Example 2b: Predicting Categorical (Ordinal, Nominal) Outcomes via SAS GLIMMIX and LOGISTIC; STATA OLOGIT, GOLOGIT2, and MLOGIT; and R GLM and VGLM
(complete syntax data, and output available for SAS, STATA, and R electronically))

The (fake) data for this example came from: https://stats.idre.ucla.edu/sas/dae/ordinal-logistic-regression/. In this example we will predict a student's **categorical decision** of the likelihood that they will apply to grad school (0=no, 1=eh, or 2=very) using undergraduate GPA (centered at 3.0), whether at least one of their parents has a graduate degree (0=no, 1=yes), and whether they attended a private university (0=no, 1=yes). We will examine three types of models that each use a multinomial conditional response distribution: (1) a standard "proportional odds ordinal regression" (i.e., using a "cumulative logit" link and assuming equal predictor slopes across submodels), (2) a modified ordinal regression for "nonproportional" or "partial-proportional" odds (still with a cumulative logit link, but allowing different predictor slopes across submodels), and (3) a "nominal" or "multinomial" regression (i.e., using a "baseline category" or "generalized logit" link to predict each outcome category in relation to a reference category).

Because STATA LOGIT does not have denominator degrees of freedom, in SAS GLIMMIX they were set to "none" so that the SAS Wald test results (still labeled as t or F) will match those of STATA and R (using z or χ^2). The standard STATA package for ordinal regression, OLOGIT, provides thresholds instead of intercepts and it does not have any means to test or specify non-proportional odds models. To solve these problems, we will be using the custom STATA program GOLOGIT2. In R, we will be using GLM and VGLM (the latter is from the VGAM package).

For syntax for importing and preparing the example data for analysis, please see PSQF 6270 Example 2a.

Syntax and SAS Output for Descriptive Statistics:

```
TITLE1 "SAS Descriptive Statistics";
PROC MEANS NDEC=2 DATA=work.Example2;
     VAR gpa3 parD priv;
RUN;
PROC FREQ DATA=work.Example2;
     TABLE apply3;
RUN;

display "STATA Descriptive Statistics"
summarize gpa3 parD priv
tabulate apply3

print("R Descriptive Statistics")
describe(x=Example2[ , c("gpa3","parD","priv")])
prop.table(table(x=Example2$apply3))
```

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
gpa3	gpa3: Student GPA (0=3)	400	-0.00	0.40	-1.10	1.00
parD	parD: Parent Has Graduate Degree (0=N,1=Y)	400	0.16	0.36	0.00	1.00
priv	priv: Student Attends Private University (0=N,1=Y)	400	0.86	0.35	0.00	1.00

apply3: 0=Not, 1=Eh, 2=Very				
apply3	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	220	55.00	220	55.00
1	140	35.00	360	90.00
2	40	10.00	400	100.00

So now we know that **55% of the respondents have apply3=0, 35% have apply3=1, and 10% have apply3=2**. This information will come in handy in making sure we understand which value our categorical regression models are predicting!

Empty Ordinal Model predicting the cumulative logit of 3-category apply using INTERCEPTS:

$$Logit(Apply3_i > 0) = \beta_{00} \rightarrow Probability(Apply3_i > 0) = \frac{exp(\beta_{00})}{1+exp(\beta_{00})} = \frac{exp(-0.2007)}{[1+exp(-0.2007)]} \,/ = .450$$

$$Logit(Apply3_i > 1) = \beta_{01} \rightarrow Probability(Apply3_i > 1) = \frac{exp(\beta_{01})}{1+exp(\beta_{01})} = \frac{exp(-2.1972)}{[1+exp(-2.1972)]} \,/ = .100$$

SAS Syntax and (condensed) Output:

```
TITLE1 "SAS Empty Model Predicting Ordinal Apply3 using MSPL=ML";
PROC GLIMMIX DATA=work.Example2 NOCLPRINT GRADIENT METHOD=MSPL;
MODEL apply3 (DESCENDING) = / SOLUTION DDFM=NONE LINK=CLOGIT DIST=MULT;
ESTIMATE "Intercept for y>1 (=2)"   int 1 0  / ILINK; * Highest is first;
ESTIMATE "Intercept for y>0 (=1+2)" int 0 1  / ILINK;
RUN; TITLE1;
```

DESCENDING makes the model predict the **probability of the higher category (using intercepts)** instead of the lower category (using thresholds).

```
            Fit Statistics
-2 Log Likelihood            741.21 → -2LL value for model
AIC  (smaller is better)     745.21
BIC  (smaller is better)     753.19
```

Parameter Estimates

Effect	apply3	Estimate	Standard Error	DF	t Value	Pr > |t|	Gradient	
Intercept	2	**-2.1972**	0.1667	Infty	-13.18	<.0001	6.19E-15	→ intercept for y>1
Intercept	1	**-0.2007**	0.1005	Infty	-2.00	0.0459	3.66E-15	→ intercept for y>0

Estimates

Estimate = predicted logit   **Mean = probability**

Label	Estimate	Standard Error	DF	t Value	Pr > |t|	Mean	Standard Error Mean
Intercept for y>1 (=2)	-2.1972	0.1667	Infty	-13.18	<.0001	**0.1000**	0.01500
Intercept for y>0 (=1+2)	-0.2007	0.1005	Infty	-2.00	0.0459	**0.4500**	0.02487

STATA Syntax and (condensed) Output:

```
display "STATA Empty Model Predicting Ordinal Apply"
display "GOLOGIT2 Gives Intercepts (Logit of Higher Category), not Thresholds"
gologit2 apply3
estat ic, n(400)   // AIC and BIC to match SAS
margins            // All 3 probabilities
```

```
Generalized Ordered Logit Estimates                 Number of obs =      400
                                                    LR chi2(0)    =    -0.00
                                                    Prob > chi2   =        .
Log likelihood = -370.60264                         Pseudo R2     =  -0.0000
```

STATA gives LL (so you need to *–2)

```
------------------------------------------------------------------------------
      apply3 | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
0      _cons |  -.2006707   .1005038    -2.00   0.046    -.3976545   -.0036869 → intercept for y>0
1      _cons |  -2.197225   .1666667   -13.18   0.000    -2.523885   -1.870564 → intercept for y>1
------------------------------------------------------------------------------
```

```
------------------------------------------------------------------------------
             |            Delta-method
    _predict |     Margin   std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
           1 |        .55   .0248747    22.11   0.000     .5012465    .5987535
           2 |        .35   .0238485    14.68   0.000     .3032578    .3967422
           3 |         .1       .015     6.67   0.000     .0706005    .1293995
------------------------------------------------------------------------------
```

Margins computes predicted probability of each response (not just for the probability for each submodel).

```
Akaike's information criterion and Bayesian information criterion
-----------------------------------------------------------------------------
       Model |          N   ll(null)  ll(model)      df        AIC        BIC
-------------+---------------------------------------------------------------
           . |        400  -370.6026  -370.6026       2   745.2053   753.1882
-----------------------------------------------------------------------------
```

For comparison, using STATA OLOGIT instead (much more common):

```
display "STATA Empty Model Predicting Ordinal Apply"
display "OLOGIT Gives Thresholds (Logit of Lower Category), not Intercepts"
ologit apply3
estat ic, n(400)   // AIC and BIC to match SAS
margins            // All 3 probabilities
```

STATA gives LL (so you need to *−2)

```
Ordered logistic regression                             Number of obs =     400
Log likelihood = -370.60264                             Pseudo R2     = -0.0000

------------------------------------------------------------------------------
      apply3 | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
       /cut1 |   .2006707   .1005038                      .0036869    .3976545 → threshold for y<1
       /cut2 |   2.197225   .1666667                      1.870564    2.523885 → threshold for y<2
------------------------------------------------------------------------------
```

```
print("R Empty Model Predicting Ordinal Apply3 using VGLM")
Model3Empty = vglm(data=Example2, family=cumulative(link="logitlink",reverse=TRUE,parallel=TRUE),
                   formula=apply3~1)
summary(Model3Empty); AIC(Model3Empty); BIC(Model3Empty) # Get AIC and BIC too
print("Convert logits to probability to check interpretation")
Model3EmptyProb=1/(1+exp(-1*coefficients(Model3Empty))); Model3EmptyProb
```

Reverse=TRUE provides **intercepts** (for y>0 and y>1) instead of thresholds

```
Call:
vglm(formula = apply3 ~ 1, family = cumulative(link = "logitlink",
    reverse = TRUE, parallel = TRUE), data = Example2)

Coefficients:
              Estimate Std. Error  z value Pr(>|z|)
(Intercept):1 -0.20067    0.10050  -1.9966  0.04586 → logit of y>0
(Intercept):2 -2.19722    0.16667 -13.1833  < 2e-16 → logit of y>1

Names of linear predictors: logitlink(P[Y>=2]), logitlink(P[Y>=3])
```

Um, NO, R. These are NOT the names of the linear predictors…

```
Residual deviance: 741.20528 on 798 degrees of freedom → model -2LL
Log-likelihood: -370.60264 on 798 degrees of freedom   → model LL instead (like STATA)

> AIC(Model3Empty)
[1] 745.20528
> BIC(Model3Empty)
[1] 753.18821
```

For comparison, using reverse=FALSE:

```
Call:
vglm(formula = apply3 ~ 1, family = cumulative(link = "logitlink",
    reverse = FALSE, parallel = TRUE), data = Example2)

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept):1  0.20067    0.10050  1.9966  0.04586 → threshold for y<1
(Intercept):2  2.19722    0.16667 13.1833  < 2e-16 → threshold for y<2

Names of linear predictors: logitlink(P[Y<=1]), logitlink(P[Y<=2])
```

Names are correct IF you re-order the 0,1,2 as 1,2,3… (ugh)

<u>Let's add some predictors, starting with main effects only...</u>

$$Logit(Apply3_i > 0) = \beta_{00} + \beta_1(GPA_i - 3) + \beta_2(ParentGD_i) + \beta_3(Private_i)$$

$$Logit(Apply3_i > 1) = \beta_{01} + \beta_1(GPA_i - 3) + \beta_2(ParentGD_i) + \beta_3(Private_i)$$

SAS Syntax and (condensed) Output:

```
TITLE1 "SAS Main-Effects Proportional Odds Model Predicting Ordinal Apply3";
PROC GLIMMIX DATA=work.Example2 NOCLPRINT GRADIENT METHOD=QUAD;
MODEL apply3 (DESCENDING) = gpa3 parD priv
      / SOLUTION DDFM=NONE LINK=CLOGIT DIST=MULT ODDSRATIO(AT gpa3=0 LABEL);
* Fake people will not work to get predicted outcomes;
CONTRAST "Multiv Wald Test of Model R2" gpa3 1, parD 1, priv 1 / CHISQ;
ESTIMATE "y>1 Yhat: Ndeg Pub GPA=2"  int 1 0 gpa3 -1 parD 0 priv 0 / ILINK;
ESTIMATE "y>1 Yhat: Ndeg Pub GPA=3"  int 1 0 gpa3  0 parD 0 priv 0 / ILINK;
ESTIMATE "y>1 Yhat: Ndeg Pub GPA=4"  int 1 0 gpa3  1 parD 0 priv 0 / ILINK;
ESTIMATE "y>1 Yhat: Ndeg Pri GPA=2"  int 1 0 gpa3 -1 parD 0 priv 1 / ILINK;
ESTIMATE "y>1 Yhat: Ndeg Pri GPA=3"  int 1 0 gpa3  0 parD 0 priv 1 / ILINK;
ESTIMATE "y>1 Yhat: Ndeg Pri GPA=4"  int 1 0 gpa3  1 parD 0 priv 1 / ILINK;
ESTIMATE "y>1 Yhat: Ydeg Pub GPA=2"  int 1 0 gpa3 -1 parD 1 priv 0 / ILINK;
ESTIMATE "y>1 Yhat: Ydeg Pub GPA=3"  int 1 0 gpa3  0 parD 1 priv 0 / ILINK;
ESTIMATE "y>1 Yhat: Ydeg Pub GPA=4"  int 1 0 gpa3  1 parD 1 priv 0 / ILINK;
ESTIMATE "y>1 Yhat: Ydeg Pri GPA=2"  int 1 0 gpa3 -1 parD 1 priv 1 / ILINK;
ESTIMATE "y>1 Yhat: Ydeg Pri GPA=3"  int 1 0 gpa3  0 parD 1 priv 1 / ILINK;
ESTIMATE "y>1 Yhat: Ydeg Pri GPA=4"  int 1 0 gpa3  1 parD 1 priv 1 / ILINK;
ESTIMATE "y>0 Yhat: Ndeg Pub GPA=2"  int 0 1 gpa3 -1 parD 0 priv 0 / ILINK;
ESTIMATE "y>0 Yhat: Ndeg Pub GPA=3"  int 0 1 gpa3  0 parD 0 priv 0 / ILINK;
ESTIMATE "y>0 Yhat: Ndeg Pub GPA=4"  int 0 1 gpa3  1 parD 0 priv 0 / ILINK;
ESTIMATE "y>0 Yhat: Ndeg Pri GPA=2"  int 0 1 gpa3 -1 parD 0 priv 1 / ILINK;
ESTIMATE "y>0 Yhat: Ndeg Pri GPA=3"  int 0 1 gpa3  0 parD 0 priv 1 / ILINK;
ESTIMATE "y>0 Yhat: Ndeg Pri GPA=4"  int 0 1 gpa3  1 parD 0 priv 1 / ILINK;
ESTIMATE "y>0 Yhat: Ydeg Pub GPA=2"  int 0 1 gpa3 -1 parD 1 priv 0 / ILINK;
ESTIMATE "y>0 Yhat: Ydeg Pub GPA=3"  int 0 1 gpa3  0 parD 1 priv 0 / ILINK;
ESTIMATE "y>0 Yhat: Ydeg Pub GPA=4"  int 0 1 gpa3  1 parD 1 priv 0 / ILINK;
ESTIMATE "y>0 Yhat: Ydeg Pri GPA=2"  int 0 1 gpa3 -1 parD 1 priv 1 / ILINK;
ESTIMATE "y>0 Yhat: Ydeg Pri GPA=3"  int 0 1 gpa3  0 parD 1 priv 1 / ILINK;
ESTIMATE "y>0 Yhat: Ydeg Pri GPA=4"  int 0 1 gpa3  1 parD 1 priv 1 / ILINK;
ESTIMATE "GPA Slope" gpa3 1 / ILINK; * Example of non-sense ILINK for a slope;
ODS OUTPUT Estimates=work.Pred3PO;   * Save estimates table as dataset to do math;
RUN;
```

Fit Statistics

-2 Log Likelihood	**717.02**	→ **-2LL value for model**
AIC (smaller is better)	727.02	
BIC (smaller is better)	746.98	

Contrasts

Label	Num DF	Den DF	Chi-Square	F Value	Pr > ChiSq	Pr > F
Multiv Wald Test of Model R2	3	Infty	**23.61**	7.87	<.0001	<.0001

Parameter Estimates

Effect	apply3	Estimate	Standard Error	DF	t Value	Pr > |t|	Gradient	
Intercept	2	-2.5102	0.3192	Infty	-7.86	<.0001	7.07E-12	**Beta01**
Intercept	1	-0.4148	0.2830	Infty	-1.47	0.1427	-262E-13	**Beta00**
gpa3		0.6157	0.2606	Infty	2.36	0.0182	5.41E-13	**Beta1**
parD		1.0477	0.2658	Infty	3.94	<.0001	1.92E-12	**Beta2**
priv		0.05868	0.2979	Infty	0.20	0.8438	6.07E-12	**Beta3**

Interpret each fixed effect…

Intercept for 2:

Intercept for 1:

GPA3:

parentGD:

private:

Odds Ratio Estimates

Comparison	Estimate	DF	95% Confidence Limits		
unit change of gpa3 from gpa3=0	1.851	Infty	1.111	3.085	**exp(Beta1)**
unit change of parD from gpa3=0	2.851	Infty	1.693	4.800	**exp(Beta2)**
unit change of priv from gpa3=0	1.060	Infty	0.591	1.901	**exp(Beta3)**

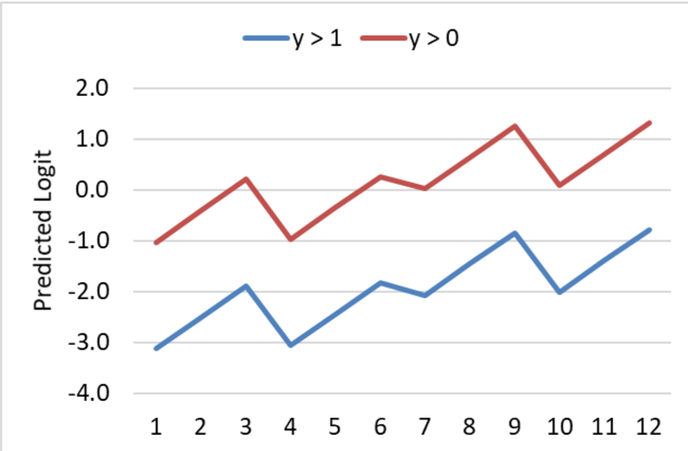
Effects of continuous variables are assessed as units offsets from the reference value.

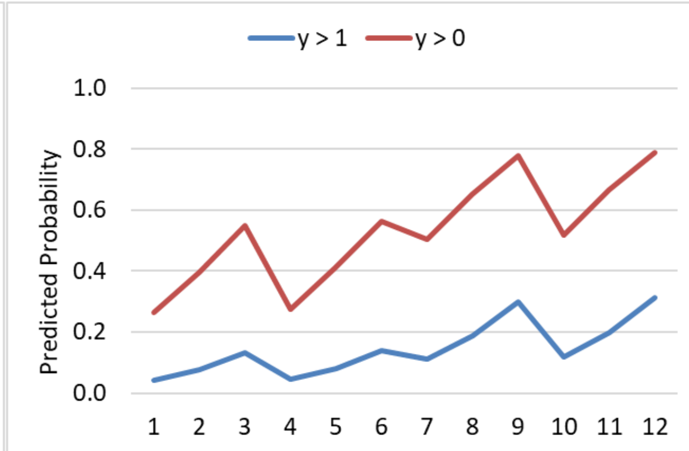
Estimates

Mean = probability saved as "mu" in Estimates

Label	Estimate	Standard Error	DF	t Value	Pr > |t|	**Mean**	Standard Error Mean	
y>1 Yhat: Ndeg Pub GPA=2	-3.1260	0.4525	Infty	-6.91	<.0001	0.04205	0.01823	
y>1 Yhat: Ndeg Pub GPA=3	-2.5102	0.3192	Infty	-7.86	<.0001	0.07515	0.02218	
y>1 Yhat: Ndeg Pub GPA=4	-1.8945	0.3672	Infty	-5.16	<.0001	0.1307	0.04173	
y>1 Yhat: Ndeg Pri GPA=2	-3.0673	0.3211	Infty	-9.55	<.0001	0.04448	0.01365	
y>1 Yhat: Ndeg Pri GPA=3	-2.4515	0.1870	Infty	-13.11	<.0001	0.07933	0.01366	
y>1 Yhat: Ndeg Pri GPA=4	-1.8358	0.3204	Infty	-5.73	<.0001	0.1376	0.03801	
y>1 Yhat: Ydeg Pub GPA=2	-2.0783	0.4971	Infty	-4.18	<.0001	0.1112	0.04914	
y>1 Yhat: Ydeg Pub GPA=3	-1.4625	0.3608	Infty	-4.05	<.0001	0.1881	0.05510	
y>1 Yhat: Ydeg Pub GPA=4	-0.8468	0.3862	Infty	-2.19	0.0283	0.3001	0.08111	
y>1 Yhat: Ydeg Pri GPA=2	-2.0196	0.3893	Infty	-5.19	<.0001	0.1172	0.04027	
y>1 Yhat: Ydeg Pri GPA=3	-1.4039	0.2634	Infty	-5.33	<.0001	0.1972	0.04170	
y>1 Yhat: Ydeg Pri GPA=4	-0.7881	0.3508	Infty	-2.25	0.0247	0.3126	0.07538	
y>0 Yhat: Ndeg Pub GPA=2	-1.0305	0.4206	Infty	-2.45	0.0143	0.2630	0.08153	
y>0 Yhat: Ndeg Pub GPA=3	-0.4148	0.2830	Infty	-1.47	0.1427	0.3978	0.06779	
y>0 Yhat: Ndeg Pub GPA=4	0.2010	0.3451	Infty	0.58	0.5603	0.5501	0.08540	
y>0 Yhat: Ndeg Pri GPA=2	-0.9718	0.2754	Infty	-3.53	0.0004	0.2745	0.05486	
y>0 Yhat: Ndeg Pri GPA=3	-0.3561	0.1173	Infty	-3.04	0.0024	0.4119	0.02840	
y>0 Yhat: Ndeg Pri GPA=4	0.2597	0.2958	Infty	0.88	0.3800	0.5646	0.07271	
y>0 Yhat: Ydeg Pub GPA=2	0.01715	0.4839	Infty	0.04	0.9717	0.5043	0.1210	
y>0 Yhat: Ydeg Pub GPA=3	0.6329	0.3511	Infty	1.80	0.0714	0.6531	0.07954	
y>0 Yhat: Ydeg Pub GPA=4	1.2486	0.3850	Infty	3.24	0.0012	0.7771	0.06670	
y>0 Yhat: Ydeg Pri GPA=2	0.07583	0.3731	Infty	0.20	0.8390	0.5189	0.09315	
y>0 Yhat: Ydeg Pri GPA=3	0.6916	0.2511	Infty	2.75	0.0059	0.6663	0.05583	
y>0 Yhat: Ydeg Pri GPA=4	1.3073	0.3504	Infty	3.73	0.0002	0.7871	0.05872	
GPA Slope	0.6157	0.2606	Infty	2.36	0.0182	**0.6493**	**0.05935**	**Nope!**

The last line illustrates why you cannot "un-logit" a slope into probability… the difference between the predicted outcomes per unit GPA in logits is a constant 0.6157, but the difference in probability is not constant. Similarly, the difference between the parent degree and university type 0–1 groups is constant in logits, but is NOT constant in probability (it depends where you are in probability, which depends upon the submodel and the other predictor values).





To convert the predicted submodel probabilities into predicted probabilities for each response, I had to do the following in SAS—I don't know of an automatic option to do so (as there is in STATA or R):

```
* Convert predicted logits into probabilities;
DATA work.Pred3PO_y2; SET work.Pred3PO; WHERE INDEX(Label,"y>1")>0;
     RENAME Estimate=Logit2 Mu=Prob2; KEEP Label Estimate Mu; RUN;
DATA work.Pred3PO_y1; SET work.Pred3PO; WHERE INDEX(Label,"y>0")>0;
     RENAME Estimate=Logit12 Mu=Prob12; KEEP Label Estimate Mu; RUN;
DATA work.Pred2PO_Prob; MERGE work.Pred3PO_y2 work.Pred3PO_y1;
     Prob0=1-Prob12; Prob1=Prob12-Prob2;
     Label=SUBSTR(Label,11); RUN;
TITLE1 "Predicted Logits and Probabilities for Fake People";
PROC PRINT NOOBS DATA=work.Pred2PO_Prob;
     VAR Label Logit12 Logit2 Prob0 Prob1 Prob2;
RUN; TITLE1;
```

Predicted Logits and Probabilities for Fake People

Label	Logit12	Logit2	Prob0	Prob1	Prob2
Ndeg Pub GPA=2	-1.0305	-3.1260	0.73702	0.22094	0.04205
Ndeg Pub GPA=3	-0.4148	-2.5102	0.60223	0.32262	0.07515
Ndeg Pub GPA=4	0.2010	-1.8945	0.44992	0.41934	0.1307
Ndeg Pri GPA=2	-0.9718	-3.0673	0.72548	0.23004	0.04448
Ndeg Pri GPA=3	-0.3561	-2.4515	0.58809	0.33258	0.07933
Ndeg Pri GPA=4	0.2597	-1.8358	0.43545	0.42700	0.1376
Ydeg Pub GPA=2	0.01715	-2.0783	0.49571	0.39306	0.1112
Ydeg Pub GPA=3	0.6329	-1.4625	0.34685	0.46507	0.1881
Ydeg Pub GPA=4	1.2486	-0.8468	0.22294	0.47696	0.3001
Ydeg Pri GPA=2	0.07583	-2.0196	0.48105	0.40179	0.1172
Ydeg Pri GPA=3	0.6916	-1.4039	0.33368	0.46911	0.1972
Ydeg Pri GPA=4	1.3073	-0.7881	0.21294	0.47449	0.3126

STATA Syntax and (condensed) Output:

```
display "STATA Main-Effects Proportional Odds Model Predicting Ordinal Apply"
// If you get an error with the code below, remove the c. from each predictor
gologit2 apply3 c.gpa3 c.parD c.priv, pl
estat ic, n(400)   // AIC and BIC to match SAS
test (c.gpa3=0) (c.parD=0) (c.priv=0)  // Multiv Wald Test of Model R2
margins, at(c.gpa3=(-1(1)1) c.parD=(0(1)1) c.priv=(0(1)1)) predict(xb) // Yhat>0 in logits
margins, at(c.gpa3=(-1(1)1) c.parD=(0(1)1) c.priv=(0(1)1))     // Each Yhat in probability

display "STATA Main-Effects Proportional Odds Model Predicting Ordinal Apply"
display "Get Odds Ratios Instead of Logit Fixed Effects"
gologit2 apply3 c.gpa3 c.parD c.priv, pl or
```

```
Generalized Ordered Logit Estimates                  Number of obs =    400
                                                     LR chi2(3)    =  24.18 → Is LRT
                                                     Prob > chi2   = 0.0000
Log likelihood = -358.51244 * -2 = -2LL              Pseudo R2     = 0.0326
------------------------------------------------------------------------------
      apply3 | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
0            |
        gpa3 |   .6157458   .2606311     2.36   0.018     .1049183    1.126573  Beta1
        parD |   1.047664   .2657891     3.94   0.000     .5267266    1.568601  Beta2
        priv |   .0586828   .2978589     0.20   0.844    -.5251098    .6424754  Beta3
       _cons |  -.4147686   .2829697    -1.47   0.143     -.969379    .1398418  Beta00
-------------+----------------------------------------------------------------
1            |
        gpa3 |   .6157458   .2606311     2.36   0.018     .1049183    1.126573  Beta1
        parD |   1.047664   .2657891     3.94   0.000     .5267266    1.568601  Beta2
        priv |   .0586828   .2978589     0.20   0.844    -.5251098    .6424754  Beta3
       _cons |  -2.510213   .3191656    -7.86   0.000    -3.135766    -1.88466  Beta01
------------------------------------------------------------------------------

Akaike's information criterion and Bayesian information criterion
-----------------------------------------------------------------------------
       Model |          N   ll(null)  ll(model)      df        AIC        BIC
-------------+---------------------------------------------------------------
           . |        400  -370.6026  -358.5124       5   727.0249   746.9822
-----------------------------------------------------------------------------

. test (c.gpa3=0) (c.parD=0) (c.priv=0)  // Multiv Wald Test of Model R2
           chi2(  3) =   23.61
         Prob > chi2 =    0.0000

------------------------------------------------------------------------------
 Logit y>0   |            Delta-method
             |     Margin   std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
         _at |
          1  |  -1.030514   .4206353    -2.45   0.014    -1.854944   -.2060845
          2  |  -.9718316   .2754386    -3.53   0.000    -1.511681   -.4319819
          3  |   .0171493   .4839078     0.04   0.972    -.9312926    .9655911
          4  |    .075832   .3731219     0.20   0.839    -.6554735    .8071376
          5  |  -.4147686   .2829697    -1.47   0.143     -.969379    .1398418
          6  |  -.3560858   .1172506    -3.04   0.002    -.5858927   -.1262789
          7  |   .6328951   .3510936     1.80   0.071    -.0552358    1.321026
          8  |   .6915779   .2511243     2.75   0.006     .1993834    1.183772
          9  |   .2009772   .3450606     0.58   0.560    -.4753291    .8772836
         10  |     .25966   .2957805     0.88   0.380    -.3200591    .8393791
         11  |   1.248641   .3849985     3.24   0.001     .4940577    2.003224
         12  |   1.307324   .3503769     3.73   0.000     .6205976     1.99405
------------------------------------------------------------------------------
------------------------------------------------------------------------------
 Prob of     |            Delta-method
 0, 1, 2     |     Margin   std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
_predict#_at |
       1  1  |   .7370156    .081529     9.04   0.000     .5772216    .8968096
       1  2  |   .7254844   .0548555    13.23   0.000     .6179697    .8329992
       1  3  |   .4957128   .1209681     4.10   0.000     .2586198    .7328058
       1  4  |   .4810511   .0931465     5.16   0.000     .2984873    .6636149
       1  5  |   .6022307   .0677851     8.88   0.000     .4693744     .735087
       1  6  |   .5880926   .0284027    20.71   0.000     .5324242    .6437609
       1  7  |   .3468544    .079539     4.36   0.000     .1909608     .502748
       1  8  |   .3336822   .0558346     5.98   0.000     .2242484    .4431159
       1  9  |   .4499241   .0853999     5.27   0.000     .2825434    .6173048
       1 10  |   .4354473   .0727126     5.99   0.000     .2929332    .5779613
```

```
      1 11  |   .2229355   .0666953     3.34   0.001     .0922151    .3536559
      1 12  |    .212935    .058721     3.63   0.000     .0978441     .328026
      2  1  |   .2209353   .0649554     3.40   0.001     .0936251    .3482455
      2  2  |   .2300381   .0438069     5.25   0.000     .1441782     .315898
      2  3  |   .3930628    .076493     5.14   0.000     .2431394    .5429862
      2  4  |   .4017898   .0593016     6.78   0.000     .2855609    .5180188
      2  5  |   .3226239   .0502672     6.42   0.000      .224102    .4211459
      2  6  |   .3325807   .0250959    13.25   0.000     .2833935    .3817678
      2  7  |   .4650679   .0396047    11.74   0.000      .387444    .5426917
      2  8  |   .4691145   .0344348    13.62   0.000     .4016236    .5366055
      2  9  |   .4193399   .0518195     8.09   0.000     .3177755    .5209043
      2 10  |   .4270021   .0448068     9.53   0.000     .3391823    .5148218
      2 11  |   .4769606   .0351094    13.58   0.000     .4081475    .5457738
      2 12  |   .4744926   .0359436    13.20   0.000     .4040444    .5449408
      3  1  |   .0420491   .0182265     2.31   0.021     .0063259    .0777723
      3  2  |   .0444775   .0136454     3.26   0.001     .0177329     .071222
      3  3  |   .1112244   .0491389     2.26   0.024      .014914    .2075348
      3  4  |   .1171591   .0402677     2.91   0.004     .0382358    .1960824
      3  5  |   .0751453   .0221815     3.39   0.001     .0316703    .1186203
      3  6  |   .0793267   .0136552     5.81   0.000     .0525631    .1060904
      3  7  |   .1880777   .0550973     3.41   0.001      .080089    .2960665
      3  8  |   .1972033   .0417048     4.73   0.000     .1154634    .2789432
      3  9  |    .130736   .0417322     3.13   0.002     .0489424    .2125296
      3 10  |   .1375506   .0380143     3.62   0.000     .0630439    .2120574
      3 11  |   .3001039   .0811121     3.70   0.000      .141127    .4590807
      3 12  |   .3125724   .0753839     4.15   0.000     .1648227    .4603221
------------------------------------------------------------------------------

Generalized Ordered Logit Estimates (ODDS RATIOS)       Number of obs =     400
                                                        LR chi2(3)    =   24.18
                                                        Prob > chi2   =  0.0000
Log likelihood = -358.51244                             Pseudo R2     =  0.0326
------------------------------------------------------------------------------
      apply3 | Odds ratio   Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
0            |
        gpa3 |   1.851037   .4824377     2.36   0.018      1.11062    3.085067  exp(Beta1)
        parD |   2.850983   .7577602     3.94   0.000      1.69338    4.799927  exp(Beta2)
        priv |   1.060439   .3158611     0.20   0.844     .5914904    1.901181  exp(Beta3)
       _cons |   .6604931   .1868995    -1.47   0.143     .3793185    1.150092  exp(Beta00)
-------------+----------------------------------------------------------------
1            |
        gpa3 |   1.851037   .4824377     2.36   0.018      1.11062    3.085067  exp(Beta1)
        parD |   2.850983   .7577602     3.94   0.000      1.69338    4.799927  exp(Beta2)
        priv |   1.060439   .3158611     0.20   0.844     .5914904    1.901181  exp(Beta3)
       _cons |   .0812509   .0259325    -7.86   0.000     .0434665    .1518807  exp(Beta01)
------------------------------------------------------------------------------
```

R Syntax and (condensed) Output:

```
print("R Main-Effects Proportional Odds Model Predicting Ordinal Apply3")
Model3PO = vglm(data=Example2, family=cumulative(link="logitlink",reverse=TRUE,parallel=TRUE),
                 formula=apply3~1+gpa3+parD+priv)
summary(Model3PO); AIC(Model3PO); BIC(Model3PO) # Get AIC and BIC too
print("Get odds ratios and 95% CIs -- will not match SAS,STATA exactly")
exp(cbind(OR = coefficients(Model3PO), confint(Model3PO)))

print("Multiv Wald Test of Model R2 -- does not match SAS and STATA")
POR2 = glht(model=Model3PO, linfct=c("gpa3=0","parD=0","priv=0"))
summary(POR2, test=Chisqtest()) # Joint chi-square test

print("Get Yhat for specific values of predictors in fake people")
print("Y column = predicted yhat, Yprob = predicted probability")
PredPO = data.frame(FP, Y=predict(object=Model3PO, newdata=FP, type="link"),
                Yprob=predict(object=Model3PO, newdata=FP, type="response"))
```

```
print("Rename columns into something meaningful")
names(PredPO)[names(PredPO)=='Y.logitlink.P.Y..2..']<-'YlogitGE0'
names(PredPO)[names(PredPO)=='Y.logitlink.P.Y..3..']<-'YlogitGE1'
PredPO

Coefficients:
               Estimate Std. Error z value  Pr(>|z|)
(Intercept):1 -0.414757   0.273224 -1.5180   0.12901  Beta00
(Intercept):2 -2.510201   0.310320 -8.0891 6.013e-16  Beta01
gpa3           0.615754   0.262578  2.3450   0.01903  Beta1
parD           1.047655   0.268448  3.9026 9.515e-05  Beta2
priv           0.058672   0.288610  0.2033   0.83891  Beta3

Residual deviance: 717.02487 on 795 degrees of freedom → model -2LL
Log-likelihood: -358.51244 on 795 degrees of freedom   → model   LL

Exponentiated coefficients:
     gpa3      parD      priv
1.8510513 2.8509581 1.0604268 → exp(Beta)

> AIC(Model3PO)
[1] 727.02487

> BIC(Model3PO)
[1] 746.98219

[1] "Get odds ratios and 95% CIs -- will not match SAS,STATA exactly"
                       OR       2.5 %     97.5 %
(Intercept):1 0.660500669 0.386638231 1.12834453  exp(Beta00)
(Intercept):2 0.081251906 0.044227087 0.14927215  exp(Beta01)
gpa3          1.851051322 1.106397840 3.09688873  exp(Beta1)
parD          2.850958137 1.684562634 4.82496889  exp(Beta2)
priv          1.060426849 0.602303377 1.86700780  exp(Beta3)

[1] "Multiv Wald Test of Model R2 -- does not match SAS and STATA"

Linear Hypotheses:
          Estimate
gpa3 == 0 0.615754
parD == 0 1.047655
priv == 0 0.058672

Global Test:
   Chisq DF Pr(>Chisq)
1 24.332  3 2.1295e-05


[1] "Get Yhat for specific values of predictors in fake people"
[1] "Y column = predicted yhat, Yprob = predicted probability"

   gpa3 parD priv    YlogitGE0   YlogitGE1    Yprob.0    Yprob.1     Yprob.2
1    -1    0    0 -1.030510902 -3.12595476 0.73701493 0.22093582 0.042049252
2     0    0    0 -0.414757142 -2.51020100 0.60222800 0.32262586 0.075146139
3     1    0    0  0.200996618 -1.89444724 0.44991934 0.41934244 0.130738229
4    -1    0    1 -0.971839387 -3.06728325 0.72548597 0.23003688 0.044477144
5     0    0    1 -0.356085627 -2.45152949 0.58809255 0.33258068 0.079326772
6     1    0    1  0.259668133 -1.83577573 0.43544529 0.42700305 0.137551656
7    -1    1    0  0.017144225 -2.07829964 0.49571405 0.39306201 0.111223942
8     0    1    0  0.632897985 -1.46254588 0.34685372 0.46506803 0.188078249
9     1    1    0  1.248651744 -0.84679212 0.22293362 0.47696017 0.300106217
10   -1    1    1  0.075815739 -2.01962812 0.48105514 0.40178741 0.117157449
11    0    1    1  0.691569499 -1.40387436 0.33368402 0.46911395 0.197202028
12    1    1    1  1.307323259 -0.78812060 0.21293510 0.47449254 0.312572355
```

These ordinal models rely on an assumption of proportional odds: that all predictor slopes are equal across sub-models. Here is an alternative, a nonproportional (or partial-proportional) odds model, which allows us to test the difference between each predictor slope across submodels:

$$Logit(Apply3_i > 0) = \beta_{00} + \beta_{10}(GPA_i - 3) + \beta_{20}(ParentGD_i) + \beta_{30}(Private_i)$$

$$Logit(Apply3_i > 1) = \beta_{01} + \beta_{11}(GPA_i - 3) + \beta_{21}(ParentGD_i) + \beta_{31}(Private_i)$$

SAS Syntax and (condensed) Output:

UNEQUALSLOPES → nonproportional odds

```
TITLE1 "SAS Main-Effects Ordinal Model Testing Proportional Odds for Apply3";
TITLE2 "Must Use LOGISTIC Instead of GLIMMIX";
PROC LOGISTIC DATA=work.Example2;
MODEL apply3 (DESCENDING) = gpa3 parD priv / LINK=CLOGIT UNEQUALSLOPES;
* Multiv Wald test of model R2;
Model_R2: TEST gpa3_2=0, gpa3_1=0, parD_2=0, parD_1=0, priv_2=0, priv_1=0;
* Univ tests of submodel slope differences and DF=3 multiv test;
gpa3_SlopePO: TEST gpa3_2=gpa3_1;
parD_SlopePO: TEST parD_2=parD_1;
priv_SlopePO: TEST priv_2=priv_1;
All_SlopesPO: TEST gpa3_2=gpa3_1, parD_2=parD_1, priv_2=priv_1;
RUN;
```

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates	
AIC	745.205	729.011	
-2 Log L	741.205	**713.011**	→ -2LL value for model

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq	
Likelihood Ratio	28.1942	6	<.0001	→ As given by STATA
Score	29.7341	6	<.0001	
Wald	**27.7201**	6	0.0001	→ Model R2 test given already

Analysis of Maximum Likelihood Estimates

Parameter	apply3	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq	
Intercept	2	1	-2.0276	0.4050	25.0617	<.0001	Beta01
Intercept	1	1	-0.5685	0.2889	3.8725	0.0491	Beta00
gpa3	2	1	0.7190	0.4537	2.5117	0.1130	Beta11
gpa3	1	1	0.5921	0.2690	4.8431	0.0278	Beta10
parD	2	1	0.9947	0.3741	7.0696	0.0078	Beta21
parD	1	1	1.0831	0.2959	13.3946	0.0003	Beta20
priv	**2**	**1**	**-0.5367**	**0.4293**	**1.5628**	**0.2113**	**Beta31**
priv	**1**	**1**	**0.2307**	**0.3063**	**0.5677**	**0.4512**	**Beta30**

Odds Ratio Estimates

Effect	apply3	Point Estimate	95% Wald Confidence Limits		
gpa3	2	2.052	0.844	4.994	exp(Beta11)
gpa3	1	1.808	1.067	3.063	exp(Beta10)
parD	2	2.704	1.299	5.629	exp(Beta21)
parD	1	2.954	1.654	5.276	exp(Beta20)
priv	**2**	**0.585**	**0.252**	**1.356**	**exp(Beta31)**
priv	**1**	**1.260**	**0.691**	**2.296**	**exp(Beta30)**

```
   Linear Hypotheses Testing Results (from TEST statements)
                     Wald
  Label          Chi-Square     DF    Pr > ChiSq
  Model_R2          27.7201      6        0.0001  → the model R2 is significant (= multiv Wald test)
  Gpa3_SlopePO       0.0839      1        0.7721  → these slopes are not different
  parD_SlopePO       0.0522      1        0.8193  → these slopes are not different
  priv_SlopePO       3.5800      1        0.0585  → these slopes are almost not different
  All_SlopesPO       4.7298      3        0.1927  → Omnibus test says no proportional odds violations
```

STATA Syntax and (condensed) Output:

```
display "STATA Main-Effects Ordinal Model Testing Proportional Odds"
display "Directly provides each slope and differences in slopes across submodels"
gologit2 apply3 c.gpa3 c.parD c.priv, gamma
estat ic, n(400)  // AIC and BIC to match SAS
// Multiv Wald Test of Model R2
test ([0]gpa3=0) ([1]gpa3=0) ([0]parD=0) ([1]parD=0) ([0]priv=0) ([1]priv=0)
// Univ tests of submodel differences between slopes
test ([0]gpa3=[1]gpa3)  // gpa3 slope diff
test ([0]parD=[1]parD)  // parD slope diff
test ([0]priv=[1]priv)  // priv slope diff
// DF=3 multiv test of proportional odds
test ([0]gpa3=[1]gpa3) ([0]parD=[1]parD) ([0]priv=[1]priv)
```

```
Generalized Ordered Logit Estimates                      Number of obs =     400
                                                         LR chi2(6)    =   28.19 → LRT
                                                         Prob > chi2   = 0.0001
Log likelihood = -356.50556 * -2 = -2LL                  Pseudo R2     = 0.0380
------------------------------------------------------------------------------
       apply |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
0            |
        gpa3 |   .5920653   .2690337     2.20   0.028     .0647689    1.119362  Beta10
        parD |   1.083129   .2959475     3.66   0.000     .5030823    1.663175  Beta20
        priv |   .2307488   .3062506     0.75   0.451    -.3694912    .8309889  Beta30
       _cons |  -.5684777   .2888819    -1.97   0.049    -1.134676   -.0022796  Beta00
-------------+----------------------------------------------------------------
1            |
        gpa3 |   .7190314   .4536953     1.58   0.113    -.1701951    1.608258  Beta11
        parD |   .9946781   .3740984     2.66   0.008     .2614588    1.727897  Beta21
        priv |  -.5366997   .4293132    -1.25   0.211    -1.378138    .3047388  Beta31
       _cons |  -2.027556    .405012    -5.01   0.000    -2.821365   -1.233747  Beta01
------------------------------------------------------------------------------

Alternative parameterization: Gammas are deviations from proportionality
------------------------------------------------------------------------------
       apply |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Beta         |
        gpa3 |   .5920653   .2690337     2.20   0.028     .0647689    1.119362  Beta10
        parD |   1.083129   .2959475     3.66   0.000     .5030823    1.663175  Beta20
        priv |   .2307488   .3062506     0.75   0.451    -.3694912    .8309889  Beta30
-------------+----------------------------------------------------------------
Gamma_2      |
        gpa3 |   .1269661   .4383381     0.29   0.772    -.7321607     .986093  Beta11 – Beta10
        parD |  -.0884506   .3871321    -0.23   0.819    -.8472157    .6703144  Beta21 – Beta20
        priv |  -.7674485   .4056115    -1.89   0.058    -1.562432    .0275354  Beta31 – Beta30
-------------+----------------------------------------------------------------
Alpha        |
     _cons_1 |  -.5684777   .2888819    -1.97   0.049    -1.134676   -.0022796  Beta00
     _cons_2 |  -2.027556    .405012    -5.01   0.000    -2.821365   -1.233747  Beta01
------------------------------------------------------------------------------
```

```
Akaike's information criterion and Bayesian information criterion
-----------------------------------------------------------------------------
       Model |          N   ll(null)  ll(model)      df        AIC        BIC
-------------+---------------------------------------------------------------
           . |        400  -370.6026  -356.5056       8   729.0111   760.9428
-----------------------------------------------------------------------------

. // Multiv Wald Test of Model R2
. test ([0]gpa3=0) ([1]gpa3=0) ([0]parD=0) ([1]parD=0) ([0]priv=0) ([1]priv=0)
         chi2(  6) =   27.72
       Prob > chi2 =    0.0001

. // Univ tests of submodel slope differences
. test ([0]gpa3=[1]gpa3)  // gpa3 slope diff
         chi2(  1) =    0.08
       Prob > chi2 =    0.7721

. test ([0]parD=[1]parD)  // parD slope diff
         chi2(  1) =    0.05
       Prob > chi2 =    0.8193

. test ([0]priv=[1]priv)  // priv slope diff
         chi2(  1) =    3.58
       Prob > chi2 =    0.0585

. // DF=3 multiv test of proportional odds
. test ([0]gpa3=[1]gpa3) ([0]parD=[1]parD) ([0]priv=[1]priv)
         chi2(  3) =    4.73
       Prob > chi2 =    0.1927
```

R Syntax and (condensed) Output:

```
print("R Main-Effects Ordinal Model Testing Proportional Odds for Apply3")
Model3NPO = vglm(data=Example2, family=cumulative(link="logitlink",reverse=TRUE,parallel=FALSE),
                formula=apply3~1+gpa3+parD+priv)
summary(Model3NPO); AIC(Model3NPO); BIC(Model3NPO) # Get AIC and BIC too
print("Odds ratios and 95% CIs -- will not match SAS,STATA exactly")
exp(cbind(OR = coefficients(Model3NPO), confint(Model3NPO)))
```

parallel=FALSE → nonproportional odds

```
print("Multiv Wald Test of Model R2 -- does not match SAS and STATA")
NPOR2 = glht(model=Model3NPO,
             linfct=c("gpa3:1=0","gpa3:2=0","parD:1=0","parD:2=0","priv:1=0","priv:2=0"))
summary(NPOR2, test=Chisqtest()) # Joint chi-square test

print("Univ tests of submodel slope differences -- does not match SAS and STATA")
NPOuniv = (summary(glht(model=Model3NPO, linfct=rbind(
  "gpa3 Slope PO" = c(0,0,-1,1, 0,0, 0,0),  # in order of fixed effects
  "parD Slope PO" = c(0,0, 0,0,-1,1, 0,0),
  "priv Slope PO" = c(0,0, 0,0, 0,0,-1,1))),test=adjusted("none")))
NPOuniv
print("Multiv Wald Test of All Slopes PO -- does not match SAS and STATA")
NPOmultiv = glht(model=Model3NPO, linfct=rbind(
  c(0,0,-1,1,0,0,0,0),c(0,0,0,0,-1,1,0,0),c(0,0,0,0,0,0,-1,1)))
summary(NPOmultiv, test=Chisqtest()) # Joint chi-square test
```

```
Coefficients:
              Estimate Std. Error z value  Pr(>|z|)
(Intercept):1 -0.56848    0.28717 -1.9796 0.0477492  Beta00
(Intercept):2 -2.02757    0.39878 -5.0845 3.686e-07  Beta01
gpa3:1         0.59207    0.27247  2.1729 0.0297843  Beta10
gpa3:2         0.71902    0.45280  1.5879 0.1123017  Beta11
parD:1         1.08312    0.29826  3.6314 0.0002819  Beta20
parD:2         0.99470    0.37695  2.6388 0.0083192  Beta21
priv:1         0.23075    0.30485  0.7569 0.4491039  Beta30
priv:2        -0.53669    0.42006 -1.2776 0.2013748  Beta30
```

```
Residual deviance: 713.01111 on 792 degrees of freedom → Model -2LL
Log-likelihood: -356.50556 on 792 degrees of freedom   → Model   LL

Exponentiated coefficients:
   gpa3:1    gpa3:2    parD:1    parD:2    priv:1    priv:2
1.8077234 2.0524197 2.9538950 2.7039030 1.2595402 0.5846818  exp(Beta)

> AIC(Model3NPO)
[1] 729.01111

> BIC(Model3NPO)
[1] 760.94283

[1] "Odds ratios and 95% CIs -- will not match SAS,STATA exactly"
                      OR       2.5 %     97.5 %
(Intercept):1 0.56638835 0.322609206 0.99437883
(Intercept):2 0.13165501 0.060255664 0.28765831
gpa3:1        1.80772336 1.059748942 3.08362066
gpa3:2        2.05241966 0.844967669 4.98531080
parD:1        2.95389504 1.646309532 5.30003364
parD:2        2.70390298 1.291619026 5.66040851
priv:1        1.25954025 0.692979042 2.28930681
priv:2        0.58468180 0.256663288 1.33191156

[1] "Multiv Wald Test of Model R2 -- does not match SAS and STATA"

Linear Hypotheses:
            Estimate
gpa3:1 == 0  0.59207
gpa3:2 == 0  0.71902
parD:1 == 0  1.08312
parD:2 == 0  0.99470
priv:1 == 0  0.23075
priv:2 == 0 -0.53669

Global Test:
  Chisq DF Pr(>Chisq)
1  28.7  6 6.9322e-05

[1] "Univ tests of submodel slope differences -- does not match SAS and STATA"

Linear Hypotheses:
                    Estimate Std. Error z value Pr(>|z|)
gpa3 Slope PO == 0  0.126951   0.440271  0.2883  0.77308
parD Slope PO == 0 -0.088428   0.390153 -0.2267  0.82070
priv Slope PO == 0 -0.767434   0.395425 -1.9408  0.05228
(Adjusted p values reported -- none method)

[1] "Multiv Wald Test of All Slopes PO -- does not match SAS and STATA"

Linear Hypotheses:
        Estimate
1 == 0  0.126951
2 == 0 -0.088428
3 == 0 -0.767434

Global Test:
   Chisq DF Pr(>Chisq)
1 4.4532  3     0.2165
```

Both SAS PROC LOGISTIC and STATA GLOLOGIT2 can automate the selection of which slopes should differ—see the online files for what happens when we let them do it while requesting that all predictors remain in the model even if nonsignificant. Btw, I did not try to figure this out in R…

Here is the final model they came up with—now only the slope for private differs across submodels:

$$Logit(Apply3_i > 0) = \beta_{00} + \beta_1(GPA_i - 3) + \beta_2(ParentGD_i) + \boldsymbol{\beta_{30}}(Private_i)$$

$$Logit(Apply3_i > 1) = \beta_{01} + \beta_1(GPA_i - 3) + \beta_2(ParentGD_i) + \boldsymbol{\beta_{31}}(Private_i)$$

Here is how to specify this same model in which YOU select which slopes are held equal:

SAS Syntax and (condensed) Output:

```
TITLE1 "SAS Main-Effects Ordinal Model -- Custom Non-Proportional Odds";
PROC LOGISTIC DATA=work.Example2;
MODEL apply3 (DESCENDING) = gpa3 parD priv / LINK=CLOGIT UNEQUALSLOPES=priv;
priv_SlopePO: TEST priv_2=priv_1; * Test submodel slope difference;
RUN;
```

UNEQUALSLOPES only for slope of private

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates	
AIC	745.205	725.142	
SC	753.188	749.090	
-2 Log L	741.205	**713.142**	**→ -2LL value for model**

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq	
Likelihood Ratio	28.0637	4	<.0001	→ As given by STATA
Score	29.5992	4	<.0001	
Wald	27.6992	4	<.0001	→ Model R2 test given already

Analysis of Maximum Likelihood Estimates

Parameter	apply3	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq	
Intercept	2	1	-2.0055	0.3707	29.2650	<.0001	Beta01
Intercept	1	1	-0.5691	0.2877	3.9127	0.0479	Beta00
gpa3		1	0.6106	0.2608	5.4821	0.0192	Beta1
parD		1	1.0576	0.2665	15.7449	<.0001	Beta2
priv	**2**	**1**	**-0.5733**	**0.4106**	**1.9490**	**0.1627**	**Beta31**
priv	**1**	**1**	**0.2350**	**0.3053**	**0.5927**	**0.4414**	**Beta30**

Odds Ratio Estimates

Effect	apply3	Point Estimate	95% Wald Confidence Limits		
gpa3		1.842	1.105	3.070	exp(Beta1)
parD		2.880	1.708	4.855	exp(Beta2)
priv	**2**	**0.564**	**0.252**	**1.261**	**exp(Beta31)**
priv	**1**	**1.265**	**0.695**	**2.301**	**exp(Beta30)**

Linear Hypotheses Testing Results

Label	Wald Chi-Square	DF	Pr > ChiSq	
priv_SlopePO	**4.5707**	1	**0.0325**	→ priv slope now differs across submodels

STATA Syntax and (condensed) Output:

```
display "STATA Main-Effects Ordinal Model -- Custom Non-Proportional Odds"
gologit2 apply3 c.gpa3 c.parD c.priv, npl(c.priv) gamma
estat ic, n(400)        // AIC and BIC to match SAS
test ([0]priv=[1]priv)  // priv slope diff
```

npl = nonproportional odds only for private slope

```
display "STATA Main-Effects Ordinal Model -- Custom Non-Proportional Odds"
display "Get Odds Ratios Instead of Logit Fixed Effects"
gologit2 apply3 c.gpa3 c.parD c.priv, npl(c.priv) gamma or
```

```
Generalized Ordered Logit Estimates                        Number of obs =    400
                                                           LR chi2(4)    =  28.06 → LRT
                                                           Prob > chi2   = 0.0000
Log likelihood = -356.57077   * -2 = -2LL                  Pseudo R2     = 0.0379
------------------------------------------------------------------------------
      apply3 | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
0            |
        gpa3 |   .6105983   .2607849     2.34   0.019     .0994694    1.121727  Beta1
        parD |   1.057633   .2665412     3.97   0.000     .5352216    1.580044  Beta2
        priv |   .2350038   .3052548     0.77   0.441    -.3632847    .8332922  Beta30
       _cons |  -.5690629   .2876884    -1.98   0.048    -1.132922    -.005204  Beta00
-------------+----------------------------------------------------------------
1            |
        gpa3 |   .6105983   .2607849     2.34   0.019     .0994694    1.121727  Beta1
        parD |   1.057633   .2665412     3.97   0.000     .5352216    1.580044  Beta2
        priv |  -.5732671   .4106292    -1.40   0.163    -1.378086    .2315513  Beta31
       _cons |  -2.005542     .37073    -5.41   0.000     -2.73216   -1.278925  Beta01
------------------------------------------------------------------------------

Alternative parameterization: Gammas are deviations from proportionality
------------------------------------------------------------------------------
      apply3 | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
Beta         |
        gpa3 |   .6105983   .2607849     2.34   0.019     .0994694    1.121727  Beta1
        parD |   1.057633   .2665412     3.97   0.000     .5352216    1.580044  Beta2
        priv |   .2350038   .3052548     0.77   0.441    -.3632847    .8332922  Beta30
-------------+----------------------------------------------------------------
Gamma_2      |
        priv |  -.8082709   .3780655    -2.14   0.033    -1.549266   -.0672762  Beta31 - Beta30
-------------+----------------------------------------------------------------
Alpha        |
     _cons_1 |  -.5690629   .2876884    -1.98   0.048    -1.132922    -.005204  Beta00
     _cons_2 |  -2.005542     .37073    -5.41   0.000     -2.73216   -1.278925  Beta01
------------------------------------------------------------------------------

Akaike's information criterion and Bayesian information criterion
-----------------------------------------------------------------------------
       Model |          N   ll(null)  ll(model)      df         AIC        BIC
-------------+---------------------------------------------------------------
           . |        400  -370.6026  -356.5708       6    725.1415   749.0903
-----------------------------------------------------------------------------

. test ([0]priv=[1]priv)  // priv slope diff
           chi2(  1) =    4.57
         Prob > chi2 =    0.0325


Generalized Ordered Logit Estimates   (ODDS RATIOS)        Number of obs =    400
                                                           LR chi2(4)    =  28.06
                                                           Prob > chi2   = 0.0000
Log likelihood = -356.57077                                Pseudo R2     = 0.0379
```

```
------------------------------------------------------------------------------
      apply3 | Odds ratio   Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
0            |
        gpa3 |   1.841533    .480244     2.34   0.019     1.104585    3.070153  exp(Beta1)
        parD |   2.879546   .7675177     3.97   0.000     1.707827    4.855169  exp(Beta2)
        priv |   1.264914   .3861209     0.77   0.441     .6953885    2.300881  exp(Beta30)
       _cons |   .5660557   .1628476    -1.98   0.048     .3220908    .9948095  exp(Beta00)
-------------+----------------------------------------------------------------
1            |
        gpa3 |   1.841533    .480244     2.34   0.019     1.104585    3.070153  exp(Beta1)
        parD |   2.879546   .7675177     3.97   0.000     1.707827    4.855169  exp(Beta2)
        priv |   .5636808   .2314638    -1.40   0.163     .2520606    1.260554  exp(Beta31)
       _cons |   .1345873   .0498956    -5.41   0.000     .0650786    .2783364  exp(Beta01)
------------------------------------------------------------------------------
```

R Syntax and (condensed) Output:

```
print("R Main-Effects Ordinal Model -- Custom Non-Proportional Odds for Apply3")
Model3CPO = vglm(data=Example2,
family=cumulative(link="logitlink",reverse=TRUE,parallel=FALSE~priv),
                  formula=apply3~1+gpa3+parD+priv)
summary(Model3CPO); AIC(Model3CPO); BIC(Model3CPO) # Get AIC and BIC too
print("Odds ratios and 95% CIs -- will not match SAS,STATA exactly")
exp(cbind(OR = coefficients(Model3CPO), confint(Model3CPO)))
print("Univ test of submodel slope difference -- does not match SAS and STATA")
CPOuniv = (summary(glht(model=Model3CPO, linfct=rbind(
  "priv Slope PO" = c(0,0,0,0,-1,1))),test=adjusted("none")))
CPOuniv
```

FALSE~priv → np odds only for private slope

```
Coefficients:
              Estimate Std. Error z value  Pr(>|z|)
(Intercept):1 -0.56906    0.28652 -1.9861   0.04702  Beta00
(Intercept):2 -2.00553    0.37084 -5.4081 6.370e-08  Beta01
gpa3           0.61061    0.26289  2.3227   0.02019  Beta1
parD           1.05763    0.26920  3.9288 8.536e-05  Beta2
priv:1         0.23501    0.30433  0.7722   0.43998  Beta30
priv:2        -0.57328    0.40935 -1.4004   0.16138  Beta31

Residual deviance: 713.14154 on 794 degrees of freedom → model -2LL
Log-likelihood: -356.57077 on 794 degrees of freedom   → model   LL

Exponentiated coefficients:
      gpa3       parD     priv:1     priv:2
1.84155529 2.87952956 1.26491688 0.56367392 → exp(Beta)

> AIC(Model3CPO)
[1] 725.14154

> BIC(Model3CPO)
[1] 749.09032

[1] "Odds ratios and 95% CIs -- will not match SAS,STATA exactly"
                      OR       2.5 %     97.5 %
(Intercept):1 0.56605450 0.322828909 0.99253100  exp(Beta00)
(Intercept):2 0.13458872 0.065065401 0.27839869  exp(Beta01)
gpa3          1.84155529 1.100058681 3.08285906  exp(Beta1)
parD          2.87952956 1.698955367 4.88046400  exp(Beta2)
priv:1        1.26491688 0.696656968 2.29670383  exp(Beta30)
priv:2        0.56367392 0.252688216 1.25739258  exp(Beta31)
[1] "Univ test of submodel slope difference -- does not match SAS and STATA"
Linear Hypotheses:
                   Estimate Std. Error z value Pr(>|z|)
priv Slope PO == 0 -0.80829    0.37927 -2.1312  0.03308
```

Btw, here is how I figured out what each of these programs was doing… I wrote it myself!

$$Logit(Apply3_i > 0) = \beta_{00} + \beta_{10}(GPA_i - 3) + \beta_{20}(ParentGD_i) + \beta_{30}(Private_i)$$

$$Logit(Apply3_i > 1) = \beta_{01} + \beta_{11}(GPA_i - 3) + \beta_{21}(ParentGD_i) + \beta_{31}(Private_i)$$

```
* MY TESTS TO MAKE SURE I UNDERSTOOD WHAT EACH PROGRAM DOES;
TITLE1 "SAS Main-Effects Ordinal Model Testing Proportional Odds";
TITLE2 "Using NLMIXED to Estimate a Custom Model";
PROC NLMIXED DATA=work.Example2 METHOD=GAUSS TECH=QUANEW GCONV=1e-12;
* Must list all parms to be estimated here with start values;
* B00 and B01 = intercepts for each equation;
* Bs = fixed effects, now separate per equation;
PARMS B01=-2.5 B11gpa3=0 B21parD=0 B31priv=0
      B00=-0.4 B10gpa3=0 B20parD=0 B30priv=0;
* Linear predictor for y>1 and y>0;
Y1 = B01 + B11gpa3*gpa3 + B21parD*parD + B31priv*priv;
Y0 = B00 + B10gpa3*gpa3 + B20parD*parD + B30priv*priv;
* Model for probability of response - writing it the shorter way;
     IF (apply3=0) THEN P = 1 - (1/(1 + EXP(-Y0)));
ELSE IF (apply3=1) THEN P =     (1/(1 + EXP(-Y0))) - (1/(1 + EXP(-Y1)));
ELSE IF (apply3=2) THEN P =                          (1/(1 + EXP(-Y1)));
* Define log-likelihood using response probabilities;
LL = LOG(P); MODEL apply3 ~ GENERAL(LL);
* Testing proportional odds;
ESTIMATE "GPA3 Slope Diff"  B11gpa3 - B10gpa3;
ESTIMATE "ParD Slope Diff"  B21parD - B20parD;
ESTIMATE "Priv Slope Diff"  B31priv - B30priv;
CONTRAST "All Slopes Diff"  B11gpa3-B10gpa3, B21parD-B20parD, B31priv-B30priv;
RUN;
```

SAS Output (condensed):

Fit Statistics

-2 Log Likelihood	713.0
AIC (smaller is better)	729.0
BIC (smaller is better)	760.9

Parameter Estimates

Parameter	Estimate	Standard Error	DF	t Value	Pr > |t|	95% Confidence Limits		Gradient
B01	-2.0276	0.4050	400	-5.01	<.0001	-2.8238	-1.2313	-2.77E-7
B11gpa3	0.7190	0.4537	400	1.58	0.1138	-0.1729	1.6110	-1.17E-7
B21parD	0.9947	0.3741	400	2.66	0.0082	0.2592	1.7301	-1.98E-7
B31priv	-0.5367	0.4293	400	-1.25	0.2120	-1.3807	0.3073	-1.22E-7
B00	-0.5685	0.2889	400	-1.97	0.0498	-1.1364	-0.00056	-2.62E-7
B10gpa3	0.5921	0.2690	400	2.20	0.0283	0.06317	1.1210	2.973E-8
B20parD	1.0831	0.2959	400	3.66	0.0003	0.5013	1.6649	7.492E-9
B30priv	0.2307	0.3063	400	0.75	0.4516	-0.3713	0.8328	-3.95E-7

Contrasts

Label	Num DF	Den DF	F Value	Pr > F
All Slopes Diff	3	400	1.58	0.1945

Additional Estimates

Label	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper
GPA3 Slope Diff	0.1270	0.4383	400	0.29	0.7722	0.05	-0.7348	0.9887
ParD Slope Diff	-0.08845	0.3871	400	-0.23	0.8194	0.05	-0.8495	0.6726
Priv Slope Diff	-0.7674	0.4056	400	-1.89	0.0592	0.05	-1.5648	0.02995

Let's examine one last set of models—treating our 3-category outcome as "nominal" or "multinomial" instead (i.e., unordered categories in which one category is the reference by which to compare each of the other categories). For convenience we will choose Apply3=1 to be the reference.

Empty Nominal Model predicting the generalized logit of 3-category apply (like dummy-coding):

$$Logit(Apply3_i = 0\ instead\ of\ 1) = \beta_{00} \rightarrow \quad Probability(Apply_i = 0\ instead\ of\ 1) = \frac{exp(\beta_{00})}{1+exp(\beta_{00})}$$

$$Logit(Apply3_i = 2\ instead\ of\ 1) = \beta_{02} \rightarrow \quad Probability(Apply_i = 2\ instead\ of\ 1) = \frac{exp(\beta_{02})}{1+exp(\beta_{02})}$$

SAS Syntax and (condensed) Output:

```
TITLE1 "SAS Empty Model Predicting Nominal Apply3 using MSPL=ML";
PROC GLIMMIX DATA=work.Example2 NOCLPRINT GRADIENT METHOD=MSPL;
MODEL apply3 (REFERENCE="1") = / SOLUTION DDFM=NONE LINK=GLOGIT DIST=MULT;
* ILINK requests logit estimate to be transformed into probability;
  ESTIMATE "Intercept" int 1 / ILINK BYCAT; * BYCAT gives per submodel;
RUN; TITLE1;
```

```
           Fit Statistics
-2 Log Likelihood            741.21  → -2LL value for model
AIC  (smaller is better)     745.21
BIC  (smaller is better)     753.19
```

Parameter Estimates

Effect	apply3	Estimate	Standard Error	DF	t Value	Pr > |t|	Gradient	
Intercept	0	0.4520	0.1081	Infty	4.18	<.0001	-107E-16	intercept for y=0
Intercept	2	-1.2528	0.1793	Infty	-6.99	<.0001	1.35E-14	intercept for y=2

Estimates

Label	apply3	Estimate	Standard Error	DF	t Value	Pr > |t|	Mean	Standard Error Mean	
Intercept	0	0.4520	0.1081	Infty	4.18	<.0001	**0.5500**	0.02487	???
Intercept	2	-1.2528	0.1793	Infty	-6.99	<.0001	**0.1000**	0.01500	???

The logits translate into conditional probabilities, but the predicted "mean" is the marginal probability… like this:

$$\textbf{Given that } y = 0 \textbf{ or } y = 1: \; Probability(Apply_i = 0) = \frac{exp(0.4520)}{[1 + exp(0.4520)]} /= .6111$$

$$\textbf{Given that } y = 2 \textbf{ or } y = 1: \; Probability(Apply_i = 2) = \frac{exp(-1.2528)}{[1 + exp(-1.2528)]} /= .2222$$

apply: 0=Not, 1=Maybe, 2=Yes

APPLY	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0No	220	55.00	220	55.00
1Maybe	140	35.00	360	90.00
2Yes	40	10.00	400	100.00

Probability that y=0 or y=1: .90, so y=0 is .55/.90 = .6111
Probability that y=2 or y=1: .45, so y=2 is .10/.45 = .2222

STATA Syntax and (condensed) Output:

```
display "STATA Empty Model Predicting Nominal Apply"
mlogit apply3, baseoutcome(1)
estat ic, n(400)  // AIC and BIC to match SAS
margins           // All 3 probabilities

Multinomial logistic regression                        Number of obs =    400
                                                       LR chi2(0)    =   0.00
                                                       Prob > chi2   =      .
Log likelihood = -370.60264 * -2 = -2LL                Pseudo R2     = 0.0000
------------------------------------------------------------------------------
      apply3 | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
0      _cons |   .4519851   .1081125     4.18   0.000     .2400885    .6638817
-------------+----------------------------------------------------------------
1            |  (base outcome)
-------------+----------------------------------------------------------------
2      _cons |  -1.252763   .1792843    -6.99   0.000    -1.604154   -.9013722
------------------------------------------------------------------------------
Akaike's information criterion and Bayesian information criterion
-----------------------------------------------------------------------------
       Model |          N   ll(null)  ll(model)      df         AIC        BIC
-------------+---------------------------------------------------------------
           . |        400  -370.6026  -370.6026       2    745.2053   753.1882
-----------------------------------------------------------------------------
------------------------------------------------------------------------------
 Marginal    |            Delta-method
 Probability |     Margin   std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
          1  |        .55   .0248747    22.11   0.000     .5012465    .5987535
          2  |        .35   .0238485    14.68   0.000     .3032578    .3967422
          3  |         .1       .015     6.67   0.000     .0706005    .1293995
------------------------------------------------------------------------------
```

R Syntax and (condensed) Output:

```
print("R Empty Model Predicting Nominal Apply3 -- ref is SECOND category")
Model3NomEmpty = vglm(data=Example2, family=multinomial(refLevel=2),reverse=TRUE,
                      formula=apply3~1)
summary(Model3NomEmpty); AIC(Model3NomEmpty); BIC(Model3NomEmpty) # Get AIC and BIC too
print("Convert logits to probability to check interpretation")
Model3NomEmptyProb=1/(1+exp(-1*coefficients(Model3NomEmpty))); Model3NomEmptyProb

Coefficients:
              Estimate Std. Error z value  Pr(>|z|)
(Intercept):1  0.45199    0.10811  4.1807 2.906e-05  intercept for y=0
(Intercept):2 -1.25276    0.17928 -6.9876 2.797e-12  intercept for y=2

Names of linear predictors: log(mu[,1]/mu[,2]), log(mu[,3]/mu[,2])

Residual deviance: 741.20528 on 798 degrees of freedom → model -2LL
Log-likelihood: -370.60264 on 798 degrees of freedom   → model LL instead (like STATA)

Reference group is level  2  of the response → so y=1 is reference (in refLevel=2)

> AIC(Model3NomEmpty)
[1] 745.20528

> BIC(Model3NomEmpty)
[1] 753.18821

[1] "Convert logits to probability to check interpretation"
(Intercept):1 (Intercept):2
   0.61111111    0.22222222
```

Name is correct IF you re-order the 0,1,2 as 1,2,3… (ugh)

Let's add main effects of our three predictors...

$$Logit(Apply3_i = 0\ instead\ of\ 1) = \beta_{00} + \beta_{10}(GPA_i - 3) + \beta_{20}(ParentGD_i) + \beta_{30}(Private_i)$$
$$Logit(Apply3_i = 2\ instead\ of\ 1) = \beta_{02} + \beta_{12}(GPA_i - 3) + \beta_{22}(ParentGD_i) + \beta_{32}(Private_i)$$

SAS Syntax and (condensed) Output:

```
TITLE1 "SAS Main-Effects Nominal Model";
PROC GLIMMIX DATA=work.Example2 NOCLPRINT GRADIENT METHOD=MSPL;
MODEL apply3 (REFERENCE="1") = gpa3 parD priv
      / SOLUTION DDFM=NONE LINK=GLOGIT DIST=MULT ODDSRATIO(AT gpa3=0 LABEL);
CONTRAST "Multiv Wald Test of Model R2"  gpa3 1, parD 1, priv 1 / CHISQ;
ESTIMATE "Yhat: Ndeg, Pub, GPA=2" int 1 gpa3 -1 parD 0 priv 0 / ILINK BYCAT;
ESTIMATE "Yhat: Ndeg, Pub, GPA=3" int 1 gpa3  0 parD 0 priv 0 / ILINK BYCAT;
ESTIMATE "Yhat: Ndeg, Pub, GPA=4" int 1 gpa3  1 parD 0 priv 0 / ILINK BYCAT;
ESTIMATE "Yhat: Ndeg, Pri, GPA=2" int 1 gpa3 -1 parD 0 priv 1 / ILINK BYCAT;
ESTIMATE "Yhat: Ndeg, Pri, GPA=3" int 1 gpa3  0 parD 0 priv 1 / ILINK BYCAT;
ESTIMATE "Yhat: Ndeg, Pri, GPA=4" int 1 gpa3  1 parD 0 priv 1 / ILINK BYCAT;
ESTIMATE "Yhat: Ydeg, Pub, GPA=2" int 1 gpa3 -1 parD 1 priv 0 / ILINK BYCAT;
ESTIMATE "Yhat: Ydeg, Pub, GPA=3" int 1 gpa3  0 parD 1 priv 0 / ILINK BYCAT;
ESTIMATE "Yhat: Ydeg, Pub, GPA=4" int 1 gpa3  1 parD 1 priv 0 / ILINK BYCAT;
ESTIMATE "Yhat: Ydeg, Pri, GPA=2" int 1 gpa3 -1 parD 1 priv 1 / ILINK BYCAT;
ESTIMATE "Yhat: Ydeg, Pri, GPA=3" int 1 gpa3  0 parD 1 priv 1 / ILINK BYCAT;
ESTIMATE "Yhat: Ydeg, Pri, GPA=4" int 1 gpa3  1 parD 1 priv 1 / ILINK BYCAT;
RUN; TITLE1;
```

Fit Statistics

-2 Log Likelihood	713.99	→ **-2LL value for model**
AIC (smaller is better)	729.99	
BIC (smaller is better)	761.93	

Parameter Estimates

Effect	apply3	Estimate	Standard Error	DF	t Value	Pr > |t|	Gradient	
Intercept	0	0.9515	0.3258	Infty	2.92	0.0035	-2.37E-8	Beta00
Intercept	2	-0.7641	0.4511	Infty	-1.69	0.0903	3.684E-8	Beta02
gpa3	0	-0.4488	0.2902	Infty	-1.55	0.1220	8.02E-11	Beta10
gpa3	2	0.4753	0.4871	Infty	0.98	0.3292	-597E-12	Beta12
parD	**0**	**-0.9516**	**0.3171**	**Infty**	**-3.00**	**0.0027**	**-1.6E-8**	**Beta20**
parD	2	0.4225	0.4083	Infty	1.03	0.3007	2.634E-8	Beta22
priv	0	-0.4188	0.3433	Infty	-1.22	0.2225	-1.27E-8	Beta30
priv	2	-0.7789	0.4706	Infty	-1.66	0.0979	2.266E-8	Beta32

Type III Tests of Fixed Effects → **Joint test of each predictor across submodels**

Effect	Num DF	Den DF	F Value	Pr > F
gpa3	2	Infty	2.48	0.0840
parD	2	Infty	6.93	0.0010
priv	2	Infty	1.52	0.2191

Odds Ratio Estimates

Comparison	Estimate	DF	95% Confidence Limits		
0: unit change of gpa3 from gpa3=0	0.638	Infty	0.361	1.128	exp(Beta10)
2: unit change of gpa3 from gpa3=0	1.608	Infty	0.619	4.179	exp(Beta12)
0: unit change of parD from gpa3=0	0.386	Infty	0.207	0.719	exp(Beta20)
2: unit change of parD from gpa3=0	1.526	Infty	0.685	3.396	exp(Beta22)
0: unit change of priv from gpa3=0	0.658	Infty	0.336	1.289	exp(Beta30)
2: unit change of priv from gpa3=0	0.459	Infty	0.182	1.154	exp(Beta32)

Effects of continuous variables are assessed as units offsets from the reference value.

Contrasts

Label	Num DF	Den DF	Chi-Square	F Value	Pr > ChiSq	Pr > F
Multiv Wald Test of Model R2	6	Infty	**25.84**	4.31	0.0002	0.0002

Estimates

Mean = marginal probability

Label	apply3	Estimate	Standard Error	DF	t Value	Pr > |t|	Mean	Standard Error Mean
Yhat: Ndeg, Pub, GPA=2	0	1.4003	0.4714	Infty	2.97	0.0030	0.7588	0.07966
Yhat: Ndeg, Pub, GPA=2	2	-1.2393	0.7516	Infty	-1.65	0.0991	0.05417	0.03582
Yhat: Ndeg, Pub, GPA=3	0	0.9515	0.3258	Infty	2.92	0.0035	0.6386	0.06687
Yhat: Ndeg, Pub, GPA=3	2	-0.7641	0.4511	Infty	-1.69	0.0903	0.1149	0.04087
Yhat: Ndeg, Pub, GPA=4	0	0.5028	0.3982	Infty	1.26	0.2067	0.4859	0.09194
Yhat: Ndeg, Pub, GPA=4	2	-0.2888	0.5628	Infty	-0.51	0.6079	0.2202	0.08758
Yhat: Ndeg, Pri, GPA=2	0	0.9815	0.3014	Infty	3.26	0.0011	0.7020	0.05998
Yhat: Ndeg, Pri, GPA=2	2	-2.0182	0.5496	Infty	-3.67	0.0002	0.03496	0.01755
Yhat: Ndeg, Pri, GPA=3	0	0.5327	0.1261	Infty	4.23	<.0001	0.5839	0.02887
Yhat: Ndeg, Pri, GPA=3	2	-1.5429	0.2322	Infty	-6.64	<.0001	0.07327	0.01487
Yhat: Ndeg, Pri, GPA=4	0	0.08396	0.3307	Infty	0.25	0.7996	0.4473	0.07667
Yhat: Ndeg, Pri, GPA=4	2	-1.0677	0.5295	Infty	-2.02	0.0438	0.1414	0.06020
Yhat: Ydeg, Pub, GPA=2	0	0.4486	0.5564	Infty	0.81	0.4201	0.5207	0.1294
Yhat: Ydeg, Pub, GPA=2	2	-0.8168	0.8380	Infty	-0.97	0.3297	0.1469	0.09749
Yhat: Ydeg, Pub, GPA=3	0	-0.00012	0.4111	Infty	-0.00	0.9998	0.3689	0.08788
Yhat: Ydeg, Pub, GPA=3	2	-0.3416	0.5184	Infty	-0.66	0.5100	0.2622	0.09139
Yhat: Ydeg, Pub, GPA=4	0	-0.4489	0.4438	Infty	-1.01	0.3118	0.2295	0.07504
Yhat: Ydeg, Pub, GPA=4	2	0.1337	0.5567	Infty	0.24	0.8101	0.4110	0.1242
Yhat: Ydeg, Pri, GPA=2	0	0.02981	0.4335	Infty	0.07	0.9452	0.4614	0.1018
Yhat: Ydeg, Pri, GPA=2	2	-1.5957	0.6744	Infty	-2.37	0.0180	0.09080	0.05246
Yhat: Ydeg, Pri, GPA=3	0	-0.4189	0.2974	Infty	-1.41	0.1589	0.3315	0.06217
Yhat: Ydeg, Pri, GPA=3	2	-1.1204	0.3669	Infty	-3.05	0.0023	0.1644	0.04752
Yhat: Ydeg, Pri, GPA=4	0	-0.8677	0.3967	Infty	-2.19	0.0287	0.2160	0.06392
Yhat: Ydeg, Pri, GPA=4	2	-0.6451	0.5375	Infty	-1.20	0.2301	0.2698	0.1005

STATA Syntax and (condensed) Output:

```
display "STATA Main-Effects Nominal Model"
mlogit apply3 c.gpa3 c.parD c.priv, baseoutcome(1),
estat ic, n(400)  // AIC and BIC to match SAS
test ([0]c.gpa3=0) ([0]c.parD=0) ([0]c.priv=0)  /// Multiv Wald Test of Model R2
     ([2]c.gpa3=0) ([2]c.parD=0) ([2]c.priv=0)
margins, at(c.gpa3=(-1(1)1) c.parD=(0(1)1) c.priv=(0(1)1)) predict(xb) // Yhat logits for 1 vs 0
margins, at(c.gpa3=(-1(1)1) c.parD=(0(1)1) c.priv=(0(1)1))             // All 3 probabilities

display "STATA Main-Effects Nominal Model"
display "Get Odds (Relative Risk) Ratios Instead of Logit Fixed Effects"
mlogit apply3 c.gpa3 c.parD c.priv, baseoutcome(1) rrr
```

There appears to be some controversy in what to call the EXP(logit slope) terms across programs: SAS says they are still "odds ratios" whereas STATA insists they are "relative risk" ratios. The values provided by each are the same, though….

```
Multinomial logistic regression                   Number of obs =    400
                                                  LR chi2(6)    =  27.21
                                                  Prob > chi2   = 0.0001
```

Log likelihood = **-356.99698 * -2 = -2LL** Pseudo R2 = 0.0367

	apply3	Coefficient	Std. err.	z	P>|z|	[95% conf.	interval]	
0	gpa3	-.4487507	.2902058	-1.55	0.122	-1.017544	.1200421	Beta10
	parD	**-.9516468**	**.3170624**	**-3.00**	**0.003**	**-1.573078**	**-.3302159**	**Beta20**
	priv	-.4188184	.3432943	-1.22	0.222	-1.091663	.2540261	Beta30
	_cons	.9515263	.3258247	2.92	0.003	.3129217	1.590131	Beta00
1		(base outcome)						
2	gpa3	.4752888	.4871448	0.98	0.329	-.4794974	1.430075	Beta12
	parD	.4225062	.4082719	1.03	0.301	-.377692	1.222704	Beta22
	priv	-.7788807	.4705994	-1.66	0.098	-1.701239	.1434771	Beta32
	_cons	-.7640601	.451101	-1.69	0.090	-1.648202	.1200817	Beta02

Akaike's information criterion and Bayesian information criterion

Model	N	ll(null)	ll(model)	df	AIC	BIC
.	400	-370.6026	-356.997	8	729.994	761.9257

```
. test ([0]c.gpa3=0) ([0]c.parD=0) ([0]c.priv=0)  /// Multiv Wald Test of Model R2
>      ([2]c.gpa3=0) ([2]c.parD=0) ([2]c.priv=0)
           chi2(  6) =   25.84
         Prob > chi2 =    0.0002
```

Logit y=0	Margin	Delta-method std. err.	z	P>|z|	[95% conf.	interval]
_at						
1	1.400277	.4714007	2.97	0.003	.4763487	2.324205
2	.9814586	.301434	3.26	0.001	.3906589	1.572258
3	.4486302	.5564095	0.81	0.420	-.6419124	1.539173
4	.0298118	.43349	0.07	0.945	-.819813	.8794366
5	.9515263	.3258247	2.92	0.003	.3129217	1.590131
6	.5327079	.1260504	4.23	0.000	.2856537	.779762
7	-.0001205	.4111398	-0.00	1.000	-.8059397	.8056987
8	-.4189389	.2973904	-1.41	0.159	-1.001813	.1639356
9	.5027756	.3981753	1.26	0.207	-.2776337	1.283185
10	.0839571	.3306867	0.25	0.800	-.564177	.7320912
11	-.4488712	.4437556	-1.01	0.312	-1.318616	.4208738
12	-.8676897	.3967459	-2.19	0.029	-1.645297	-.0900819

Prob of 0, 1, 2	Margin	Delta-method std. err.	z	P>|z|	[95% conf.	interval]
_predict#_at						
1 1	.7587733	.0796572	9.53	0.000	.6026481	.9148985
1 2	.7019679	.0599833	11.70	0.000	.5844027	.819533
1 3	.5206685	.1294389	4.02	0.000	.2669729	.774364
1 4	.461375	.1017886	4.53	0.000	.2618729	.660877
1 5	.6385658	.066868	9.55	0.000	.507507	.7696246
1 6	.5839456	.0288715	20.23	0.000	.5273586	.6405326
1 7	.3688851	.0878766	4.20	0.000	.1966501	.5411201
1 8	.331544	.0621686	5.33	0.000	.2096958	.4533923
1 9	.4859103	.0919373	5.29	0.000	.3057165	.6661042
1 10	.4473075	.0766732	5.83	0.000	.2970309	.5975842
1 11	.229503	.0750421	3.06	0.002	.0824231	.3765828
1 12	.2159523	.0639172	3.38	0.001	.0906768	.3412277
2 1	.1870594	.0701952	2.66	0.008	.0494792	.3246395

```
        2  2  |   .2630723   .0575301     4.57   0.000     .1503154    .3758292
        2  3  |   .3324479   .1178515     2.82   0.005     .1014633    .5634326
        2  4  |   .4478235    .102089     4.39   0.000     .2477328    .6479143
        2  5  |   .2465829   .0582712     4.23   0.000     .1323735    .3607923
        2  6  |   .3427838   .0277275    12.36   0.000     .2884389    .3971287
        2  7  |   .3689295   .0891977     4.14   0.000     .1941052    .5437539
        2  8  |   .5040621   .0663465     7.60   0.000     .3740253     .634099
        2  9  |   .2939026   .0794475     3.70   0.000     .1381884    .4496169
        2 10  |   .4112862   .0772504     5.32   0.000     .2598782    .5626941
        2 11  |   .3595263   .1024993     3.51   0.000     .1586312    .5604213
        2 12  |   .5142693   .0985017     5.22   0.000     .3212095     .707329
        3  1  |   .0541673   .0358216     1.51   0.130    -.0160418    .1243764
        3  2  |   .0349598   .0175543     1.99   0.046     .0005541    .0693655
        3  3  |   .1468836   .0974923     1.51   0.132    -.0441978     .337965
        3  4  |   .0908015   .0524554     1.73   0.083    -.0120092    .1936122
        3  5  |   .1148513   .0408725     2.81   0.005     .0347426      .19496
        3  6  |   .0732706   .0148659     4.93   0.000     .0441339    .1024073
        3  7  |   .2621854    .091392     2.87   0.004     .0830604    .4413104
        3  8  |   .1643938   .0475182     3.46   0.001     .0712599    .2575278
        3  9  |    .220187   .0875824     2.51   0.012     .0485286    .3918454
        3 10  |   .1414063   .0602004     2.35   0.019     .0234157    .2593969
        3 11  |   .4109708    .124167     3.31   0.001      .167608    .6543336
        3 12  |   .2697785   .1005486     2.68   0.007     .0727068    .4668502
------------------------------------------------------------------------------

Multinomial logistic regression (ODDS RATIOS)           Number of obs =    400
                                                        LR chi2(6)    =  27.21
                                                        Prob > chi2   = 0.0001
Log likelihood = -356.99698                             Pseudo R2     = 0.0367
------------------------------------------------------------------------------
      apply3 |        RRR   Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
0            |
        gpa3 |   .6384252   .1852747    -1.55   0.122     .3614818    1.127544
        parD |   .3861047   .1224193    -3.00   0.003     .2074059    .7187686
        priv |   .6578236   .2258271    -1.22   0.222     .3356578    1.289205
       _cons |   2.589659   .8437749     2.92   0.003     1.367414    4.904391
-------------+----------------------------------------------------------------
1            |  (base outcome)
-------------+----------------------------------------------------------------
2            |
        gpa3 |   1.608479   .7835619     0.98   0.329     .6190945    4.179012
        parD |   1.525781   .6229334     1.03   0.301     .6854416    3.396361
        priv |   .4589194   .2159672    -1.66   0.098     .1824574     1.15428
       _cons |   .4657715     .21011    -1.69   0.090     .1923955    1.127589
------------------------------------------------------------------------------
```

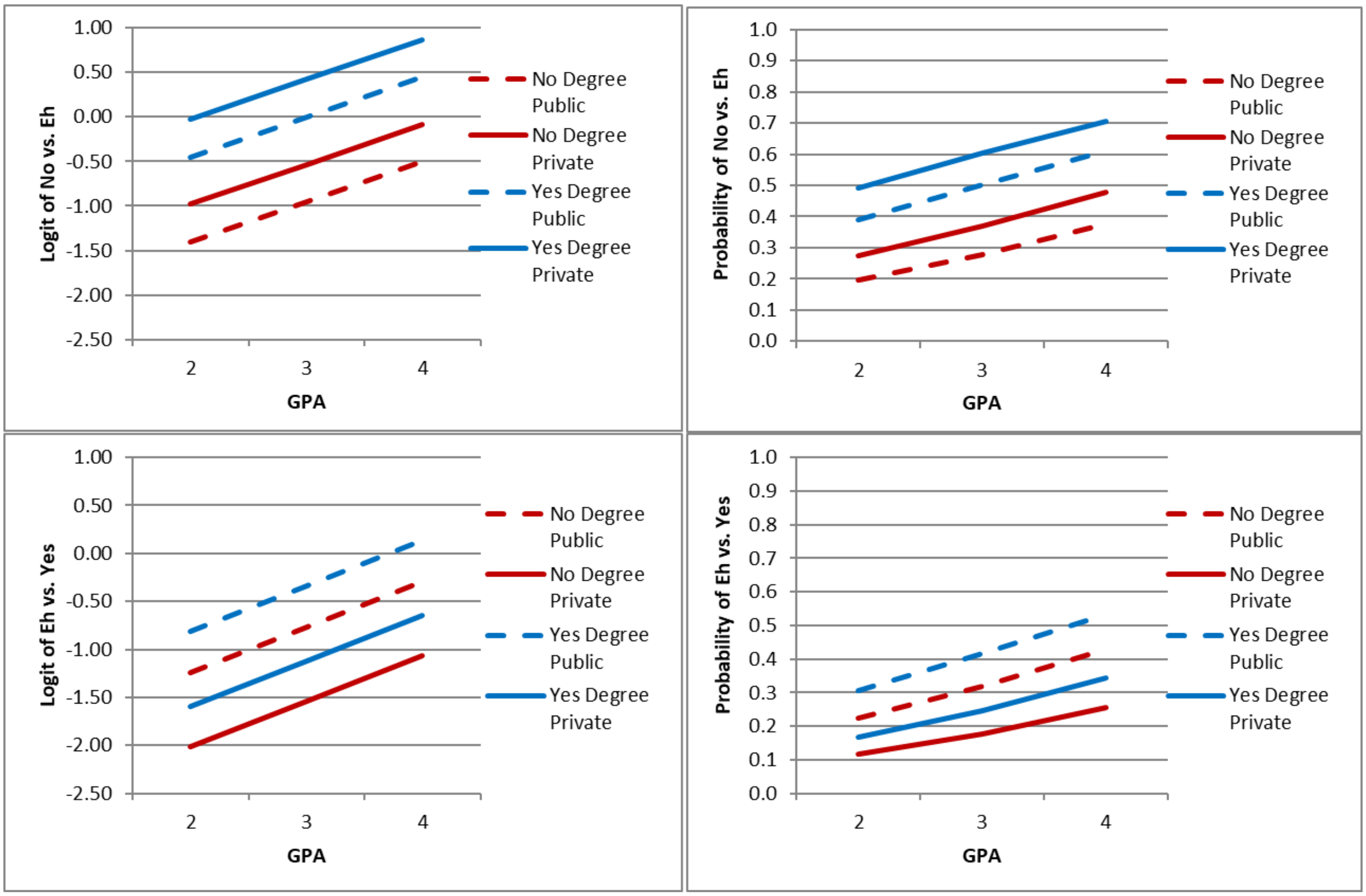
R Syntax and (condensed) Output:

```
print("R Main-Effects Nominal Model -- ref is SECOND category")
Model3NomMain = vglm(data=Example2, family=multinomial(refLevel=2),reverse=TRUE,
                     formula=apply3~1+gpa3+parD+priv)
summary(Model3NomMain); AIC(Model3NomMain); BIC(Model3NomMain) # Get AIC and BIC too
print("Get odds ratios and 95% CIs -- will not match SAS,STATA exactly")
exp(cbind(OR = coefficients(Model3NomMain), confint(Model3NomMain)))

print("Multiv Wald Test of Model R2 -- does match SAS and STATA")
NomR2 = glht(model=Model3NomMain,
             linfct=c("gpa3:1=0","gpa3:2=0","parD:1=0","parD:2=0","priv:1=0","priv:2=0"))
summary(NomR2, test=Chisqtest()) # Joint chi-square test

print("Get Yhat for specific values of predictors in fake people")
print("Y column = predicted yhat, Yprob = predicted probability")
PredNom = data.frame(FP, Y=predict(object=Model3NomMain, newdata=FP, type="link"),
                     Yprob=predict(object=Model3NomMain, newdata=FP, type="response"))
```

```
print("Rename columns into something meaningful")
names(PredNom)[names(PredNom)=='Y.log.mu..1..mu..2..']<-'Ylogit1vs0'
names(PredNom)[names(PredNom)=='Y.log.mu..3..mu..2..']<-'Ylogit1vs2'
PredNom

Coefficients:
                Estimate Std. Error z value Pr(>|z|)
(Intercept):1  0.95153    0.32582  2.9204 0.003496  Beta00
(Intercept):2 -0.76406    0.45110 -1.6938 0.090308  Beta02
gpa3:1        -0.44875    0.29021 -1.5463 0.122028  Beta10
gpa3:2         0.47529    0.48714  0.9757 0.329229  Beta12
parD:1        -0.95165    0.31706 -3.0014 0.002687  Beta20
parD:2         0.42251    0.40827  1.0349 0.300731  Beta22
priv:1        -0.41882    0.34329 -1.2200 0.222466  Beta30
priv:2        -0.77888    0.47060 -1.6551 0.097907  Beta32

Residual deviance: 713.99396 on 792 degrees of freedom  → model -2LL
Log-likelihood: -356.99698 on 792 degrees of freedom    → model   LL

Reference group is level  2  of the response → so y=1 is reference (in refLevel=2)

> AIC(Model3NomMain)
[1] 729.99396
> BIC(Model3NomMain)
[1] 761.92568

[1] "Get odds ratios and 95% CIs -- will not match SAS,STATA exactly"
                      OR      2.5 %     97.5 %
(Intercept):1 2.58965924 1.36741393 4.90439276  exp(Beta00)
(Intercept):2 0.46577148 0.19239614 1.12758539  exp(Beta02)
gpa3:1        0.63842521 0.36148171 1.12754460  exp(Beta10)
gpa3:2        1.60847863 0.61909832 4.17898647  exp(Beta12)
parD:1        0.38610466 0.20740579 0.71876879  exp(Beta20)
parD:2        1.52578072 0.68544314 3.39635289  exp(Beta22)
priv:1        0.65782362 0.33565772 1.28920588  exp(Beta30)
priv:2        0.45891938 0.18245781 1.15427777  exp(Beta32)

[1] "Multiv Wald Test of Model R2 -- does match SAS and STATA"

Linear Hypotheses:
            Estimate
gpa3:1 == 0 -0.44875
gpa3:2 == 0  0.47529
parD:1 == 0 -0.95165
parD:2 == 0  0.42251
priv:1 == 0 -0.41882
priv:2 == 0 -0.77888

Global Test:
   Chisq DF Pr(>Chisq)
1 25.839  6 0.00023855   → Closer!

[1] "Get Yhat for specific values of predictors in fake people"
[1] "Y column = predicted yhat, Yprob = predicted probability"

   gpa3 parD priv     Ylogit1vs0  Ylogit1vs2    Yprob.0    Yprob.1     Yprob.2
1    -1    0    0  1.40027704027 -1.23934893 0.75877334 0.18705937 0.054167285
2     0    0    0  0.95152629782 -0.76406014 0.63856577 0.24658293 0.114851298
3     1    0    0  0.50277555536 -0.28877136 0.48591035 0.29390265 0.220187007
4    -1    0    1  0.98145859580 -2.01822965 0.70196785 0.26307233 0.034959819
5     0    0    1  0.53270785335 -1.54294087 0.58394560 0.34278382 0.073270576
6     1    0    1  0.08395711089 -1.06765208 0.44730754 0.41128618 0.141406282
7    -1    1    0  0.44863024244 -0.81684270 0.52066846 0.33244793 0.146883615
8     0    1    0 -0.00012050002 -0.34155392 0.36888509 0.36892954 0.262185368
9     1    1    0 -0.44887124247  0.13373486 0.22950297 0.35952626 0.410970767
```

```
10   -1    1    1  0.02981179797 -1.59572343 0.46137496 0.44782354 0.090801504
11    0    1    1 -0.41893894449 -1.12043464 0.33154403 0.50406215 0.164393825
12    1    1    1 -0.86768968694 -0.64514586 0.21595225 0.51426928 0.269778473
```

It looks like parent graduate degree (no vs. yes) has a stronger effect on no vs. maybe, whereas public vs. private school has a stronger effect on maybe vs. yes. But are these effect sizes really different??? I don't know, because I haven't figured out a way to test them in these procedures….



Part 2 sample results section using SAS output:

We examined the extent to which a three-category decision for likelihood to apply to graduate school (55.00% 0=No, 35.00% 1=Eh, 10.00% 2=Very) could be predicted by a student's undergraduate GPA (M = 3.0, SD = 0.40, range = 1.90 to 4.00), whether at least one of their parents has a graduate degree (15.75% 0=No, 84.25% 1=Yes), and whether they attended a private university (14.25% 0=No, 85.75% 1=Yes). Specifically, we estimated two alternative sets of generalized linear models with conditional multinomial distributions using maximum likelihood within SAS GLIMMIX. The GPA predictor was centered such that 0 indicated a GPA = 3. Effect sizes are provided using odds ratios (OR), in which OR values between 0 and 1 indicate negative effects, 1 indicates no effect, and values above 1 indicate positive effects. SAS ESTIMATE statements were used to request simple slopes and model-implied predicted outcomes.

First, we treated the three-category outcome as ordinal using a cumulative logit link function—this parameterization requires two submodels that predict the logit of $y_i > 0$ and $y_i > 1$. By default, separate intercepts are estimated for each submodel, but all model slopes are constrained equal across submodels (i.e., proportional odds). This first ordinal model examined the main effects of the three predictors, which together resulted in a significant model, $\chi^2(3) = 23.61$, $p < .0001$. GPA had a significantly positive effect, such that for every unit greater GPA, the logit of the higher-coded response was greater by 0.616 (SE = 0.261; OR = 1.851). Likewise, the logit of the higher-coded response was significantly greater for students for whom at least one parent had a graduate degree by 1.048 (SE = 0.266, OR = 2.851). However, the logit of the higher-coded response was nonsignificantly greater for students who attended a private university by 0.059 (SE = 0.298, OR = 1.060). We then tested the proportional odds assumption by specifying an alternative model in which separate slopes were estimated for the two submodels. Only the slope for parent graduate differed across models—although neither slope was significant, the slope was significantly more negative in predicting y_i > 1 than y_i > 0.

Second, we treated the outcome as nominal using a generalized logit link function—this approach requires choosing a reference category (1=Eh). The submodels then predict the logit of choosing each other possible response (i.e., y_i = 0 given y_i = 0 or 1; y_i = 2 given y_i = 2 or 1). All parameters are estimated separately across submodels, and only two slopes were significant. First, the logit of choosing 0=No instead of 1=Eh was significantly smaller for students for whom for whom at least one parent had a graduate degree by 0.952 (SE = 0.317, OR = 0.386). Second, the logit of choosing 2=Very instead of 1=Eh was significantly smaller for students who went to a private university by 0.779 (SE = 0.471, OR = 0.459).