Odds ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     BYGENDER |   1.152157    .026181     6.23   0.000     1.101969    1.204631
------------------------------------------------------------------------------
 _cut1   |   -2.790613   .0372137              (Ancillary parameters)
 _cut2   |   -.9961043   .0219305
 _cut3   |   -.7736138   .0238228
 _cut4   |     .368887    .026111
 _cut5   |    3.392331   .0966743
------------------------------------------------------------------------------
. aic
AIC Statistic =   .9224153
```

The estimated logit regression coefficient, $\beta = 0.1416$, $z = 6.23$, $p < 0.001$, indicated that gender had a significant effect on mathematics proficiency. Substituting the value of the coefficient into the formula (8), logit $[\pi(Y=j \mid Y \geq j, \text{gender})] = \alpha_j + (-\beta_1 X_1)$, the logit $[\pi(Y=j \mid Y \geq j, \text{gender})] = \alpha_j - 0.1416$ (gender), $OR = e^{(-.1416)} = 0.8680$, was calculated indicating that male students were 0.8680 times the odds for female students of being in any category compared to being in higher categories, that is, female students were more likely than male students to drop out in a particular category, because males are coded as 1 and females are coded as 0.

To estimate the conditional probability of being beyond a category of mathematics proficiency, which is the complement of the conditional probability of being at a category, the signs before the cutpoints and the estimated logits in the equation (8) are changed and the logit $[\pi(Y>j \mid Y \geq j, \text{gender})] = -\alpha_j + 0.1416$ (gender) calculated. Exponentiating 0.1416, results in the $OR = 1.152$, which indicated that male students were 1.152 times more likely to be beyond a particular mathematics proficiency level than female students.

The CR model could also be fitted using the complementary log-log link (clog-log) with the cumulative option within the Stata *ocratio* command. The CR model with the complementary log-log link is actually the discrete-time proportional hazards model for the event history analysis or survival analysis (Allison, 1999; O'Connell, 2006). It estimates the hazard ratio (HR) rather than the odds ratio (OR) of being in a particular category relative to advancing to a higher category. Figure 2 displays the Stata output for the clog-log continuation model.

The log likelihood ratio Chi-Square test with 1 degree of freedom, LR $\chi^2_{(1)} = 51.38$, $p < 0.001$, indicating that the full model with one predictor provides a better fit than the null model with no independent variables. The Pseudo $R^2 = 0.0011$, suggested that the relationship between the response variable, mathematics proficiency, and the predictor, gender was small. The AIC statistic was 0.922

The estimated clog-log coefficient, $\beta = 0.1257$, $z = 7.17$, $p < 0.001$, indicating that

gender had a significant effect on mathematics proficiency. Since Clog-log $[\pi(Y=j \mid Y \geq j, \text{gender})] = \log(-\log(1-\pi)) = \alpha_j + (-\beta_1 X_1)$, we calculated $\log(-\log(1-\pi)) = \alpha_j - 0.1257$ (gender). By exponentiating -0.1257, the hazard ratio, $HR = e^{(-.1257)} = 0.8819$ was obtained, indicating that the hazard of being in a particular proficiency level rather than beyond for male students was 0.8819 times the hazard for female students, that is, the hazard for female students of stopping out in a particular category was 1.134 times as great as that for male students.

Proportional Odds Model with a Single Explanatory Variable

Next, for comparison purposes, a proportional odds model analysis with the same single predictor, gender was conducted using the Stata *ologit* procedure. Figure 3 displays the Stata output for the one-predictor proportional odds model.

LR $\chi^2_{(1)} = 28.13$, $p < 0.001$, indicating that the one-predictor PO model provided a better fit than the null model with no independent variables in predicting cumulative probabilities for mathematics proficiency level. The Pseudo $R^2 = 0.0006$, which was as small as that in the continuation ratio model.

The estimated logit regression coefficient, $\beta = 0.1527$, $z = 5.30$, $p < 0.001$. Because the PO model estimates the cumulative odds and cumulative probabilities of being at or below a particular category of the ordinal response outcome, logit $[\pi(Y \leq j \mid \text{gender})] = \alpha_j - 0.1527$ (gender) was calculated. By exponentiating the logit, -0.1527, the odds ratio (OR), $e^{(-.1527)} = 0.8584$ was obtained, indicating that the odds of being at or below a mathematics proficiency level were 0.8584 times as great for male students as they were for female students, thus, female students were more likely than male students to be at or below a particular proficiency level.

The PO model can estimate J–1 cumulative probabilities of being at or below a category of the ordinal response variable with j levels. When the ordinal response variable, mathematics proficiency, has six levels from 0 to 5, the proportional odds model estimates five cumulative probabilities: $P(Y \leq 0)$, $P(Y \leq 1)$,

Figure 2: Stata Continuation Ratio Model with Clog-log Link: Single Predictor, Gender

```
. ocratio Profmath BYGENDER, link (cloglog) cumulative

Ordered cloglog Estimates                           Number of obs =    51353
                                                    chi2(1)       =    51.38
                                                    Prob > chi2   =   0.0000
Log Likelihood = -23677.16                          Pseudo R2     =   0.0011


-------------------------------------------------------------------------------
    Profmath |      Coef.   Std. Err.       z     P>|z|      [95% Conf. Interval]
-------------+-----------------------------------------------------------------
    BYGENDER |   .1256615   .0175265     7.17    0.000      .0913103     .1600128
-------------------------------------------------------------------------------
 _cut1       |  -2.826367    .0356499              (Ancillary parameters)
 _cut2       |  -.9834265     .022463
 _cut3       |  -.2817271    .0217445
 _cut4       |   .5087509    .0202158
 _cut5       |   1.663668    .0274349
-------------------------------------------------------------------------------

. aic
AIC Statistic =   .9221723




. ocratio Profmath BYGENDER, link (cloglog) eform cumulative

Ordered cloglog Estimates                           Number of obs =    51353
                                                    chi2(1)       =    51.38
                                                    Prob > chi2   =   0.0000
Log Likelihood = -23677.16                          Pseudo R2     =   0.0011


-------------------------------------------------------------------------------
    Profmath | Haz. ratio   Std. Err.       z     P>|z|      [95% Conf. Interval]
-------------+-----------------------------------------------------------------
    BYGENDER |   1.133898   .0198732     7.17    0.000      1.095609     1.173526
-------------------------------------------------------------------------------
 _cut1       |  -2.826367    .0356499              (Ancillary parameters)
 _cut2       |  -.9834265     .022463
 _cut3       |  -.2817271    .0217445
 _cut4       |   .5087509    .0202158
 _cut5       |   1.663668    .0274349
-------------------------------------------------------------------------------
```

P(Y ≤ 2), P(Y ≤ 3) and P(Y ≤ 4). The cumulative probabilities of being beyond a category can also be estimated because they are the complementary probabilities of the being at or below a particular category.

Different from cumulative probabilities in the PO model, the logit CR model estimates conditional probabilities. In the gender-only CR model, it estimates conditional probabilities of being in category j, conditional on being at or beyond that category, that is, P (Y = j | Y ≥ j, gender). This CR model can also estimate the conditional probability of being beyond a category given that individual has achieved that particular category, because P (Y > j | Y ≥ j, gender) is the complementary form of P (Y = j | Y ≥ j, gender).

Another difference between the CR model and the PO model is the change in sample size. In the gender-only PO model, the sample size was 15,325, however, the number of observations increased to 51,353 in the CR model due to different comparisons between proficiency levels, which included level 0 versus levels 1, 2, 3, 4 and 5; level 1 versus levels 2, 3, 4 and 5; level 2 versus 3, 4 and 5; level 3 versus 4 and 5; and level 4 versus level 5 (Table 2 shows the comparisons between the six proficiency levels). Fitting the CR model using SAS required a restructured data set from the J−1concatenated sub-data sets from the comparisons between proficiency levels (Allison, 1999; O'Connell, 2006), though Stata can fit the CR model directly without the data restructuring procedure.

Figure 3: Stata Proportional Odds Model: Single Predictor, Gender

```
ologit Profmath BYGENDER

Iteration 0:   log likelihood = -23702.845
Iteration 1:   log likelihood = -23688.779
Iteration 2:   log likelihood = -23688.778

Ordered logistic regression                    Number of obs   =       15325
                                               LR chi2(1)      =       28.13
                                               Prob > chi2     =      0.0000
Log likelihood = -23688.778                    Pseudo R2       =      0.0006


------------------------------------------------------------------------------
   Profmath |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
------------+-----------------------------------------------------------------
   BYGENDER |   .1527419   .0288057     5.30   0.000     .0962839        .2092
------------+-----------------------------------------------------------------
      /cut1 |  -2.785918   .0381689                     -2.860728    -2.711108
      /cut2 |  -.7893203   .0224898                     -.8333995    -.7452411
      /cut3 |   .1072826   .0214844                       .065174     .1493911
      /cut4 |   1.402499   .0246227                      1.354239     1.450758
      /cut5 |   4.981085    .095611                      4.793691     5.168479
------------------------------------------------------------------------------
```

Continuation Ratio Model with 11 Explanatory Variables

A CR model was fitted with 11 explanatory variables; this was referred to as the full model. Table 3 displays the results for the fitting of the full model with all the predictors.

The log likelihood ratio Chi-Square test, LR $\chi^2_{(11)}$ = 3069.32, p < 0.001, indicating that the full model with 11 predictor provides a better fit than the null model with no independent variables in predicting conditional probability for mathematics proficiency. Although the likelihood ratio R^2_L = 0.0777, was much larger than that of the gender-only model, it was still fairly small, suggesting that the relationship between the response variable, mathematics proficiency and 11 predictors, was small. AIC Goodness-of-fit statistics were used for model comparisons using the AIC command (Hilbe, 2009). Compared with the gender-only model (0.9224), the AIC statistic indicated that the full-model fit the data better (0.8483).

Using the *eform* option, odds ratios could be obtained for all the predictors. Overall, these predictors, such as, being male students (gender), reviewing work from the previous day in math class (review), listening to teachers' lectures (listen), doing problem solving in class (probsolv), using general calculators (usecalcu), using graphing calculators (usegraph), and explaining work orally (explain), were positively associated with the odds of being beyond a particular mathematics proficiency level. Copying notes from board in class (copynote), using books besides textbooks (usebooks), using computers (usecompu), and participating in student-lead discussions (participate) were less likely to advance to a higher proficiency level, that is, they were more likely to stop out in a particular proficiency level.

In terms of odds ratios, male students had 1.359 times greater odds than female students to be beyond a given proficiency level (OR = 1.359), after controlling for the effects of other predictors in the full model. The odds of being beyond a particular proficiency level relative to being in that level were 1.166 times greater with one unit increase in the frequency of reviewing work from the previous day (OR = 1.166). Similarly, listening to teachers' lectures (OR = 1.192), doing problem solving in class

(OR = 1.077), using general calculators (OR = 1.179), using graphing calculators (OR = 1.173), and explaining work orally (OR =1.066) were more likely to be in a higher proficiency level. Conversely, for every one unit increase in copying notes from board in class, the odds of being beyond a particular category decreased by a factor of 0.96 (OR = 0.96). In other words, the more the students copied notes from board, the more likely they would stop out in a mathematics proficiency level. Similarly, the odds decreased by a factor of 0.785 (OR = 0.785), for a unit increase in using textbooks besides the mathematics textbook, they decreased by a factor of 0.833 for a unit increase in using computers in math classes, and they decreased by a factor of 0.892 in participating in student-led discussions, holding the effects of the other variables constant.

Table 3 also provides the results of the multiple regression (MR) analysis. Although the results of MR analysis looked similar to those estimated by the CR model, they were different in nature: the former estimates the linear effects the classroom practices on mathematics proficiency level, while the latter estimates the conditional probability of being in a proficiency level relative to being beyond, or its complement, the probability of advancing to a higher proficiency level rather than being in that particular level. The MR analysis could be used as a preliminary analysis before the CR model fitting.

Comparison of Results of a Single Variable CR Logit Model Using Stata and SAS

When fitting CR models with logit link, Stata and SAS use different procedures to restructure data, estimate parameters differently and produce different outputs. It is, therefore, important to understand how data sets are restructured and how to interpret these estimates. Before using the LOGISTIC procedure, SAS requires a process of data restructuring in order to estimate conditional probabilities of not advancing to a higher proficiency level. If there are j categories, J−1 sub-data sets are needed. Because the mathematics proficiency includes six levels, five sub-data sets are created. Corresponding to the category comparisons indicated in Table 2 (i.e., level 0 versus level 1

Table 2: Category Comparisons for the Continuation Odds Model with Six Mathematics Proficiency Levels (j = 0, 1, 2, …, 5).

Proficiency Category	Conditional Probability $P(Y=j \mid Y \geq j)$	Odds Ratio	Probability Comparisons
0	$P(Y=0 \mid Y \geq 0)$	$\dfrac{P(Y=0)}{P(Y>0)}$	Category 0 vs. all categories above
1	$P(Y=1 \mid Y \geq 1)$	$\dfrac{P(Y=1)}{P(Y>1)}$	Category 1 vs. Categories 2 through 5
2	$P(Y=2 \mid Y \geq 2)$	$\dfrac{P(Y=2)}{P(Y>2)}$	Category 2 vs. Categories 3 - 5
3	$P(Y=3 \mid Y \geq 3)$	$\dfrac{P(Y=3)}{P(Y>3)}$	Category 3 vs. Categories 4 and 5
4	$P(Y=4 \mid Y \geq 4)$	$\dfrac{P(Y=4)}{P(Y>4)}$	Category 4 vs. 5

Table 3: Results of the Continuation Ratio Model and the OLS Regression Model (Full Model), n = 42,992

Variable	Continuation Ratio Model (logit)		OLS Model
	b (se(b))	OR	
α_1	−1.50 (0.08)		1.15 (0.06)
α_2	0.49 (0.08)		
α_3	0.89 (0.08)		
α_4	2.27 (0.08)		
α_5	5.64 (0.13)		
Gender[δ]	0.31 (0.03)[**]	1.36	0.21 (0.02)[**]
Review	0.15 (0.01)[**]	1.17	0.12 (0.01)[**]
Listen	0.18 (0.01)[**]	1.19	0.13 (0.01)[**]
Copynote	−0.04 (0.01)[**]	0.96	−0.02 (0.01)[*]
Usebooks	−0.24 (0.01)[**]	0.79	−0.18 (0.01)[**]
Probsolv	0.07 (0.01)[**]	1.08	0.05 (0.01)[**]
Usecalcu	0.16 (0.01)[**]	1.18	0.12 (0.01)[**]
Usegraph	0.16 (0.01)[**]	1.17	0.11 (0.01)[**]
Usecompu	−0.18 (0.01)[**]	0.83	−0.14 (0.01)[**]
Explain	0.06 (0.01)[**]	1.06	0.05 (0.01)[**]
Participate	−0.11 (0.01)[**]	0.89	−0.09 (0.01)[**]
R^2	$R^2_L = 0.078$		$R^2 = 0.221$
Model Fit[a]	$\chi^2_{11} = 3039.32$ (p < 0.0001)		F(11, 12768) = 329.24[**]

[δ] gender: male=1; [a] Likelihood ratio test; *Significant at p<0.05; ** p<0.01

and above; level 1 versus level 2, and above; level 2 versus 3, 4 and 5; level 3 versus 4 and 5; and level 4 versus level 5), observations for students who did not make to the given proficiency level were dropped out of the concatenated data sets. These sub-data sets were merged into one data set with each individual having as many observations as the number of proficiency levels to which she/her could advance. A new binary variable was created in each data set with being beyond a category coded as 1 and 0 otherwise (see O'Connell, 2006 for details on data restructuring). Different from SAS, the Stata *ocratio* procedure does not require the above process because it restructures the data internally and produces the same sample size as that of the restructured data in SAS.

Table 4 presents a comparison of the results of fitting the single-variable CR model with logit link using both Stata *ocratio* and SAS PROC LOGISTIC with the descending option. In Stata, the CR model estimates the odds of being a particular category versus beyond; while this model in SAS with the descending option estimates the odds of being beyond a given category relative to being in that category, which are the reciprocal. Using Stata and SAS descending, the estimated coefficients are the same in both magnitude and sign. Using the Stata CR model equation (8), logit $[\pi(Y= j \mid Y \geq j, \text{gender})] = \alpha_j + (-\beta_1 X_1)$, logit $[\pi(Y= j \mid Y \geq j, \text{gender})] = \alpha_j - 0.1416$ (gender) was calculated, and OR $= e^{(-.1416)} = 0.8680$, indicating that male students were 0.8680 times the odds for female students of being in any category compared to being in higher categories.

To estimate the conditional probability of being beyond a category of mathematics proficiency using Stata, it is necessary to negate the signs before the cutpoints and the estimated logits in the equation (8) to get the complementary probability of being in a category conditional on being beyond, i.e., logit $[\pi(Y>j \mid Y \geq j, \text{gender})] = -\alpha_j + \beta_1 X_1$. Substituting the coefficient into the equation results in logit $[\pi(Y>j \mid Y \geq j, \text{gender})] = -\alpha_j + 0.1416$ (gender). Exponentiating 0.1416, resulted in the OR of 1.152, which indicated that

male students were 1.152 times more likely to be beyond a particular mathematics proficiency level than female students. Using equation (7) for the SAS CR logit model, it was found that logit $[\pi(Y>j \mid Y \geq j, \text{gender})] = \alpha_j + 0.1416$ (gender). Exponentiating the logit coefficient 0.1416 resulted in the same odds ratio, 1.152.

The CR model using Stata also estimates the cutpoints based on different logit comparisons; these are useful to calculate the conditional probabilities. From the left to the right direction, five cutpoints were -2.791, -0.996, -0.774, 0.369, and 3.392. The results of the CR model using SAS descending as shown in Table 4 provide the estimated intercept, and dumcr0 through dumcr3, which are dummy coded variables for logit comparisons with the final comparison as the reference group. The intercept, -3.392, was the fifth cutpoint, α_5, because it was used to find the odds of being beyond the proficiency level 4 relative to being in that level. The first cutpoint $=$ intercept $+$ dumcr0 $= -3.392 + 6.182 = 2.790$. The second cutpoint $=$ intercept $+$ dumcr1 $= -3.392 + 4.388 = 0.996$. Using the same method resulted in the third, 0.773, and the fourth cutpoints, -0.369, respectively. Comparing the results of the cutpoints estimated by the CR model using Stata and SAS descending, it was found that they were the same in magnitude but had opposite signs. SAS does not provide direct estimates of these cutpoints, but they can be calculated from the estimated intercept and dummy variables.

Although the omnibus likelihood ratio tests for the CR model using Stata and SAS indicated that the single-variable model had better fit than the null model, their degrees of freedom (df) were different because SAS estimated four extra parameters: an intercept and three dummy variables. Accordingly, the log likelihood $R^2_L = 0.254$ estimated using SAS, was much larger than that using Stata, $R^2_L = .0008$. Both CR models had the same sample size when SAS restructured the data (N = 51,353). Feature comparisons of fitting the CR model with the logit link are provided in Table 5.

Table 4: Results of the CR Logit Models with a Single Variable Using Stata and SAS:
A Comparison, $\underline{n}$ = 51,353 (Restructured Data

Model Estimates	STATA	SAS (Descending)
	$P(Y= j \mid Y \geq j)$	$P(Y> j \mid Y \geq j)$
Cutpoints (Stata)/ Intercept (SAS)	$\alpha_1 = -2.791$	Intercept= -3.392
	$\alpha_2 = -0.996$	Dumcr0 = 6.182
	$\alpha_3 = -0.774$	Dumcr1 = 4.388
	$\alpha_4 = 0.369$	Dumcr2 = 4.165
	$\alpha_5 = 3.392$	Dumcr3 = 3.023
BYGENDER$^\delta$	$0.142\ (0.023)^{**}$	$0.142\ (0.023)^{**}$
LR R²	$R^2_L = 0.0008$	$R^2_L = 0.254$
Model Fit[a]	$\chi^2_1 = 38.90\ (p < 0.0001)**$	$\chi^2_5 = 15040.557\ (p < 0.0001)^{**}$

$^\delta$BYGENDER: male=1; [a]Likelihood ratio test; Results are incomparable due to data restructuring using SAS; *Significant at p<0.05; ** p<0.01

Table 5: Feature Comparisons of the CR Model with Logit Link Using Stata and SAS

	STATA	SAS
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		
Log likelihood	√	√
Goodness-of-fit Test	√	√
Pseudo R-Square	√	√
Association of Predicted Probabilities and Observed Responses		√

Conclusion

This article illustrated the use of continuation ratio models to estimate high school students' mathematics proficiency from a set of predictors of classroom practices. Model fitting started from a single-variable CR with both logit and clog-log links and then progressed to a PO model, and finally a full CR logit model with 11 predictor variables.

Results from the CR models suggested that some classroom practices, such as reviewing work from the previous day in math class, listening to teachers' lectures, doing problem solving in class, using general calculators, using graphing calculators and explaining work orally, had positive effects on the odds of being beyond a particular mathematics proficiency level relative to being in that level; while other classroom practices, such as, copying notes from board, using books besides textbooks, using computers in class and participating in student-led discussions were associated with odds of stopping out in a particular proficiency level rather than advancing to a higher proficiency level.

Comparing Stata and SAS, it was found that both packages used different formulations to estimate the CR model and the requirements for data restructuring were also different. Compared to SAS, Stata could estimate the CR model directly without data restructuring. Compared to Stata, SAS produced different model fit statistics, because it estimated more parameters in the CR model, such as dummy coding variables. The estimated logit coefficients were the same using both packages. However, regarding the CR cutpoints, SAS provided different results in the output from those estimated by Stata. Equivalent cutpoints in magnitude could be obtained after further calculations, but they were reversed in sign, because the conditional probabilities estimated by the CR model using Stata and SAS with the descending option were complementary.

In educational research, the demand for ordinal response data analysis is increasing tremendously, it is therefore crucial for researchers to understand different statistical methods for analyzing ordinal response variables. Although comparisons have been made between statistical software packages, a preference of one package over the other is not suggested; this is left to researchers to choose. It is our hope that this article will help researchers become familiar with continuation ratio models and utilize them correctly in their research.

References

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

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

Agresti, A. (2007). *An introduction to categorical data analysis* (*2nd Ed.*). New York: 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. (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.

Fienberg, S. E. (1980). *The analysis of cross-classified categorical data.* Cambridge, MA: The MIT Press.

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

Hardin, J. W., & Hilbe, J. M. (2007). *Generalized linear models and extensions* (*2nd Ed.*). Texas: Stata Press.

Hilbe, J. M. (2009). *Logistic regression models*. Boca Raton, FL: Chapman & Hall/CRC.

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

Ingels, S. J., Pratt, D. J., Roger, J., Siegel, P. H., & Stutts, E. (2004). *ELS: 2002 base year data file user's manual*. Washington, DC: NCES (NCES 2004-405).

Ingels, S. J., Pratt, D. J., Roger, J., Siegel, P. H., & Stutts, E. (2005). *Education Longitudinal Study: 2002/04 public use base-year to first follow-up data files and electronic codebook system*. Washington DC: NCES (NCES 2006-346).

Liu, X. (2009). Ordinal regression analysis: Fitting the proportional odds model using Stata, SAS and SPSS. *Journal of Modern Applied Statistical Methods*, *8*(*2*), 632-645.

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, Series 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: SAGE.

O'Connell, A. A., & Liu, X. (2011). Model diagnostics for proportional and partial proportional odds models. *Journal of Modern Applied Statistical Methods*, *10*(*1*), 139-175.

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

Pulkstenis, E., & Robinson, T. J. (2004). Goodness-of-fit tests for ordinal response regression models. *Statistics in Medicine*, *23*(*6*), 999-1014.

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

Wolfe, R. (1998). Continuation-ratio models for ordinal response data. *Stata Technical Bulletin*, *44*, 18-21.