Examples of Modeling Binary Outcomes via SAS PROC GLIMMIX and STATA XTMELOGIT
(data, syntax, and output available for SAS and STATA electronically)

The (likely fake) data for this example came from: http://www.ats.ucla.edu/stat/sas/dae/ologit.htm.
In this example we will predict a binary decision to apply to grad school based on student GPA, whether at least one of their parents has a graduate degree, and whether they attended private school.

STATA Syntax and Output for Data Manipulation:

```
* 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 fparentgd   0 "0NoDegree" 1 "1YesDegree"
label define fprivate    0 "0public" 1 "1private"
label values apply2 f2apply
label values parentgd fparentgd
label values private fprivate
```

SAS Syntax and Output for Data Manipulation:

```
* Creating formats for categorical variables;
PROC FORMAT;
    VALUE F2apply   0="0No" 1="1Pry";
    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;

TITLE1 "DESCRIPTIVES FOR STUDY VARIABLES";
PROC MEANS DATA=work.ologit; VAR GPA; RUN;
PROC FREQ DATA=work.ologit; TABLE apply2 parentGD private; RUN; TITLE1;
```

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

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

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

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

TITLE1 "EMPTY MODEL PREDICTING BINARY DV";
PROC GLIMMIX DATA=work.ologit NOCLPRINT NOITPRINT GRADIENT METHOD=QUAD;
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'.

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

Hooray! Our estimates are usable!

Fit Statistics

-2 Log Likelihood	550.51
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

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

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

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 "MAIN EFFECTS MODEL PREDICTING BINARY DV"
melogit apply2 c.gpa3 ib(last).parentgd ib(last).private,
    estat ic, n(400),
    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

TITLE1 "MAIN EFFECTS MODEL PREDICTING BINARY DV";
PROC GLIMMIX DATA=work.ologit NOCLPRINT NOITPRINT GRADIENT METHOD=QUAD;
CLASS parentGD private;
MODEL apply2 (DESCENDING) = GPA3 parentGD private
                  / SOLUTION LINK=LOGIT DIST=BINARY ODDSRATIO(AT GPA3=0);
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 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

                                          Parameter Estimates
                          private:
              parentGD:   Student
              Parent Has  Attends
              Graduate    Private
              Degree      University                Standard
Effect        (0=N,1=Y)   (0=N,1=Y)     Estimate      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 main effect…

Intercept:

GPA3:

parentGD:

private:

Odds Ratio Estimates

parentGD: Parent Has Graduate Degree (O=N,1=Y)	private: Student Attends Private University (O=N,1=Y)	GPA3	parentGD: Parent Has Graduate Degree (O=N,1=Y)	private: Student Attends Private University (O=N,1=Y)	_GPA3	Estimate	DF	95% Confidence Limits	
		1			0	1.730	396	1.013	2.956
0NoDegree		0	1YesDegree		0	0.347	396	0.193	0.622
	0public	0		1private	0	0.818	396	0.449	1.491

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
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… the difference between the intercepts per unit GPA in logits is a constant 0.5482, but the difference in probability is not constant (and is not 0). 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...

$$Logit(Apply_i = 1) = \beta_0 + \beta_1(GPA_i - 3) + \beta_2(ParentGD_i) + \beta_3(Private_i) + \beta_4(GPA_i - 3)(ParentGD_i) + \beta_5(GPA_i - 3)(Private_i) + \beta_6(GParentGD_i)(Private_i)$$

```
display as result "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,
estat ic, n(400),
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 for no degree, public
lincom  c.gpa3*1 + c.gpa3#i1.parentgd*1 + c.gpa3#i0.private  // GPA slope for yes degree, public
lincom  c.gpa3*1 + c.gpa3#i0.parentgd*1 + c.gpa3#i1.private  // GPA slope for no degree, private
lincom  c.gpa3*1 + c.gpa3#i1.parentgd*1 + c.gpa3#i1.private  // GPA slope for yes degree, private
```

```
TITLE1 "INTERACTIONS MODEL PREDICTING BINARY DV";
PROC GLIMMIX DATA=work.ologit NOCLPRINT NOITPRINT GRADIENT METHOD=QUAD;
CLASS parentGD private;
MODEL apply2 (DESCENDING) = GPA3|parentGD|private@2
             / SOLUTION LINK=LOGIT DIST=BINARY ODDSRATIO(AT GPA3=0);
* Conditional means for plotting interactions, SLICEDIFF gives simple effects;
LSMEANS parentGD|private / ILINK SLICEDIFF=parentGD SLICEDIFF=private AT(GPA3)=(-1); * At GPA=2;
LSMEANS parentGD|private / ILINK SLICEDIFF=parentGD SLICEDIFF=private AT(GPA3)=( 0); * At GPA=3;
LSMEANS parentGD|private / ILINK SLICEDIFF=parentGD SLICEDIFF=private AT(GPA3)=( 1); * At GPA=4;
* Simple slopes for GPA and slope differences by moderators;
ESTIMATE "GPA Slope for No  Degree Public School"  GPA3 1 GPA3*parentGD  1 0 GPA3*private  1 0;
ESTIMATE "GPA Slope for Yes Degree Public School"  GPA3 1 GPA3*parentGD  0 1 GPA3*private  1 0;
ESTIMATE "GPA Slope for No  Degree Private School" GPA3 1 GPA3*parentGD  1 0 GPA3*private  0 1;
ESTIMATE "GPA Slope for Yes Degree Private School" GPA3 1 GPA3*parentGD  0 1 GPA3*private  0 1;
ESTIMATE "GPA Slope Degree Diff"                          GPA3*parentGD -1 1                 ;
ESTIMATE "GPA Slope School Diff"                                          GPA3*private -1 1;
RUN; TITLE1;
```

SAS Output (condensed 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
```

Parameter Estimates

Effect	parentGD: Parent Has Graduate Degree (0=N,1=Y)	private: Student Attends Private University (0=N,1=Y)	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

parentGD: Parent Has Graduate Degree (0=N,1=Y)	private: Student Attends Private University (0=N,1=Y)	GPA3	parentGD: Parent Has Graduate Degree (0=N,1=Y)	private: Student Attends Private University (0=N,1=Y)	_GPA3	Estimate	DF	95% Confidence Limits	
0NoDegree		0	1YesDegree		0	0.430	393	0.191	0.966
0NoDegree		1	0NoDegree		0	2.572	393	1.108	5.973
1YesDegree		1	1YesDegree		0	1.430	393	0.308	6.637
	0public	0		1private	0	0.531	393	0.228	1.237
	0public	1		0public	0	2.935	393	0.621	13.878
	1private	1		1private	0	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

Label	Estimate	Standard Error	DF	t Value	Pr > |t|
GPA Slope for No Degree for Public School	1.3704	0.7761	393	1.77	0.0782
GPA Slope for Yes Degree for Public School	0.7832	0.9846	393	0.80	0.4268
GPA Slope for No Degree for Private School	0.5193	0.3119	393	1.67	0.0967
GPA Slope for Yes Degree for Private School	-0.06789	0.7630	393	-0.09	0.9291
GPA Slope Degree Diff	-0.5872	0.8038	393	-0.73	0.4655
GPA Slope School Diff	-0.8511	0.8153	393	-1.04	0.2972

Notice that the last two estimates are just the two-way interactions among GPA and degree/private…

parentGD*private Least Squares Means

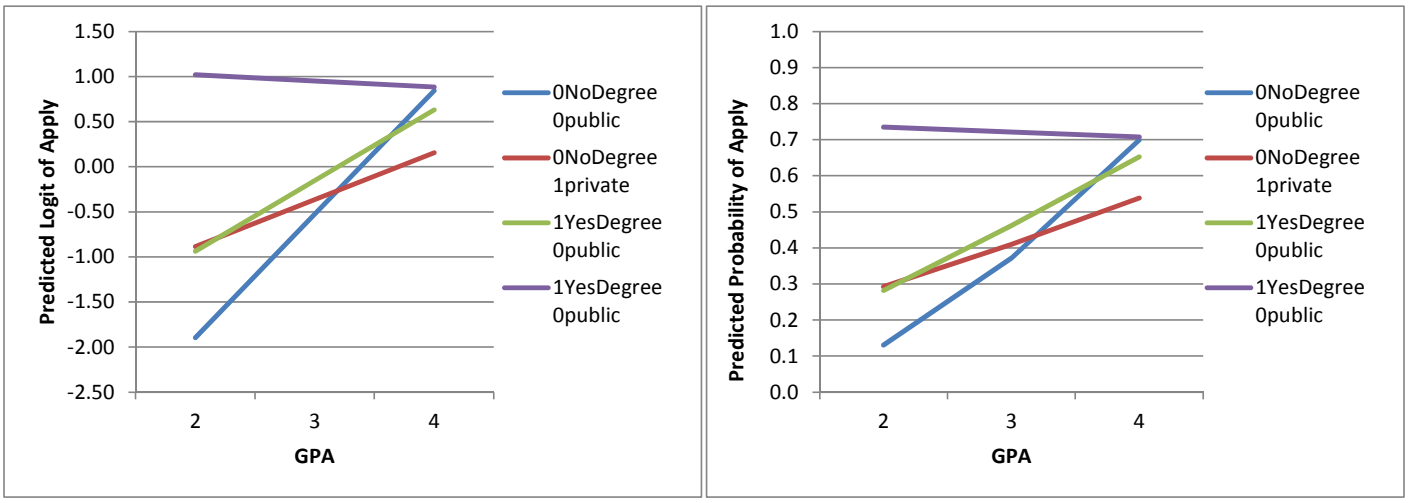
parentGD: Parent Has Graduate Degree (0=N,1=Y)	private: Student Attends Private University (0=N,1=Y)	GPA3	Estimate	Standard Error	DF	t Value	Pr > |t|	Mean	Standard Error Mean
0NoDegree	0public	-1.00	-1.8949	0.9793	393	-1.94	0.0537	0.1307	0.1113
0NoDegree	1private	-1.00	-0.8844	0.3199	393	-2.76	0.0060	0.2923	0.06617
1YesDegree	0public	-1.00	-0.9369	1.4797	393	-0.63	0.5270	0.2815	0.2993
1YesDegree	1private	-1.00	1.0200	0.9071	393	1.12	0.2615	0.7350	0.1767
0NoDegree	0public	0.00	-0.5245	0.3488	393	-1.50	0.1335	0.3718	0.08147
0NoDegree	1private	0.00	-0.3651	0.1206	393	-3.03	0.0026	0.4097	0.02918
1YesDegree	0public	0.00	-0.1537	0.6804	393	-0.23	0.8214	0.4617	0.1691
1YesDegree	1private	0.00	0.9521	0.3270	393	2.91	0.0038	0.7215	0.06569
0NoDegree	0public	1.00	0.8460	0.6993	393	1.21	0.2271	0.6997	0.1469
0NoDegree	1private	1.00	0.1543	0.3483	393	0.44	0.6580	0.5385	0.08655
1YesDegree	0public	1.00	0.6295	0.8218	393	0.77	0.4441	0.6524	0.1864
1YesDegree	1private	1.00	0.8842	0.7452	393	1.19	0.2361	0.7077	0.1542

Simple Effect Comparisons of parentGD*private Least Squares Means By parentGD

Simple Effect Level	private: Student Attends Private University (0=N,1=Y)	private: Student Attends Private University (0=N,1=Y)	GPA3	Estimate	Standard Error	DF	t Value	Pr > |t|
parentGD 0NoDegree	0public	1private	-1.00	-1.0105	1.0108	393	-1.00	0.3181
parentGD 1YesDegree	0public	1private	-1.00	-1.9569	1.2797	393	-1.53	0.1270
parentGD 0NoDegree	0public	1private	0.00	-0.1594	0.3695	393	-0.43	0.6664
parentGD 1YesDegree	0public	1private	0.00	-1.1058	0.7281	393	-1.52	0.1296
parentGD 0NoDegree	0public	1private	1.00	0.6917	0.7622	393	0.91	0.3647
parentGD 1YesDegree	0public	1private	1.00	-0.2547	0.8673	393	-0.29	0.7692

Simple Effect Comparisons of parentGD*private Least Squares Means By private

Simple Effect Level	parentGD: Parent Has Graduate Degree (0=N,1=Y)	parentGD: Parent Has Graduate Degree (0=N,1=Y)	GPA3	Estimate	Standard Error	DF	t Value	Pr > |t|
private 0public	0NoDegree	1YesDegree	-1.00	-0.9580	1.2689	393	-0.75	0.4507
private 1private	0NoDegree	1YesDegree	-1.00	-1.9044	0.9433	393	-2.02	0.0442
private 0public	0NoDegree	1YesDegree	0.00	-0.3708	0.7171	393	-0.52	0.6054
private 1private	0NoDegree	1YesDegree	0.00	-1.3172	0.3488	393	-3.78	0.0002
private 0public	0NoDegree	1YesDegree	1.00	0.2164	0.8428	393	0.26	0.7975
private 1private	0NoDegree	1YesDegree	1.00	-0.7300	0.8035	393	-0.91	0.3642



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 condition. Model-predicted 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.