Example 1 Part 1: Predicting Binary Outcomes via SAS GLIMMIX and STATA MELOGIT
(data, syntax, and output available for SAS and STATA electronically)

The (likely fake) data for this example came from: https://stats.idre.ucla.edu/sas/dae/ordinal-logistic-regression/ . In part 1 of this example we will predict a student's **binary decision** to apply to grad school using undergraduate GPA, whether at least one of their parents has a graduate degree, and whether they attended private school. Although not necessary for these single-level, univariate data, I am using routines that allow expansion into multilevel or multivariate data: SAS GLIMMIX and STATA MELOGIT. SAS GLIMMIX uses denominator degrees of freedom (so its Wald test results are given using t or F), whereas STATA MELOGIT does not (using z or χ^2, respectively).

STATA Syntax and Output for Data Manipulation and Descriptive Statistics:

```
* Defining global variable for file location to be replaced in code below
global filesave "C:\Dropbox\20_PSQF7375_Generalized\PSQF7375_Generalized_Example1"

* Import data, transform variables, apply value formats
use "$filesave\ologit.dta", clear
gen apply2=0
recode apply2 (0=1) if apply==1
recode apply2 (0=1) if apply==2
gen parentgd=pared
gen gpa3=gpa-3
gen private=0
recode private (0=1) if public==0
label variable apply    "apply: 0=Not, 1=Eh, 2=Very"
label variable apply2   "apply2: 0=No, 1=Pry"
label variable parentgd "parentgd: Parent Has Graduate Degree (0=N,1=Y)"
label variable private  "private: Student Attends Private University (0=N,1=Y)"
label variable gpa3     "gpa3: Student GPA (0=3)"
label define f2apply     0 "0No" 1 "1Pry"
label define f3apply     0 "0No" 1 "1Eh" 2 "2Very"
label define fparentgd   0 "0NoDegree" 1 "1YesDegree"
label define fprivate    0 "0public" 1 "1private"
label values apply2 f2apply
label values apply  f3apply
label values parentgd fparentgd
label values private fprivate
```

I recoded the three-category ordinal outcome "apply" into the two-category outcome "apply2" to use for part 1 of this example. The three-category version of the outcome will be predicted in part 2.

```
* Save results to separate file
log using $filesave\Example1_Part1_STATA_Output.log, replace name(Example1_Part1)

display as result "DESCRIPTIVES FOR STUDY VARIABLES"
summarize gpa
tabulate parentgd private
tabulate apply2 // For Part 1
tabulate apply  // For Part 2
```

SAS Syntax and Output for Data Manipulation and Descriptive Statistics:

```
* Location for original files for these models – change this path;
%LET filesave= C:\Dropbox\20_PSQF7375_Generalized\PSQF7375_Generalized_Example1;
LIBNAME filesave "&filesave.";

* Creating formats for categorical variables;
PROC FORMAT;
    VALUE F2apply   0="0No" 1="1Pry";
    VALUE F3apply   0="0No" 1="1Eh" 2="Very";
    VALUE FparentGD 0="0NoDegree" 1="1YesDegree";
    VALUE Fprivate  0="0public" 1="1private";
RUN;
```

```
* Import data into work library and center/recode predictors, apply value formats;
DATA work.ologit; SET filesave.ologit;
     IF apply=0 THEN apply2=0;
     IF apply>0 THEN apply2=1;
     parentGD=pared;
     GPA3=GPA-3;
          IF public=1 THEN private=0;
     ELSE IF public=0 THEN private=1;
     ELSE IF public=. THEN private=.;
     LABEL apply=    "apply: 0=Not, 1=Eh, 2=Very"
           apply2=   "apply2: 0=No, 1=Pry"
           parentGD= "parentGD: Parent Has Graduate Degree (0=N,1=Y)"
           private=  "private: Student Attends Private University (0=N,1=Y)"
           GPA3=     "GPA3: Student GPA (0=3)";
     FORMAT apply2 F2apply. parentGD FparentGD. private Fprivate.;
RUN;

* Open output directory to save results to;
ODS RTF FILE="&filesave.\Example1_Part1_SAS_Output.rtf" STYLE=HTMLBlue STARTPAGE=NO;

TITLE1 "DESCRIPTIVES FOR STUDY VARIABLES";
PROC MEANS NDEC=2 DATA=work.ologit; VAR GPA; RUN;
PROC FREQ DATA=work.ologit;
     TABLE parentGD*private apply2 apply / NOROW NOCOL;
RUN; TITLE1;
```

SAS Output for Descriptive Statistics:

Analysis Variable: GPA

N	Mean	Std Dev	Minimum	Maximum
400	3.00	0.40	1.90	4.00

Frequency Percent	0public	1private	Total
0NoDegree	44 11.00	293 73.25	337 84.25
1YesDegree	13 3.25	50 12.50	63 15.75
Total	57 14.25	343 85.75	400 100.00

apply2: 0=No, 1=Pry

apply2	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0No	220	55.00	220	55.00
1Pry	180	45.00	400	100.00

apply: 0=Not, 1=Eh, 2=Very

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

Empty Model Predicting the logit of the binary version of apply:

$$Logit(Apply_i = 1) = \beta_0 \quad \rightarrow \quad Probability(Apply_i = 1) = \frac{exp(\beta_0)}{1+exp(\beta_0)}$$

```
display as result "STATA EMPTY MODEL PREDICTING BINARY DV"
melogit apply2 ,
estat ic, n(400),
nlcom 1/(1+exp(-1*(_b[_cons]))) // intercept in probability

TITLE1 "SAS EMPTY MODEL PREDICTING BINARY DV";
PROC GLIMMIX DATA=work.ologit NOCLPRINT GRADIENT METHOD=MSPL;
* DESCENDING means predict prob=1 rather than prob=0;
MODEL apply2 (DESCENDING) = / SOLUTION LINK=LOGIT DIST=BINARY;
* ILINK requests logit estimate to be transformed into probability;
ESTIMATE "Intercept" intercept 1 / ILINK;
RUN; TITLE1;
```

SAS Output:

Response Profile

Ordered Value	apply2	Total Frequency
1	1Pry	180
2	0No	220

SAS is trying to help explain what it's doing… see, it's still predicting down, but it re-ordered your data so that up is now down, and down is now up…. That's not confusing at all! This is why we start with an empty model, to make sure we know what SAS is predicting. Fortunately, in the current version of GLIMMIX, it now tells us directly below what it is modeling…

The GLIMMIX procedure is modeling the probability that apply2='1Pry'.

Iteration History

Iteration	Restarts	Evaluations	Objective Function	Change	Max Gradient
0	0	4	275.34490191	.	4.200299
1	0	3	275.25552601	0.08937590	0.010221
2	0	3	**275.25552549**	0.00000053	5.275E-8

Convergence criterion (GCONV=1E-8) satisfied.

Hooray! The overall model converged! But we still need to watch for crazy SEs and gradients ≠ 0 for problems with individual model parameters…

Fit Statistics

-2 Log Likelihood	**550.51**	→ -2LL value for sample = 2*(-1)*275.256, last value above
AIC (smaller is better)	552.51	
AICC (smaller is better)	552.52	
BIC (smaller is better)	556.50	
CAIC (smaller is better)	557.50	
HQIC (smaller is better)	554.09	
Pearson Chi-Square	400.00	
Pearson Chi-Square / DF	1.00	→ Indicates perfect distribution fit (always happens for binary)

$$\text{Probability of } (Apply_i = 1) = \frac{exp(-0.2007)}{[1+exp(-0.2007)]} = 0.450$$

Parameter Estimates

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|	Gradient
Intercept	-0.2007	0.1005	399	-2.00	0.0465	5.275E-8

Estimates

Estimate = predicted logit

Mean = probability

Label	Estimate	Standard Error	DF	t Value	Pr > |t|	Mean	Standard Error Mean
Intercept	-0.2007	0.1005	399	-2.00	0.0465	0.4500	0.02487 → **0.450 matches!**

Let's add some predictors, starting with main effects only...

$$Logit(Apply_i = 1) = \beta_0 + \beta_1(GPA_i - 3) + \beta_2(ParentGD_i) + \beta_3(Private_i)$$

```
display as result "STATA MAIN EFFECTS MODEL PREDICTING BINARY DV"
melogit apply2 c.gpa3 ib(last).parentgd ib(last).private,
estat ic, n(400),
test (c.gpa3=0)(i0.parentgd=i1.parentgd)(i0.private=i1.private)     // demo of Wald test of R2
margins i.parentgd, over(i.private) at(c.gpa3=(-1(1)1)) predict(xb) // logits
margins i.parentgd, over(i.private) at(c.gpa3=(-1(1)1))             // probabilities

* Must re-run with 'or' added to first line to get odds ratios
display as result "STATA MAIN EFFECTS MODEL PREDICTING BINARY DV"
display as result "STATA ODDS RATIOS INSTEAD OF LOGITS"
melogit apply2 c.gpa3 ib(last).parentgd ib(last).private, or

TITLE1 "SAS MAIN EFFECTS MODEL PREDICTING BINARY DV";
PROC GLIMMIX DATA=work.ologit NOCLPRINT NOITPRINT METHOD=MSPL;
CLASS parentGD private;
MODEL apply2 (DESCENDING) = GPA3 parentGD private
                              / SOLUTION LINK=LOGIT DIST=BINARY ODDSRATIO(AT GPA3=0 LABEL);
CONTRAST "Wald test of R2" GPA3 1, parentGD -1 1, private -1 1 / CHISQ;
ESTIMATE "Intercept No Degree,  Public,  GPA=2" int 1 GPA3 -1 parentGD 1 0 private 1 0 / ILINK;
ESTIMATE "Intercept No Degree,  Private, GPA=2" int 1 GPA3 -1 parentGD 1 0 private 0 1 / ILINK;
ESTIMATE "Intercept Yes Degree, Public,  GPA=2" int 1 GPA3 -1 parentGD 0 1 private 1 0 / ILINK;
ESTIMATE "Intercept Yes Degree, Private, GPA=2" int 1 GPA3 -1 parentGD 0 1 private 0 1 / ILINK;
ESTIMATE "Intercept No Degree,  Public,  GPA=3" int 1 GPA3  0 parentGD 1 0 private 1 0 / ILINK;
ESTIMATE "Intercept No Degree,  Private, GPA=3" int 1 GPA3  0 parentGD 1 0 private 0 1 / ILINK;
ESTIMATE "Intercept Yes Degree, Public,  GPA=3" int 1 GPA3  0 parentGD 0 1 private 1 0 / ILINK;
ESTIMATE "Intercept Yes Degree, Private, GPA=3" int 1 GPA3  0 parentGD 0 1 private 0 1 / ILINK;
ESTIMATE "Intercept No Degree,  Public,  GPA=4" int 1 GPA3  1 parentGD 1 0 private 1 0 / ILINK;
ESTIMATE "Intercept No Degree,  Private, GPA=4" int 1 GPA3  1 parentGD 1 0 private 0 1 / ILINK;
ESTIMATE "Intercept Yes Degree, Public,  GPA=4" int 1 GPA3  1 parentGD 0 1 private 1 0 / ILINK;
ESTIMATE "Intercept Yes Degree, Private, GPA=4" int 1 GPA3  1 parentGD 0 1 private 0 1 / ILINK;
ESTIMATE "Slope for GPA" GPA3 1 / ILINK; * Example of non-sense ILINK for a slope;
RUN; TITLE1;
```

SAS Output (condensed and re-arranged for convenience):

Convergence criterion (GCONV=1E-8) satisfied.

Fit Statistics

-2 Log Likelihood	529.92
AIC (smaller is better)	537.92
BIC (smaller is better)	553.89

Contrasts

Label	Num DF	Den DF	Chi-Square	F Value	Pr > ChiSq	Pr > F
Wald test of R2	3	396	18.95	6.32	0.0003	0.0003

Parameter Estimates

Effect	parentGD	private	Estimate	Standard Error	DF	t Value	Pr > |t|	Gradient
Intercept			0.7214	0.2801	396	2.58	0.0104	-868E-14
GPA3			0.5482	0.2724	396	2.01	0.0449	-608E-14
parentGD	0NoDegree		-1.0596	0.2974	396	-3.56	0.0004	3.26E-12
parentGD	1YesDegree		0	.	.	.	.	.
private		0public	-0.2006	0.3053	396	-0.66	0.5117	-143E-14
private		1private	0	.	.	.	.	.

Interpret each fixed effect…

Intercept:

GPA3:

parentGD:

private:

Odds Ratio Estimates

Comparison	Estimate	DF	95% Confidence Limits		
unit change of GPA3 from GPA3=0	1.730	396	1.013	2.956	
parentGD 0NoDegree vs 1YesDegree at GPA3=0	0.347	396	0.193	0.622	1/0.347 to reverse
private 0public vs 1private at GPA3=0	0.818	396	0.449	1.491	1/0.818 to reverse

Effects of continuous variables are assessed as units offsets from the reference value.
The UNIT suboption modifies the offsets.

Because the magnitude of odds ratios less than 1 can be harder to understand, you can flip the direction of the coefficient (i.e., to go from x = 0 to 1 instead of x = 1 to 0) by taking the reciprocal. I did this in the sample results section to match how the categorical predictor variables were originally coded.

Estimates

Label	Estimate	Standard Error	DF	t Value	Pr > |t|	Mean	Standard Error Mean
Intercept for No Degree, Public, GPA=2	-1.0870	0.4312	396	-2.52	0.0121	0.2522	0.08132
Intercept for No Degree, Private, GPA=2	-0.8865	0.2843	396	-3.12	0.0020	0.2918	0.05877
Intercept for Yes Degree, Public, GPA=2	-0.02742	0.5123	396	-0.05	0.9573	0.4931	0.1280
Intercept for Yes Degree, Private, GPA=2	0.1731	0.4078	396	0.42	0.6714	0.5432	0.1012
Intercept for No Degree, Public, GPA=3	-0.5388	0.2874	396	-1.87	0.0616	0.3685	0.06688
Intercept for No Degree, Private, GPA=3	-0.3382	0.1187	396	-2.85	0.0046	0.4162	0.02885
Intercept for Yes Degree, Public, GPA=3	0.5208	0.3714	396	1.40	0.1616	0.6273	0.08683
Intercept for Yes Degree, Private, GPA=3	0.7214	0.2801	396	2.58	0.0104	0.6729	0.06164
Intercept for No Degree, Public, GPA=4	0.009455	0.3574	396	0.03	0.9789	0.5024	0.08934
Intercept for No Degree, Private, GPA=4	0.2100	0.3095	396	0.68	0.4978	0.5523	0.07652
Intercept for Yes Degree, Public, GPA=4	1.0691	0.4024	396	2.66	0.0082	0.7444	0.07656
Intercept for Yes Degree, Private, GPA=4	1.2696	0.3728	396	3.41	0.0007	0.7807	0.06383
Slope for GPA	0.5482	0.2724	396	2.01	0.0449	0.6337	0.06324

The last line illustrates why you cannot "un-logit" a slope all the way back into probability… the difference between the intercepts per unit GPA in logits is a constant 0.5482, but the corresponding difference in probability is not constant between GPA units. Similarly, the difference between the groups is constant in logits, but is NOT constant in probability—it depends where you are on the probability scale.

Let's see how to add interactions, such as all possible two-way interactions...

$$\begin{aligned} Logit(Apply_i = 1) = \beta_0 &+ \beta_1(GPA_i - 3) + \beta_2(ParentGD_i) + \beta_3(Private_i) \\ &+\beta_4(GPA_i - 3)(ParentGD_i) + \beta_5(GPA_i - 3)(Private_i) \\ &+\beta_6(GParentGD_i)(Private_i) \end{aligned}$$

Simple effect of GPA: $\beta_1 + \beta_4(ParentGD_i) + \beta_5(Private_i)$
Simple effect of ParentGD: $\beta_2 + \beta_4(GPA_i - 3) \quad + \beta_6(Private_i)$
Simple effect of Private: $\beta_3 + \beta_5(GPA_i - 3) \quad + \beta_6(ParentGD_i)$

```
display as result "STATA INTERACTIONS EFFECTS MODEL PREDICTING BINARY DV"
melogit apply2 c.gpa3 ib(last).parentgd ib(last).private       ///
               c.gpa3#ib(last).parentgd c.gpa3#ib(last).private ///
               ib(last).parentgd#ib(last).private, // add ,or for odds ratios instead
estat ic, n(400),
test (c.gpa3=0)(i0.parentgd=i1.parentgd)(i0.private=i1.private) ///
     (c.gpa3#i0.parentgd=c.gpa3#i1.parentgd) ///
     (c.gpa3#i0.private=c.gpa3#i1.private)   ///
     (i0.parentgd#i0.private=i1.parentgd#i1.private) // demo of Wald test of R2
test (c.gpa3#i0.parentgd=c.gpa3#i1.parentgd) ///
     (c.gpa3#i0.private=c.gpa3#i1.private)   ///
     (i0.parentgd#i0.private=i1.parentgd#i1.private) // Wald test for change in R2
margins i.parentgd#i.private, at(c.gpa3=(-1(1)1)) predict(xb)// logits
margins i.parentgd#i.private, at(c.gpa3=(-1(1)1))             // probabilities
margins i.private@i.parentgd, at(c.gpa3=(-1(1)1)) predict(xb)// simple effects in logits
margins i.parentgd@i.private, at(c.gpa3=(-1(1)1)) predict(xb)// simple effects in logits
lincom  c.gpa3*1 + c.gpa3#i0.parentgd*1 + c.gpa3#i0.private  // GPA slope: no degree, public
lincom  c.gpa3*1 + c.gpa3#i1.parentgd*1 + c.gpa3#i0.private  // GPA slope: yes degree, public
lincom  c.gpa3*1 + c.gpa3#i0.parentgd*1 + c.gpa3#i1.private  // GPA slope: no degree, private
lincom  c.gpa3*1 + c.gpa3#i1.parentgd*1 + c.gpa3#i1.private  // GPA slope: yes degree, private
// Could not figure out how to get odds ratios for simple effects without retyping them in nlcom

* Close log
log close Example1_Part1
```

```
TITLE1 "SAS INTERACTIONS MODEL PREDICTING BINARY DV";
PROC GLIMMIX DATA=work.ologit NOCLPRINT GRADIENT METHOD=MSPL;
CLASS parentGD private;
MODEL apply2 (DESCENDING) = GPA3|parentGD|private@2 / SOLUTION LINK=LOGIT DIST=BINARY
                                                      ODDSRATIO(AT GPA3=0 DIFF=ALL LABEL);
CONTRAST "Wald test of R2" GPA3 1, parentGD -1 1, private -1 1,
                           GPA3*parentGD -1 1, GPA3*private -1 1,
                           parentGD*private -1 1 1 -1 / CHISQ;
CONTRAST "Wald test of change in R2" GPA3*parentGD -1 1, GPA3*private -1 1,
                                     parentGD*private -1 1 1 -1 / CHISQ;

* Conditional means for plotting interactions, SLICEDIFF gives simple effects;
LSMEANS parentGD*private / ILINK SLICEDIFF=parentGD SLICEDIFF=private ODDSRATIO CL AT(GPA3)=(-1); * At GPA=2;
LSMEANS parentGD*private / ILINK SLICEDIFF=parentGD SLICEDIFF=private ODDSRATIO CL AT(GPA3)=( 0); * At GPA=3;
LSMEANS parentGD*private / ILINK SLICEDIFF=parentGD SLICEDIFF=private ODDSRATIO CL AT(GPA3)=( 1); * At GPA=4;

* Simple slopes for GPA and slope differences by moderators (EXP to get odds ratios);
ESTIMATE "GPA Slope for No  Degree for Public School"  GPA3 1 GPA3*parentGD  1 0 GPA3*private  1 0 / CL EXP;
ESTIMATE "GPA Slope for Yes Degree for Public School"  GPA3 1 GPA3*parentGD  0 1 GPA3*private  1 0 / CL EXP;
ESTIMATE "GPA Slope for No  Degree for Private School" GPA3 1 GPA3*parentGD  1 0 GPA3*private  0 1 / CL EXP;
ESTIMATE "GPA Slope for Yes Degree for Private School" GPA3 1 GPA3*parentGD  0 1 GPA3*private  0 1 / CL EXP;
ODS OUTPUT LSMEANS=work.Meansave SliceDiffs=work.Diffsave; * Save to combine output tables;
RUN; TITLE1;

* Close output;
ODS RTF CLOSE;
```

SAS Output (condensed and re-arranged for convenience):

Convergence criterion (GCONV=1E-8) satisfied.

Fit Statistics

-2 Log Likelihood	526.66
AIC (smaller is better)	540.66
BIC (smaller is better)	568.60

Contrasts

Label	Num DF	Den DF	Chi-Square	F Value	Pr > ChiSq	Pr > F
Wald test of R2	6	393	21.66	3.61	0.0014	0.0017
Wald test of change in R2	3	393	3.18	1.06	0.3642	0.3654

Parameter Estimates

Effect	parentGD	private	Estimate	Standard Error	DF	t Value	Pr > |t|	Gradient
Intercept			0.9521	0.3270	393	2.91	0.0038	-176E-13
GPA3			-0.06789	0.7630	393	-0.09	0.9291	-385E-14
parentGD	0NoDegree		-1.3172	0.3488	393	-3.78	0.0002	4.79E-12
parentGD	1YesDegree		0	.	.	.	.	.
GPA3*parentGD	0NoDegree		0.5872	0.8038	393	0.73	0.4655	-259E-14
GPA3*parentGD	1YesDegree		0	.	.	.	.	.
private		0public	-1.1058	0.7281	393	-1.52	0.1296	3.04E-12
private		1private	0	.	.	.	.	.
GPA3*private		0public	0.8511	0.8153	393	1.04	0.2972	-17E-13
GPA3*private		1private	0	.	.	.	.	.
parentGD*private	0NoDegree	0public	0.9464	0.7703	393	1.23	0.2199	3.09E-12
parentGD*private	0NoDegree	1private	0	.	.	.	.	.
parentGD*private	1YesDegree	0public	0	.	.	.	.	.
parentGD*private	1YesDegree	1private	0	.	.	.	.	.

Interpret each simple effect and interaction…

GPA3:

parentGD:

private:

GPA3*parentGD:

GPA3*private:

parentGD*private:

Odds Ratio Estimates (marginalizing over third variable, so less useful)

Comparison	Estimate	DF	95% Confidence Limits	
parentGD 0NoDegree vs 1YesDegree at GPA3=0	0.430	393	0.191	0.966
unit change of GPA3 from GPA3=0 for parentGD 0NoDegree	2.572	393	1.108	5.973
unit change of GPA3 from GPA3=0 for parentGD 1YesDegree	1.430	393	0.308	6.637
private 0public vs 1private at GPA3=0	0.531	393	0.228	1.237
unit change of GPA3 from GPA3=0 for private 0public	2.935	393	0.621	13.878
unit change of GPA3 from GPA3=0 for private 1private	1.253	393	0.547	2.874

Effects of continuous variables are assessed as units offsets from the reference value.
The UNIT suboption modifies the offsets.

Estimates → Simple Effects of GPA

Label	Estimate	Standard Error	DF	t Value	Pr > |t|	Lower	Upper
GPA Slope for No Degree for Public School	1.3704	0.7761	393	1.77	0.0782	-0.1554	2.8963
GPA Slope for Yes Degree for Public School	0.7832	0.9846	393	0.80	0.4268	-1.1525	2.7189
GPA Slope for No Degree for Private School	0.5193	0.3119	393	1.67	0.0967	-0.09381	1.1325
GPA Slope for Yes Degree for Private School	-0.06789	0.7630	393	-0.09	0.9291	-1.5679	1.4322

Label	Exponentiated Estimate	Exponentiated Lower	Exponentiated Upper
GPA Slope for No Degree for Public School	3.9370	0.8561	18.1064
GPA Slope for Yes Degree for Public School	2.1885	0.3159	15.1637
GPA Slope for No Degree for Private School	1.6809	0.9105	3.1033
GPA Slope for Yes Degree for Private School	0.9344	0.2085	4.1877

`LSMEANS output (cleaned up from ODS OUTPUT):`

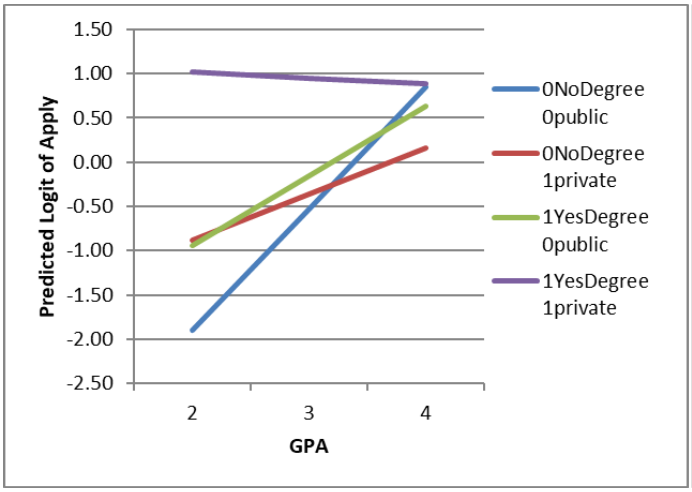
parentGD	private	GPA3	Estimate	Standard Error	DF	t Value	Pr > |t|	Lower	Upper	Mean	Standard Error	Lower Mean	Upper Mean
0NoDegree	0public	-1	-1.895	0.979	393	-1.940	0.054	-3.820	0.030	0.131	0.111	0.021	0.508
0NoDegree	1private	-1	-0.884	0.320	393	-2.760	0.006	-1.513	-0.255	0.292	0.066	0.180	0.437
1YesDegree	0public	-1	-0.937	1.480	393	-0.630	0.527	-3.846	1.972	0.282	0.299	0.021	0.878
1YesDegree	1private	-1	1.020	0.907	393	1.120	0.262	-0.763	2.803	0.735	0.177	0.318	0.943
0NoDegree	0public	0	-0.525	0.349	393	-1.500	0.134	-1.210	0.161	0.372	0.081	0.230	0.540
0NoDegree	1private	0	-0.365	0.121	393	-3.030	0.003	-0.602	-0.128	0.410	0.029	0.354	0.468
1YesDegree	0public	0	-0.154	0.680	393	-0.230	0.821	-1.491	1.184	0.462	0.169	0.184	0.766
1YesDegree	1private	0	0.952	0.327	393	2.910	0.004	0.309	1.595	0.722	0.066	0.577	0.831
0NoDegree	0public	1	0.846	0.699	393	1.210	0.227	-0.529	2.221	0.700	0.147	0.371	0.902
0NoDegree	1private	1	0.154	0.348	393	0.440	0.658	-0.530	0.839	0.539	0.087	0.370	0.698
1YesDegree	0public	1	0.630	0.822	393	0.770	0.444	-0.986	2.245	0.652	0.186	0.272	0.904
1YesDegree	1private	1	0.884	0.745	393	1.190	0.236	-0.581	2.349	0.708	0.154	0.359	0.913

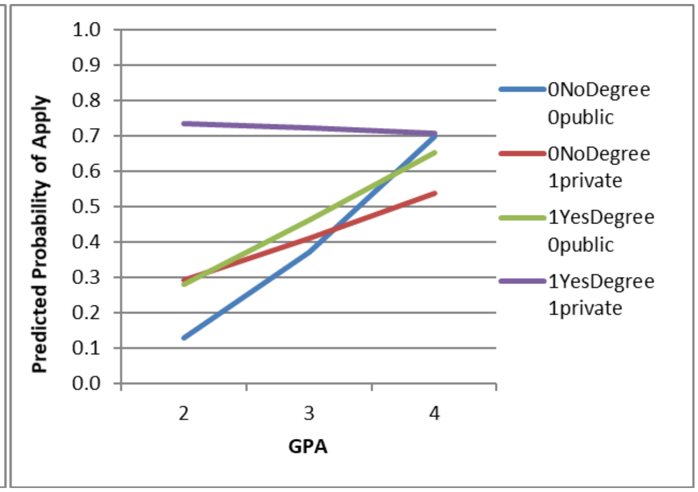
LSMEANS provides the model-predicted outcomes in two metrics: in logits (Estimate, SE, t-Value, p-value, Lower, and Upper) and in probability (mean, SE, lower mean, upper mean).

`SLICEDIFF output (cleaned up from ODS OUTPUT):`

Simple Effect Level	Alt	Ref	GPA3	Estimate	Standard Error	DF	t Value	Pr > |t|	Logit Lower	Logit Upper	Odds Ratio	Odds Lower	Odds Upper
parentGD 0NoDegree	0public	1private	-1	-1.011	1.011	393	-1.000	0.318	-2.998	0.977	0.364	0.050	2.656
parentGD 1YesDegree	0public	1private	-1	-1.957	1.280	393	-1.530	0.127	-4.473	0.559	0.141	0.011	1.749
private 0public	0NoDegree	1YesDegree	-1	-0.958	1.269	393	-0.750	0.451	-3.453	1.537	0.384	0.032	4.650
private 1private	**0NoDegree**	**1YesDegree**	**-1**	**-1.904**	**0.943**	**393**	**-2.020**	**0.044**	**-3.759**	**-0.050**	**0.149**	**0.023**	**0.951**
parentGD 0NoDegree	0public	1private	0	-0.159	0.370	393	-0.430	0.666	-0.886	0.567	0.853	0.412	1.763
parentGD 1YesDegree	0public	1private	0	-1.106	0.728	393	-1.520	0.130	-2.537	0.326	0.331	0.079	1.385
private 0public	0NoDegree	1YesDegree	0	-0.371	0.717	393	-0.520	0.605	-1.781	1.039	0.690	0.169	2.826
private 1private	**0NoDegree**	**1YesDegree**	**0**	**-1.317**	**0.349**	**393**	**-3.780**	**0.000**	**-2.003**	**-0.632**	**0.268**	**0.135**	**0.532**
parentGD 0NoDegree	0public	1private	1	0.692	0.762	393	0.910	0.365	-0.807	2.190	1.997	0.446	8.937
parentGD 1YesDegree	0public	1private	1	-0.255	0.867	393	-0.290	0.769	-1.960	1.451	0.775	0.141	4.265
private 0public	0NoDegree	1YesDegree	1	0.216	0.843	393	0.260	0.798	-1.440	1.873	1.242	0.237	6.510
private 1private	0NoDegree	1YesDegree	1	-0.730	0.804	393	-0.910	0.364	-2.310	0.850	0.482	0.099	2.339

SLICEDIFF provides the pairwise differences between cells at each value of the interacting predictors in two metrics: in logits (Estimate, SE, t-Value, p-value, Lower, and Upper) and in odds ratios (OR, lower odds, upper odds).





The model provides direct tests of the differences in logits amongst the degree and school conditions, as well as for the simple slopes of GPA for each degree and school combination. Model-predicted logit outcomes can then be converted through an inverse link ("un-logit") into predicted probabilities for ease of interpretation, but the slopes or mean differences themselves cannot be converted in differences in probabilities, only odds ratios.

Part 1 sample results section using SAS output, reversing direction of main effects to go from 0 to 1:

We examined the extent to which a binary decision to apply to graduate school (55.00% 0=No, 45.00% 1=Pry) 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 private school (14.25% 0=No, 85.75% 1=Yes). Specifically, we estimated generalized linear models using maximum likelihood, in which the conditional probability of applying to graduate school was predicted using a logit link function and a conditional Bernoulli distribution. SAS GLIMMIX was used for estimation, such that all fixed effects were tested using residual denominator degrees of freedom (i.e., as is the case in traditional linear regression). 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 and LSMEANS statements were used to request simple effects and model-implied predicted outcomes.

The first model examined only the main effects of the three predictors, which together resulted in a significant model, $F(3, 396) = 6.32$, $p < .0003$. GPA had a significantly positive effect, such that for every unit greater GPA, the logit of applying to graduate school was greater by 0.548 (SE = 0.272; OR = 1.730). Likewise, the logit of applying to graduate school was significantly greater for students for whom at least one parent had a graduate degree by 1.060 (SE = 0.297, OR = 2.882). However, the logit of applying to graduate school was nonsignificantly greater for students who attended private school by 0.200 (SE = 0.305, OR = 1.224).

The second model then included all three possible two-way interactions among the predictors to examine moderation. The augmented model remained significant, $F(6, 393) = 3.61$, $p < .0017$, but it was not a significant improvement over the main effects model, $F(3, 393) = 1.06$, $p < .3654$. None of the individual interaction terms was significant, but the augmented model results in terms of simple effects and simple effect differences are reported for completeness. The effect of GPA was not significant in any of the four subgroups (i.e., formed by parent graduate degree by school type). Group differences were explored conditioning on GPA = 2, 3, and 4. Only two significant differences were found: for private school students, those who had at least one parent with a graduate degree were significantly more likely to apply to graduate school at GPA = 2 (OR = 6.711) and GPA = 3 (OR= 3.731), but not at GPA = 4 (OR = 2.074).

Example 1 Part 2: Predicting Ordinal or Nominal Outcomes via SAS GLIMMIX and STATA GOLOGIT2 *(data, syntax, and output available for SAS and STATA electronically)*

In part 2 of the example we will predict the **three-category decision** to apply to grad school instead. Because we can't use LSMEANS with ordinal models, I am treating *parentGD* and *private* as quantitative this time.

The standard STATA package for single-level ordinal regression, ologit, insists on providing thresholds instead of intercepts, it also does not have any means to test or specify non-proportional odds models, and it's unclear what its margins is providing in the logit metric. The same is true for their multilevel version, meologit. To solve these problems, we will be using the custom STATA program *Gologit2: A Program for Generalized Logistic Regression/Partial Proportional Odds Models for Ordinal Dependent Variables* written by Richard Williams. The manual is here: https://www.stata.com/meeting/4nasug/gologit2.pdf

Getting STATA ready for part 2:

```
* Save results to separate file
log using $filesave\Example1_Part2_STATA_Output.log, replace name(Example1_Part2)

* Need to find and install a new package -- first run this:
search gologit2
* Now click on first link under "web resources" in new window to install, then can use gologit2
```

Getting SAS ready for part 2:

```
* Open output directory to save results to;
ODS RTF FILE="&filesave.\Example1_Part2_SAS_Output.rtf" STYLE=HTMLBlue STARTPAGE=NO;
```

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

$$Logit(Apply_i > 0) = \beta_{00} \rightarrow \; Probability(Apply_i > 0) = \frac{exp(\beta_{00})}{1+exp(\beta_{00})}$$

$$Logit(Apply_i > 1) = \beta_{01} \rightarrow \; Probability(Apply_i > 1) = \frac{exp(\beta_{01})}{1+exp(\beta_{01})}$$

```
display as result "STATA EMPTY MODEL PREDICTING ORIDNAL DV"
display as result "(ME)OLOGIT USES THRESHOLDS (GIVE LOGIT OF LOWER CATEGORY), NOT INTERCEPTS"
meologit apply ,
estat ic, n(400),
nlcom 1/(1+exp(-1*(_b[cut2:_cons]))) // threshold for y<2 in probability
nlcom 1/(1+exp(-1*(_b[cut1:_cons]))) // threshold for y<1 in probability


display as result "STATA EMPTY MODEL PREDICTING ORDINAL DV"
display as result "GOLOGIT2 USES INTERCEPTS (GIVE LOGIT OF HIGHER CATEGORY), NOT THRESHOLDS"
gologit2 apply ,
estat ic, n(400),
margins


TITLE1 "SAS EMPTY MODEL PREDICTING ORDINAL DV";
PROC GLIMMIX DATA=work.ologit NOCLPRINT GRADIENT METHOD=MSPL;
MODEL apply (DESCENDING) = / SOLUTION LINK=CLOGIT DIST=MULT;
ESTIMATE "Intercept for y>1 (=2)"   intercept 1 0 / ILINK; * Highest is first;
ESTIMATE "Intercept for y>0 (=1+2)" intercept 0 1 / ILINK;
RUN; TITLE1;
```

STATA Output (condensed):

```
. meologit apply ,
Ordered logistic regression                     Number of obs     =        400
                                                 chi2()           =          .
Log likelihood = -370.60264                     Prob > chi2       =          .
------------------------------------------------------------------------------
       apply |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       /cut1 |   .2006707   .1005038     2.00   0.046     .0036869    .3976545 → threshold logit of y<1
       /cut2 |   2.197225   .1666667    13.18   0.000     1.870564    2.523885 → threshold logit of y<2
------------------------------------------------------------------------------

. gologit2 apply ,
Generalized Ordered Logit Estimates             Number of obs     =        400
                                                LR chi2(0)        =      -0.00
                                                Prob > chi2       =          .
Log likelihood = -370.60264                     Pseudo R2         =    -0.0000
------------------------------------------------------------------------------
       apply |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
0No          |
       _cons |  -.2006707   .1005038    -2.00   0.046    -.3976545   -.0036869 → intercept logit of y>0
-------------+----------------------------------------------------------------
1Eh          |
       _cons |  -2.197225   .1666667   -13.18   0.000    -2.523885   -1.870564 → intercept logit of y>1
------------------------------------------------------------------------------
```

SAS Output (condensed):

```
                                Iteration History
                                           Objective                      Max
Iteration    Restarts    Evaluations        Function        Change    Gradient
        0           0              4    370.60264132

      Convergence criterion (ABSGCONV=0.00001) satisfied.
```

Hooray! The overall model converged! But we still need to watch for crazy SEs and gradients ≠ 0 for problems with individual model parameters…

```
           Fit Statistics
-2 Log Likelihood            741.21 → is 2 times last value above
AIC  (smaller is better)     745.21
BIC  (smaller is better)     753.19
```

Parameter Estimates

Effect	apply	Estimate	Standard Error	DF	t Value	Pr > |t|	Gradient
Intercept	2Very	-2.1972	0.1667	398	-13.18	<.0001	6.19E-15 → logit of y>1
Intercept	1Eh	-0.2007	0.1005	398	-2.00	0.0465	3.66E-15 → logit of 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	-2.1972	0.1667	398	-13.18	<.0001	0.1000	0.01500
Intercept for y>0	-0.2007	0.1005	398	-2.00	0.0465	0.4500	0.02487

$$Probability(Apply_i > 0) = \frac{exp(-0.2007)}{[1+exp(-0.2007)]} /= .450$$

$$Probability(Apply_i > 1) = \frac{exp(-2.1972)}{[1+exp(-2.1972)]} /= .100$$

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

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

Let's add some predictors, starting with main effects only...

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

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

```
display as result "STATA MAIN EFFECTS PROPORTIONAL ODDS MODEL PREDICTING ORDINAL DV"
gologit2 apply c.gpa3 c.parentgd c.private, p1
estat ic, n(400),
test (c.gpa3=0)(c.parentgd=0)(c.private=0)  // demo of Wald test of R2
margins, at(c.gpa3=(-1(1)1) c.parentgd=(0(1)1) c.private=(0(1)1)) predict(xb) // logits of y>0
margins, at(c.gpa3=(-1(1)1) c.parentgd=(0(1)1) c.private=(0(1)1))         // all probabilities

display as result "STATA MAIN EFFECTS PROPOTIONAL ODDS MODEL PREDICTING ORDINAL DV"
display as result "ODDS RATIOS INSTEAD OF LOGITS"
gologit2 apply c.gpa3 c.parentgd c.private, pl or
```

```
TITLE1 "SAS MAIN EFFECTS PROPORTIONAL ODDS MODEL PREDICTING ORDINAL DV";
PROC GLIMMIX DATA=work.ologit NOCLPRINT GRADIENT METHOD=MSPL;
MODEL apply (DESCENDING) = GPA3 parentGD private / SOLUTION LINK=CLOGIT DIST=MULT
                                                   ODDSRATIO(AT GPA3=0 LABEL);
CONTRAST "Wald test of R2" GPA3 1, parentGD 1, private 1 / CHISQ;
ESTIMATE "y > 1 Int for No  Degree, Public,  GPA=2" int 1 0 GPA3 -1 parentGD 0 private 0 / ILINK;
ESTIMATE "y > 1 Int for No  Degree, Private, GPA=2" int 1 0 GPA3 -1 parentGD 0 private 1 / ILINK;
ESTIMATE "y > 1 Int for Yes Degree, Public,  GPA=2" int 1 0 GPA3 -1 parentGD 1 private 0 / ILINK;
ESTIMATE "y > 1 Int for Yes Degree, Private, GPA=2" int 1 0 GPA3 -1 parentGD 1 private 1 / ILINK;
ESTIMATE "y > 1 Int for No  Degree, Public,  GPA=3" int 1 0 GPA3  0 parentGD 0 private 0 / ILINK;
ESTIMATE "y > 1 Int for No  Degree, Private, GPA=3" int 1 0 GPA3  0 parentGD 0 private 1 / ILINK;
ESTIMATE "y > 1 Int for Yes Degree, Public,  GPA=3" int 1 0 GPA3  0 parentGD 1 private 0 / ILINK;
ESTIMATE "y > 1 Int for Yes Degree, Private, GPA=3" int 1 0 GPA3  0 parentGD 1 private 1 / ILINK;
ESTIMATE "y > 1 Int for No  Degree, Public,  GPA=4" int 1 0 GPA3  1 parentGD 0 private 0 / ILINK;
ESTIMATE "y > 1 Int for No  Degree, Private, GPA=4" int 1 0 GPA3  1 parentGD 0 private 1 / ILINK;
ESTIMATE "y > 1 Int for Yes Degree, Public,  GPA=4" int 1 0 GPA3  1 parentGD 1 private 0 / ILINK;
ESTIMATE "y > 1 Int for Yes Degree, Private, GPA=4" int 1 0 GPA3  1 parentGD 1 private 1 / ILINK;

ESTIMATE "y > 0 Int for No  Degree, Public,  GPA=2" int 0 1 GPA3 -1 parentGD 0 private 0 / ILINK;
ESTIMATE "y > 0 Int for No  Degree, Private, GPA=2" int 0 1 GPA3 -1 parentGD 0 private 1 / ILINK;
ESTIMATE "y > 0 Int for Yes Degree, Public,  GPA=2" int 0 1 GPA3 -1 parentGD 1 private 0 / ILINK;
ESTIMATE "y > 0 Int for Yes Degree, Private, GPA=2" int 0 1 GPA3 -1 parentGD 1 private 1 / ILINK;
ESTIMATE "y > 0 Int for No  Degree, Public,  GPA=3" int 0 1 GPA3  0 parentGD 0 private 0 / ILINK;
ESTIMATE "y > 0 Int for No  Degree, Private, GPA=3" int 0 1 GPA3  0 parentGD 0 private 1 / ILINK;
ESTIMATE "y > 0 Int for Yes Degree, Public,  GPA=3" int 0 1 GPA3  0 parentGD 1 private 0 / ILINK;
ESTIMATE "y > 0 Int for Yes Degree, Private, GPA=3" int 0 1 GPA3  0 parentGD 1 private 1 / ILINK;
ESTIMATE "y > 0 Int for No  Degree, Public,  GPA=4" int 0 1 GPA3  1 parentGD 0 private 0 / ILINK;
ESTIMATE "y > 0 Int for No  Degree, Private, GPA=4" int 0 1 GPA3  1 parentGD 0 private 1 / ILINK;
ESTIMATE "y > 0 Int for Yes Degree, Public,  GPA=4" int 0 1 GPA3  1 parentGD 1 private 0 / ILINK;
ESTIMATE "y > 0 Int for Yes Degree, Private, GPA=4" int 0 1 GPA3  1 parentGD 1 private 1 / ILINK;
ESTIMATE "Slope for GPA" GPA3 1 / ILINK; * Example of non-sense ILINK for a slope;
RUN; TITLE1;
```

SAS Output (condensed and re-arranged for convenience):

```
           Fit Statistics
-2 Log Likelihood             717.02
AIC  (smaller is better)      727.02
BIC  (smaller is better)      746.98

                                      Contrasts
                       Num      Den
Label                   DF       DF     Chi-Square     F Value       Pr > ChiSq     Pr > F
Wald test of R2          3      395          23.61        7.87           <.0001     <.0001
```

Parameter Estimates

Effect	apply	Estimate	Standard Error	DF	t Value	Pr > |t|	Gradient	
Intercept	2Very	-2.5102	0.3192	395	-7.86	<.0001	7.07E-12	Beta01
Intercept	1Eh	-0.4148	0.2830	395	-1.47	0.1435	-262E-13	Beta00
GPA3		0.6157	0.2606	395	2.36	0.0186	5.41E-13	Beta1
parentGD		1.0477	0.2658	395	3.94	<.0001	1.92E-12	Beta2
private		0.05868	0.2979	395	0.20	0.8439	6.07E-12	Beta3

Interpret each fixed effect…

Intercept for 2Yes:

Intercept for 1Maybe:

GPA3:

parentGD:

private:

Odds Ratio Estimates

Comparison	Estimate	DF	95% Confidence Limits		
unit change of GPA3 from GPA3=0	1.851	395	1.109	3.090	EXP(Beta1)
unit change of parentGD from GPA3=0	2.851	395	1.691	4.808	EXP(Beta2)
unit change of private from GPA3=0	1.060	395	0.590	1.905	EXP(Beta3)

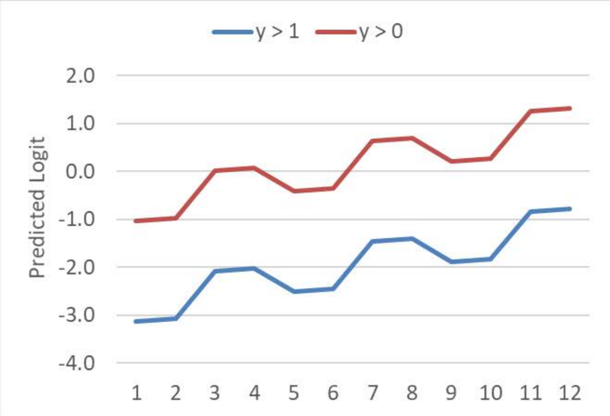
Effects of continuous variables are assessed as units offsets from the reference value.
The UNIT suboption modifies the offsets.

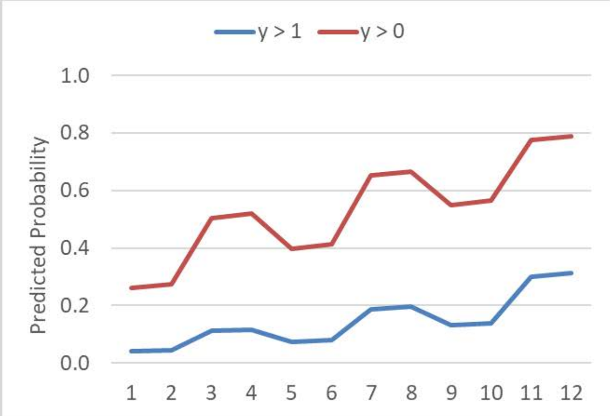
Estimates

Label	Estimate	Standard Error	DF	t Value	Pr > |t|	Mean	Standard Error Mean
y > 1 Int for No Degree, Public, GPA=2	-3.1260	0.4525	395	-6.91	<.0001	0.04205	0.01823
y > 1 Int for No Degree, Private, GPA=2	-3.0673	0.3211	395	-9.55	<.0001	0.04448	0.01365
y > 1 Int for Yes Degree, Public, GPA=2	-2.0783	0.4971	395	-4.18	<.0001	0.1112	0.04914
y > 1 Int for Yes Degree, Private, GPA=2	-2.0196	0.3893	395	-5.19	<.0001	0.1172	0.04027
y > 1 Int for No Degree, Public, GPA=3	-2.5102	0.3192	395	-7.86	<.0001	0.07515	0.02218
y > 1 Int for No Degree, Private, GPA=3	-2.4515	0.1870	395	-13.11	<.0001	0.07933	0.01366
y > 1 Int for Yes Degree, Public, GPA=3	-1.4625	0.3608	395	-4.05	<.0001	0.1881	0.05510
y > 1 Int for Yes Degree, Private, GPA=3	-1.4039	0.2634	395	-5.33	<.0001	0.1972	0.04170
y > 1 Int for No Degree, Public, GPA=4	-1.8945	0.3672	395	-5.16	<.0001	0.1307	0.04173
y > 1 Int for No Degree, Private, GPA=4	-1.8358	0.3204	395	-5.73	<.0001	0.1376	0.03801
y > 1 Int for Yes Degree, Public, GPA=4	-0.8468	0.3862	395	-2.19	0.0289	0.3001	0.08111
y > 1 Int for Yes Degree, Private, GPA=4	-0.7881	0.3508	395	-2.25	0.0252	0.3126	0.07538
y > 0 Int for No Degree, Public, GPA=2	-1.0305	0.4206	395	-2.45	0.0147	0.2630	0.08153
y > 0 Int for No Degree, Private, GPA=2	-0.9718	0.2754	395	-3.53	0.0005	0.2745	0.05486
y > 0 Int for Yes Degree, Public, GPA=2	0.01715	0.4839	395	0.04	0.9717	0.5043	0.1210
y > 0 Int for Yes Degree, Private, GPA=2	0.07583	0.3731	395	0.20	0.8391	0.5189	0.09315
y > 0 Int for No Degree, Public, GPA=3	-0.4148	0.2830	395	-1.47	0.1435	0.3978	0.06779
y > 0 Int for No Degree, Private, GPA=3	-0.3561	0.1173	395	-3.04	0.0025	0.4119	0.02840
y > 0 Int for Yes Degree, Public, GPA=3	0.6329	0.3511	395	1.80	0.0722	0.6531	0.07954
y > 0 Int for Yes Degree, Private, GPA=3	0.6916	0.2511	395	2.75	0.0062	0.6663	0.05583
y > 0 Int for No Degree, Public, GPA=4	0.2010	0.3451	395	0.58	0.5606	0.5501	0.08540
y > 0 Int for No Degree, Private, GPA=4	0.2597	0.2958	395	0.88	0.3805	0.5646	0.07271
y > 0 Int for Yes Degree, Public, GPA=4	1.2486	0.3850	395	3.24	0.0013	0.7771	0.06670
y > 0 Int for Yes Degree, Private, GPA=4	1.3073	0.3504	395	3.73	0.0002	0.7871	0.05872
Slope for GPA	0.6157	0.2606	395	2.36	0.0186	0.6493	0.05935

The last line illustrates why you cannot "un-logit" a slope… the difference between the intercepts per unit GPA in logits is a constant 0.6157, but the difference in probability is not constant.

Similarly, the difference between the 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).





More importantly, these models rely on an assumption of <u>proportional odds</u>: that all predictor slopes are equal across sub-models. Here is an alternative, a non-proportional odds univariate model:

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

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

```
display as result "STATA MAIN EFFECTS MODEL TESTING ORDINAL PROPORTIONAL ODDS"
display as result "DIRECTLY PROVIDES EACH SLOPE AND DIFFERENCES IN SLOPES ACROSS SUBMODELS"
gologit2 apply c.gpa3 c.parentgd c.private, gamma // test: Multiv Wald Model R2 test
test ([0No]gpa3=0)([1Eh]gpa3=0)([0No]parentgd=0)([1Eh]parentgd=0)([0No]private=0)([1Eh]private=0)
test ([0No]gpa3=[1Eh]gpa3)([0No]parentgd=[1Eh]parentgd)([0No]private=[1Eh]private)
      // Global PO test
estat ic, n(400),

TITLE1 "SAS MAIN EFFECTS MODEL TESTING ORDINAL PROPORTIONAL ODDS";
TITLE2 "MUST USE LOGISTIC INSTEAD OF GLIMMIX";
PROC LOGISTIC DATA=work.ologit;
MODEL apply (DESCENDING) = GPA3 parentGD private / LINK=CLOGIT UNEQUALSLOPES;
* Request multiv Wald test of model R2;
Model:     TEST GPA3_2Very=0, GPA3_1Eh=0, parentGD_2Very=0,
                parentGD_1Eh=0, private_2Very=0, private_1Eh=0;
* Requesting tests of differences between slopes and DF=3 test;
GPA3:      TEST GPA3_2Very=GPA3_1Eh;
parentGD:  TEST parentGD_2Very=parentGD_1Eh;
private:   TEST private_2Very=private_1Eh;
PropOdds:  TEST GPA3_2Very=GPA3_1Eh, parentGD_2Very=parentGD_1Eh, private_2Very=private_1Eh;
RUN;
```

STATA Output (condensed):

```
. gologit2 apply c.gpa3 c.parentgd c.private
Generalized Ordered Logit Estimates               Number of obs     =        400
                                                  LR chi2(6)        =      28.19
                                                  Prob > chi2       =     0.0001
Log likelihood = -356.50556                       Pseudo R2         =     0.0380
------------------------------------------------------------------------------
       apply |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
0No          |
        gpa3 |   .5920653   .2690337     2.20   0.028     .0647689    1.119362 Beta10
    parentgd |   1.083129   .2959475     3.66   0.000     .5030823    1.663175 Beta20
     private |   .2307488   .3062506     0.75   0.451    -.3694912    .8309889 Beta30
       _cons |  -.5684777   .2888819    -1.97   0.049    -1.134676   -.0022796 Beta00
-------------+----------------------------------------------------------------
```

```
1Eh          |
        gpa3 |   .7190314   .4536953     1.58   0.113    -.1701951    1.608258 Beta11
    parentgd |   .9946781   .3740984     2.66   0.008     .2614588    1.727897 Beta21
     private |  -.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
    parentgd |   1.083129   .2959475     3.66   0.000     .5030823    1.663175 Beta20
     private |   .2307488   .3062506     0.75   0.451    -.3694912    .8309889 Beta30
-------------+----------------------------------------------------------------
Gamma_2      |
        gpa3 |   .1269661   .4383381     0.29   0.772    -.7321607     .986093 Beta11 – Beta10
    parentgd |  -.0884506   .3871321    -0.23   0.819    -.8472157    .6703144 Beta21 – Beta20
     private |  -.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
------------------------------------------------------------------------------

. test ([0]gpa3=[1]gpa3)([0]parentgd=[1]parentgd)([0]private=[1]private) // Global PO test
           chi2(  3) =    4.73
         Prob > chi2 =    0.1927
```

SAS Output (condensed):

```
         Model Fit Statistics
                             Intercept
              Intercept            and
Criterion          Only     Covariates
AIC             745.205        729.011
SC              753.188        760.943
-2 Log L        741.205        713.011

                   Analysis of Maximum Likelihood Estimates
                                          Standard          Wald
Parameter    APPLY    DF    Estimate        Error    Chi-Square    Pr > ChiSq
Intercept    2Very     1     -2.0276       0.4050       25.0617        <.0001  Beta01
Intercept    1Eh       1     -0.5685       0.2889        3.8725        0.0491  Beta00
GPA3         2Very     1      0.7190       0.4537        2.5117        0.1130  Beta11
GPA3         1Eh       1      0.5921       0.2690        4.8431        0.0278  Beta10
parentGD     2Very     1      0.9947       0.3741        7.0696        0.0078  Beta21
parentGD     1Eh       1      1.0831       0.2959       13.3946        0.0003  Beta20
private      2Very     1     -0.5367       0.4293        1.5628        0.2113  Beta31
private      1Eh       1      0.2307       0.3063        0.5677        0.4512  Beta30

      Linear Hypotheses Testing Results (from TEST statements)
                   Wald
 Label       Chi-Square      DF    Pr > ChiSq
 Model          27.7201       6        0.0001 → the model is significant (multiv Wald test)
 GPA3            0.0839       1        0.7721 → these slopes are not different
 parentGD        0.0522       1        0.8193 → these slopes are not different
 private         3.5800       1        0.0585 → these slopes are almost not different
 PropOdds        4.7298       3        0.1927 → Overall test of overall proportional odds violations
```

Both programs have a way to automate the selection of which slopes should differ—let's see what happens when we let them do it (but request that all predictors remain the model even if NS):

```
display as result "STATA MAIN EFFECTS ORDINAL MODEL -- AUTOMATED SELECTION OF PROPORTIONAL ODDS"
gologit2 apply c.gpa3 c.parentgd c.private, gamma auto
estat ic, n(400)

TITLE1 "SAS MAIN EFFECTS ORDINAL MODEL -- AUTOMATED SELECTION OF PROPORTIONAL ODDS";
PROC LOGISTIC DATA=work.ologit;
MODEL apply (DESCENDING) = GPA3 parentGD private / LINK=CLOGIT
/* Start with unequal slopes and remove if p>.05, keep 3 slopes minimum */
         EQUALSLOPES UNEQUALSLOPES SELECTION=BACKWARD STOP=3 DETAILS; RUN;
```

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

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

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

STATA Output (condensed):

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

 ( 1)  [0]parentgd - [1]parentgd = 0
 ( 2)  [0]gpa3 - [1]gpa3 = 0
------------------------------------------------------------------------------
       apply |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
0No          |
        gpa3 |   .6105983   .2607849     2.34   0.019     .0994694    1.121727 Beta1
    parentgd |   1.057633   .2665412     3.97   0.000     .5352216    1.580044 Beta2
     private |   .2350038   .3052548     0.77   0.441    -.3632847    .8332922 Beta30
       _cons |  -.5690629   .2876884    -1.98   0.048    -1.132922    -.005204 Beta00
-------------+----------------------------------------------------------------
1Eh          |
        gpa3 |   .6105983   .2607849     2.34   0.019     .0994694    1.121727 Beta1
    parentgd |   1.057633   .2665412     3.97   0.000     .5352216    1.580044 Beta2
     private |  -.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
------------------------------------------------------------------------------
       apply |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Beta         |
        gpa3 |   .6105983   .2607849     2.34   0.019     .0994694    1.121727 Beta1
    parentgd |   1.057633   .2665412     3.97   0.000     .5352216    1.580044 Beta2
     private |   .2350038   .3052548     0.77   0.441    -.3632847    .8332922 Beta30
-------------+----------------------------------------------------------------
Gamma_2      |
     private |  -.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
------------------------------------------------------------------------------
```

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

```
display as result "STATA MAIN EFFECTS ORDINAL MODEL -- CUSTOM SELECTION OF PROPORTIONAL ODDS"
gologit2 apply c.gpa3 c.parentgd c.private, npl(c.private) gamma
estat ic, n(400)

display as result "STATA MAIN EFFECTS ORDINAL MODEL -- CUSTOM SELECTION OF PROPORTIONAL ODDS"
display as result "ODDS RATIOS INSTEAD OF LOGITS"
gologit2 apply c.gpa3 c.parentgd c.private, npl(c.private) gamma or
```

```
TITLE1 "SAS MAIN EFFECTS MODEL -- CUSTOM SELECTION OF PROPORTIONAL ODDS";
PROC LOGISTIC DATA=work.ologit;
MODEL apply (DESCENDING) = GPA3 parentGD private / LINK=CLOGIT UNEQUALSLOPES=private;
private: TEST private_2Very=private_1Eh; RUN;
```

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

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

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

```
TITLE1 "SAS MAIN EFFECTS ORDINAL MODEL TESTING PROPORTIONAL ODDS";
TITLE2 "USING NLMIXED TO WRITE CUSTOM MODEL";
PROC NLMIXED DATA=work.ologit 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 B00=-0.4 B11GPA3=0 B10GPA3=0
      B21parentGD=0 B20parentGD=0 B31private=0 B30private=0;
* Linear predictor written as single-level equation for y>0 and y>1;
Y0 = B00 + B10GPA3*GPA3 + B20parentGD*parentGD + B30private*private;
Y1 = B01 + B11GPA3*GPA3 + B21parentGD*parentGD + B31private*private;
* Model for probability of response - writing it the shorter way;
     IF (apply=0) THEN P =  1 - (1/(1 + EXP(-Y0)));
ELSE IF (apply=1) THEN P =      (1/(1 + EXP(-Y0))) - (1/(1 + EXP(-Y1)));
ELSE IF (apply=2) THEN P =                           (1/(1 + EXP(-Y1)));
LL = LOG(P);
MODEL apply ~ GENERAL(LL);
* Testing proportional odds;
ESTIMATE "GPA Slope Diff"       B11GPA3     - B10GPA3;
ESTIMATE "ParentGD Slope Diff"  B21parentGD - B20parentGD;
ESTIMATE "Private Slope Diff"   B31private  - B30private;
CONTRAST "Overall Proportional Odds Test" B11GPA3-B10GPA3,
          B21parentGD-B20parentGD, B31private-B30private;
RUN;
```

SAS Output (condensed):

Fit Statistics

-2 Log Likelihood	713.0
AIC (smaller is better)	729.0
AICC (smaller is better)	729.4
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	-2.75E-7
B00	-0.5685	0.2889	400	-1.97	0.0498	-1.1364	-0.00056	-2.52E-7
B11GPA3	0.7190	0.4537	400	1.58	0.1138	-0.1729	1.6110	-1.18E-7
B10GPA3	0.5921	0.2690	400	2.20	0.0283	0.06317	1.1210	2.536E-8
B21parentGD	0.9947	0.3741	400	2.66	0.0082	0.2592	1.7301	-2E-7
B20parentGD	1.0831	0.2959	400	3.66	0.0003	0.5013	1.6649	7.605E-9
B31private	-0.5367	0.4293	400	-1.25	0.2120	-1.3807	0.3073	-1.22E-7
B30private	0.2307	0.3063	400	0.75	0.4516	-0.3713	0.8328	-3.86E-7

Contrasts

Label	Num DF	Den DF	F Value	Pr > F
Overall Proportional Odds Test	3	400	1.58	0.1945

```
                                    Additional Estimates

                                  Standard
Label                 Estimate       Error      DF    t Value    Pr > |t|     Alpha       Lower       Upper
GPA Slope Diff          0.1270      0.4383     400       0.29      0.7722      0.05     -0.7348      0.9887
ParentGD Slope Diff   -0.08845      0.3871     400      -0.23      0.8194      0.05     -0.8495      0.6726
Private 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 instead (i.e., unordered categories in which one is the reference by which to compare each of the others). For convenience we will choose Apply=1 to be the reference.

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

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

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

```
display as result "STATA EMPTY MODEL PREDICTING NOMINAL DV"
mlogit apply ,  baseoutcome(1),
estat ic, n(400),
nlcom 1/(1+exp(-1*(_b[0:_cons]))) // intercept for y=0 instead of 1 in probability
nlcom 1/(1+exp(-1*(_b[2:_cons]))) // intercept for y=2 instead of 1 in probability
```

```
TITLE1 "SAS EMPTY MODEL PREDICTING NOMINAL DV";
PROC GLIMMIX DATA=work.ologit NOCLPRINT GRADIENT METHOD=MSPL;
MODEL apply (REFERENCE="1Eh") = / SOLUTION LINK=GLOGIT DIST=MULT;
* ILINK requests logit estimate to be transformed into probability;
ESTIMATE "Intercept" intercept 1 / ILINK BYCAT; * BYCAT gives per submodel;
RUN; TITLE1;
```

STATA Output (condensed):

```
Multinomial logistic regression                 Number of obs     =        400
                                                LR chi2(0)        =       0.00
                                                Prob > chi2       =          .
Log likelihood = -370.60264                     Pseudo R2         =     0.0000
------------------------------------------------------------------------------
       apply |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
0No          |
       _cons |   .4519851   .1081125     4.18   0.000     .2400885    .6638817 → logit of 0 instead of 1
-------------+----------------------------------------------------------------
1Eh          |  (base outcome)
-------------+----------------------------------------------------------------
2Very        |
       _cons |  -1.252763   .1792843    -6.99   0.000    -1.604154   -.9013722 → logit of 2 instead of 1
------------------------------------------------------------------------------
```

SAS Output (condensed):

```
           Fit Statistics
-2 Log Likelihood             741.21
AIC  (smaller is better)      745.21
BIC  (smaller is better)      753.19
```

```
                                  Parameter Estimates
                                    Standard
Effect       Apply        Estimate      Error       DF    t Value    Pr > |t|     Gradient
Intercept    0No            0.4520     0.1081      398       4.18      <.0001     -107E-16
Intercept    2Very         -1.2528     0.1793      398      -6.99      <.0001     1.35E-14
```

Estimates

Label	Apply	Estimate	Standard Error	DF	t Value	Pr > |t|	Mean	Standard Error Mean
Intercept	0No	0.4520	0.1081	398	4.18	<.0001	0.5500	0.02487
Intercept	2Very	-1.2528	0.1793	398	-6.99	<.0001	0.1000	0.01500

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

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

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

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

Let's add some predictors...

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

```
display as result "STATA MAIN EFFECTS MODEL PREDICTING NOMINAL DV"
mlogit apply c.gpa3 c.parentgd c.private, baseoutcome(1),
estat ic, n(400), // test: Model R2 multivariate Wald test
test ([0No]c.gpa3=0)([2Very]c.gpa3=0)([0No]c.parentgd=0)([2Very]c.parentgd=0) ///
     ([0No]c.private=0)([2Very]c.private=0) // test: Model R2 multivariate Wald test
margins, at(c.gpa3=(-1(1)1) c.parentgd=(0(1)1) c.private=(0(1)1)) predict(xb) // logits for 1vs0
margins, at(c.gpa3=(-1(1)1) c.parentgd=(0(1)1) c.private=(0(1)1))            // all probabilities

display as result "STATA MAIN EFFECTS MODEL PREDICTING NOMINAL DV"
display as result "ODDS RATIOS INSTEAD OF LOGITS"
mlogit apply c.gpa3 c.parentgd c.private, baseoutcome(1) rrr,

* Close log
log close Example1_Part2
```

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….

```
TITLE1 "SAS MAIN EFFECTS MODEL PREDICTING NOMINAL DV";
PROC GLIMMIX DATA=work.ologit NOCLPRINT GRADIENT METHOD=MSPL;
MODEL apply (REFERENCE="1Eh") = GPA3 parentGD private / SOLUTION LINK=GLOGIT DIST=MULT
                                                        ODDSRATIO(AT GPA3=0 LABEL);
CONTRAST "Model R2 multivariate Wald test" GPA3 1, parentGD 1, private 1 / CHISQ;
ESTIMATE "Int for No  Degree, Public,  GPA=2" int 1 GPA3 -1 parentGD 0 private 0 / ILINK BYCAT;
ESTIMATE "Int for No  Degree, Private, GPA=2" int 1 GPA3 -1 parentGD 0 private 1 / ILINK BYCAT;
ESTIMATE "Int for Yes Degree, Public,  GPA=2" int 1 GPA3 -1 parentGD 1 private 0 / ILINK BYCAT;
ESTIMATE "Int for Yes Degree, Private, GPA=2" int 1 GPA3 -1 parentGD 1 private 1 / ILINK BYCAT;
ESTIMATE "Int for No  Degree, Public,  GPA=3" int 1 GPA3  0 parentGD 0 private 0 / ILINK BYCAT;
ESTIMATE "Int for No  Degree, Private, GPA=3" int 1 GPA3  0 parentGD 0 private 1 / ILINK BYCAT;
ESTIMATE "Int for Yes Degree, Public,  GPA=3" int 1 GPA3  0 parentGD 1 private 0 / ILINK BYCAT;
ESTIMATE "Int for Yes Degree, Private, GPA=3" int 1 GPA3  0 parentGD 1 private 1 / ILINK BYCAT;
ESTIMATE "Int for No  Degree, Public,  GPA=4" int 1 GPA3  1 parentGD 0 private 0 / ILINK BYCAT;
ESTIMATE "Int for No  Degree, Private, GPA=4" int 1 GPA3  1 parentGD 0 private 1 / ILINK BYCAT;
ESTIMATE "Int for Yes Degree, Public,  GPA=4" int 1 GPA3  1 parentGD 1 private 0 / ILINK BYCAT;
```

```
ESTIMATE "Int for Yes Degree, Private, GPA=4" int 1 GPA3  1 parentGD 1 private 1 / ILINK BYCAT;
RUN; TITLE1;

* Close output;
ODS RTF CLOSE;
```

SAS Output (condensed):

Fit Statistics

-2 Log Likelihood	713.99
AIC (smaller is better)	729.99
BIC (smaller is better)	761.93

Type III Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
GPA3	2	392	2.48	0.0853
parentGD	2	392	6.93	0.0011
private	2	392	1.52	0.2204

Contrasts

Label	Num DF	Den DF	Chi-Square	F Value	Pr > ChiSq	Pr > F
Model R2 multivariate Wald test	6	392	25.84	4.31	0.0002	0.0003

Parameter Estimates

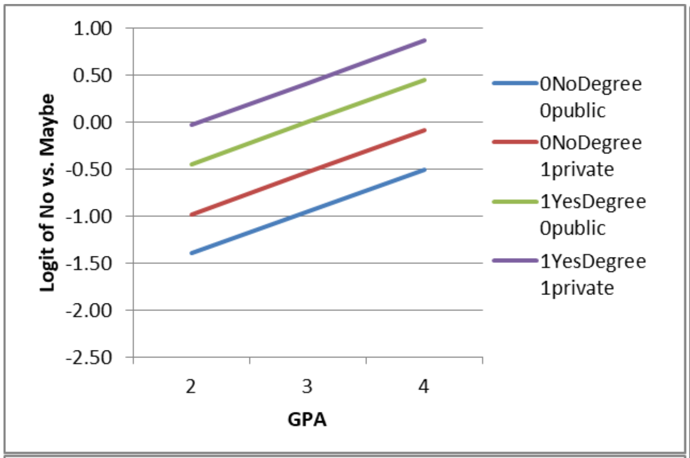
Effect	Apply	Estimate	Standard Error	DF	t Value	Pr > |t|	Gradient
Intercept	0No	0.9515	0.3258	392	2.92	0.0037	-2.37E-8
Intercept	2Very	-0.7641	0.4511	392	-1.69	0.0911	3.684E-8
GPA3	0No	-0.4488	0.2902	392	-1.55	0.1228	8.02E-11
GPA3	2Very	0.4753	0.4871	392	0.98	0.3298	-597E-12
parentGD	**0No**	**-0.9516**	**0.3171**	**392**	**-3.00**	**0.0029**	**-1.6E-8**
parentGD	2Very	0.4225	0.4083	392	1.03	0.3014	2.634E-8
private	0Very	-0.4188	0.3433	392	-1.22	0.2232	-1.27E-8
private	**2Yes**	**-0.7789**	**0.4706**	**392**	**-1.66**	**0.0987**	**2.266E-8**

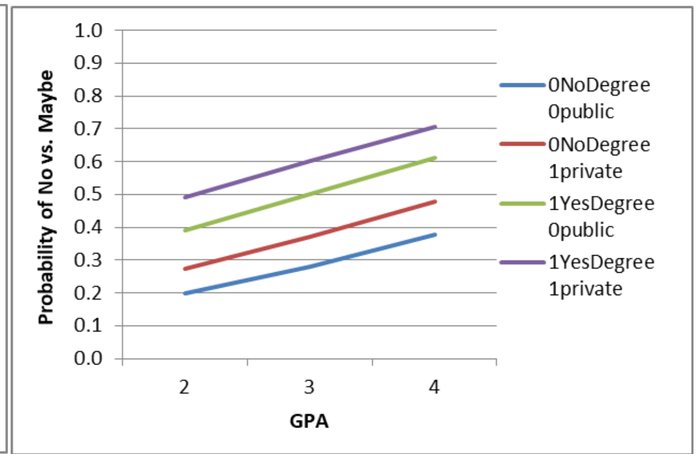
Odds Ratio Estimates

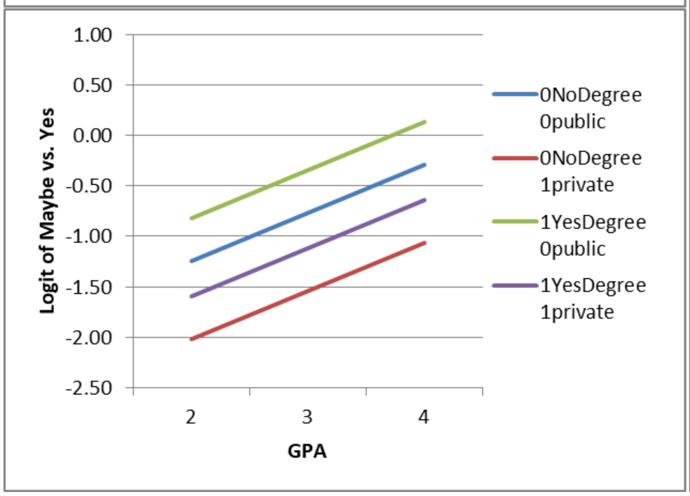
Comparison	Estimate	DF	95% Confidence Limits	
0: unit change of GPA3 from GPA3=0	0.638	392	0.361	1.130
2: unit change of GPA3 from GPA3=0	1.608	392	0.617	4.191
0: unit change of parentGD from GPA3=0	0.386	392	0.207	0.720
2: unit change of parentGD from GPA3=0	1.526	392	0.684	3.405
0: unit change of private from GPA3=0	0.658	392	0.335	1.292
2: unit change of private from GPA3=0	0.459	392	0.182	1.158

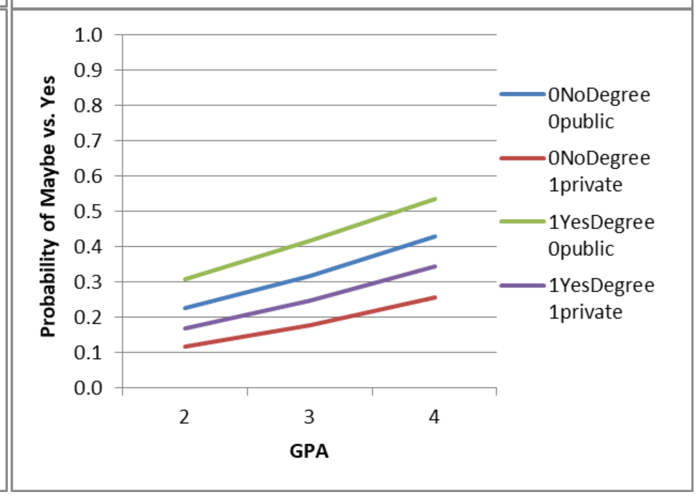
Effects of continuous variables are assessed as units offsets from
the reference value. The UNIT suboption modifies the offsets.

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….









Estimates (how plots were made, reversing direction of predictions from 1 vs 0)

Label	Apply	Estimate	Standard Error	DF	t Value	Pr > |t|	Mean	Standard Error Mean
Int: No Degree, Public, GPA=2	0No	1.4003	0.4714	392	2.97	0.0032	0.7588	0.07966
Int: No Degree, Public, GPA=2	2Very	-1.2393	0.7516	392	-1.65	0.0999	0.05417	0.03582
Int: No Degree, Private, GPA=2	0No	0.9815	0.3014	392	3.26	0.0012	0.7020	0.05998
Int: No Degree, Private, GPA=2	2Very	-2.0182	0.5496	392	-3.67	0.0003	0.03496	0.01755
Int: Yes Degree, Public, GPA=2	0No	0.4486	0.5564	392	0.81	0.4206	0.5207	0.1294
Int: Yes Degree, Public, GPA=2	2Very	-0.8168	0.8380	392	-0.97	0.3303	0.1469	0.09749
Int: Yes Degree, Private, GPA=2	0No	0.02981	0.4335	392	0.07	0.9452	0.4614	0.1018
Int: Yes Degree, Private, GPA=2	2Very	-1.5957	0.6744	392	-2.37	0.0185	0.09080	0.05246
Int: No Degree, Public, GPA=3	0No	0.9515	0.3258	392	2.92	0.0037	0.6386	0.06687
Int: No Degree, Public, GPA=3	2Very	-0.7641	0.4511	392	-1.69	0.0911	0.1149	0.04087
Int: No Degree, Private, GPA=3	0No	0.5327	0.1261	392	4.23	<.0001	0.5839	0.02887
Int: No Degree, Private, GPA=3	2Very	-1.5429	0.2322	392	-6.64	<.0001	0.07327	0.01487
Int: Yes Degree, Public, GPA=3	0No	-0.00012	0.4111	392	-0.00	0.9998	0.3689	0.08788
Int: Yes Degree, Public, GPA=3	2Very	-0.3416	0.5184	392	-0.66	0.5104	0.2622	0.09139
Int: Yes Degree, Private, GPA=3	0No	-0.4189	0.2974	392	-1.41	0.1597	0.3315	0.06217
Int: Yes Degree, Private, GPA=3	2Very	-1.1204	0.3669	392	-3.05	0.0024	0.1644	0.04752
Int: No Degree, Public, GPA=4	0No	0.5028	0.3982	392	1.26	0.2074	0.4859	0.09194
Int: No Degree, Public, GPA=4	2Very	-0.2888	0.5628	392	-0.51	0.6082	0.2202	0.08758
Int: No Degree, Private, GPA=4	0No	0.08396	0.3307	392	0.25	0.7997	0.4473	0.07667
Int: No Degree, Private, GPA=4	2Very	-1.0677	0.5295	392	-2.02	0.0444	0.1414	0.06020
Int: Yes Degree, Public, GPA=4	0No	-0.4489	0.4438	392	-1.01	0.3124	0.2295	0.07504
Int: Yes Degree, Public, GPA=4	2Very	0.1337	0.5567	392	0.24	0.8103	0.4110	0.1242
Int: Yes Degree, Private, GPA=4	0No	-0.8677	0.3967	392	-2.19	0.0293	0.2160	0.06392
Int: Yes Degree, Private, GPA=4	2Very	-0.6451	0.5375	392	-1.20	0.2308	0.2698	0.1005

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 private school (14.25% 0=No, 85.75% 1=Yes). Specifically, we estimated two alternative sets of generalized linear models with conditional multinomial distributions, as described below. SAS GLIMMIX was used for the maximum likelihood estimation, such that all fixed effects were tested using residual denominator degrees of freedom (i.e., as is the case in traditional linear regression). 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 and LSMEANS statements were used to request simple effects and model-implied predicted outcomes.

First, we treated the 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, $F(3, 395) = 7.87$, $p < .0001$. 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 private school 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 $y_i > 1$.

Second, we treated the outcome as nominal using a generalized logit link function—this approach requires choosing a reference category (1Eh). 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 0No instead of 1Eh 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 2Very instead of 1Eh was significantly smaller for students who went to private school by 0.779 (SE = 0.471, OR = 0.459).