Example 2: Predicting Count Outcomes in SAS and STATA
(syntax and output available for SAS and STATA electronically)

The real data for this example come from a study of the effects of emotion regulation strategy (none=control, cognitive reappraisal, or suppression) on aggressive behavior in persons with or without a history of intimate partner violence (IPV). The planned analysis was a 2x3 between-groups ANCOVA with factors for strategy condition (3) and IPV (2), with neutral condition aggression as a covariate. Here is the paper published using these data (with similar results as shown here, although their model and sample differed slightly):

Maldonado, R. C., DiLillo, D., & **Hoffman, L.** (2015). Can college students alter their intimate partner aggression-risk behaviors using emotion regulation strategies? An examination using I3 Theory. *Psychology of Violence*, *5*(1), 46-55. Download the paper here

Although not necessary for these single-level, univariate data, when possible in SAS I used GLIMMIX that provides all the necessary distribution fit info AND allows expansion into multilevel or multivariate data. The STATA modules used here are univariate only given that their multilevel/multivariate versions did not provide distribution fit. Note that SAS GLIMMIX uses denominator degrees of freedom (so its Wald test results are given using t or F), whereas each STATA module does not (i.e., they use z or χ^2, respectively).

STATA Syntax for Data Manipulation and Description:

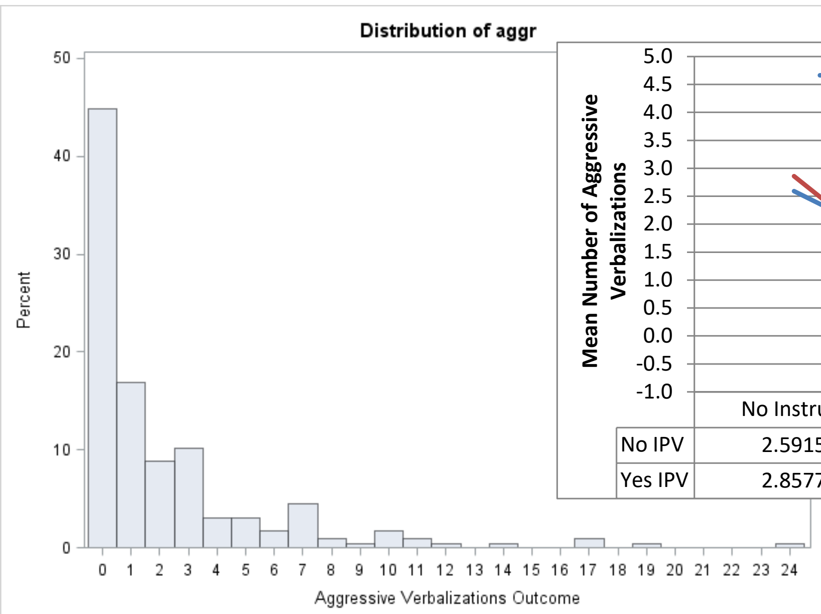
```
* Defining global variable for file location to be replaced in code below
global filesave "C:\Dropbox\20_PSQF7375_Generalized\PSQF7375_Generalized_Example2"

* Import data, transform variables, create and apply value formats
use "$filesave\Example2.dta", clear
label define fercond 1 "1_None" 2 "2_CogR" 3 "3_Supp"
label values ercond fercond

* Save results to separate file
log using $filesave\Example2_STATA_Output.log, replace name(Example2)

display as result "STATA Overall and Cell Means for Aggressive Verbalizations Outcome"
format aggr %8.3f
tabulate ipv ercond, summarize(aggr)

display as result "STATA Histogram for Aggressive Verbalizations Outcome"
hist aggr, percent normal discrete width(1) start(0)
```

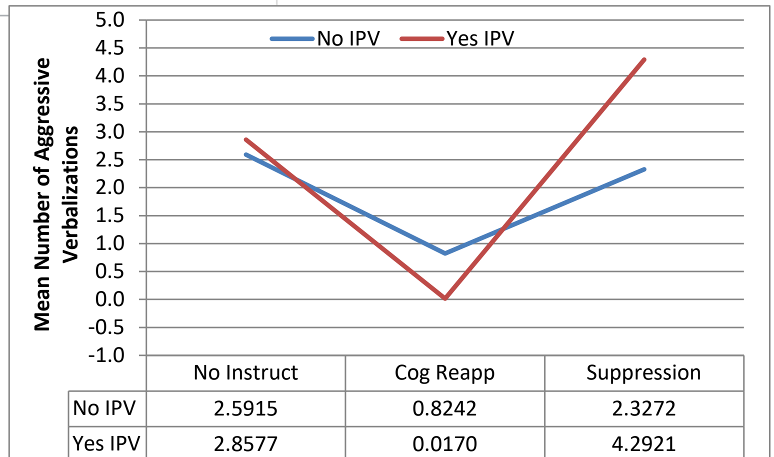
SAS Syntax and Output for Data Manipulation and Description:

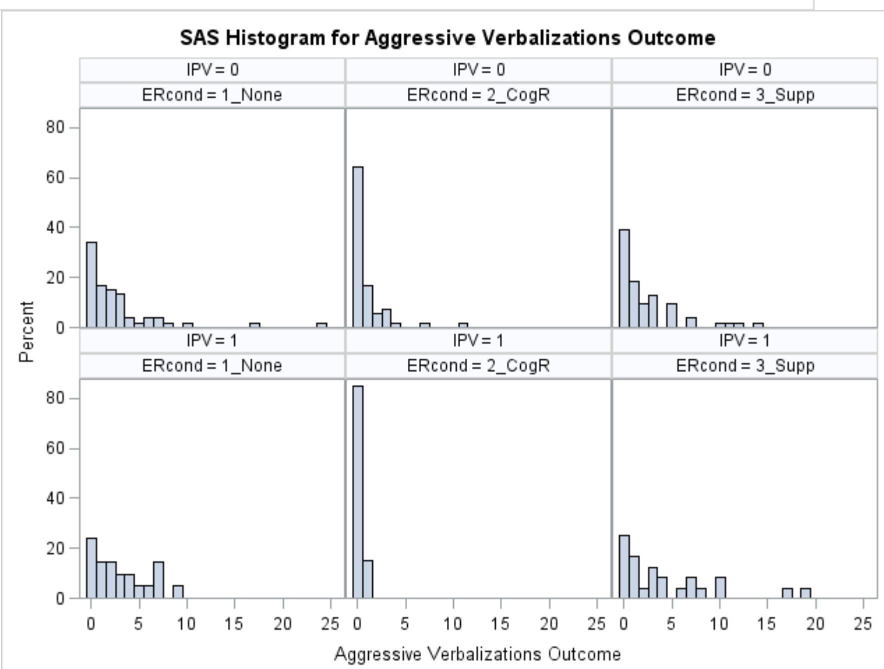
```
* Location for original files for these models – change this path;
%LET filesave= C:\Dropbox\20_PSQF7375_Generalized\PSQF7375_Generalized_Example2;

* Import the STATA version of file into SAS;
PROC IMPORT DATAFILE="&filesave.\Example2.dta" OUT=work.Example2 DBMS=DTA REPLACE; RUN;

* Label variables;
DATA work.Example2; SET work.Example2;
     LABEL IPV=      "Intimate Partner Violence (0=N,1=Y)"
           Neutral=  "Aggressive during Neutral Condition (0=N,1=Y)"
           ERcond=   "Condition (1=None, 2=CogR, 3=Supp)"
           Aggr=     "Aggressive Verbalizations Outcome"; RUN;

* Create labels for groups;
PROC FORMAT; VALUE Fercond 1="1_None" 2="2_CogR" 3="3_Supp"; RUN;

* Add format to data;
DATA work.Example2; SET work.Example2; FORMAT ERcond Fercond.; RUN;
```

```
TITLE "SAS Overall and Cell Means for Aggressive Verbalizations Outcome";
PROC MEANS NONOBS NDEC=3 DATA=work.Example2;
     CLASS IPV ERcond; WAYS 0 2; VAR Aggr; RUN;

* Make histograms to describe the count outcome;
TITLE "SAS Histogram for Aggressive Verbalizations Outcome";
PROC UNIVARIATE NOPRINT DATA=work.example11; VAR Aggr;
     HISTOGRAM Aggr / MIDPOINTS=0 TO 24 BY 1;
RUN; QUIT; TITLE;

TITLE "SAS Group-Specific Histograms for Aggressive Verbalizations Outcome";
PROC SGPANEL DATA=work.Example2;
     PANELBY IPV ERcond / ROWS=2 COLUMNS=3;
     HISTOGRAM Aggr / BINSTART=0 BINWIDTH=1;
RUN; QUIT; TITLE;
```

These are the unadjusted raw means from the data.







Look how non-aggressive our sample is! That's great for them, but not so good if we expect to use a general linear model (i.e., ANCOVA in this case) to analyze this outcome…

However, this is only the marginal distribution of Y. Maybe the residuals will look more normal?

Not likely… (see left)

Model Predicting Aggressive Verbalizations using Normal Distribution and Identity Link (ANCOVA, usually estimated with least squares which is REML, here with ML instead)

$$E(Aggr_i) = \beta_0 + \beta_1(Neutral_i) + \beta_2(IPV_i) + \beta_3(SuppVsNone_i) + \beta_4(SuppVsCog_i) + \beta_5(IPV_i)(SuppVsNone_i) + \beta_6(IPV_i)(SuppVsCog_i) + e_i$$

We are using STATA GLM (generalized linear models) to see how well our residuals fit the normal conditional distribution.

```
display as result "STATA GLM of Aggressive Verbalizations Outcome using ML"
glm aggr c.neutral c.ipv##ib(last).ercond, ///
         link(identity) family(gaussian) // normal residuals
estat ic, n(225),
* Wald Test of Model R2 (not given by default in GLM)
test (c.neutral=0)(c.ipv=0)(i1.ercond=i2.ercond)(i1.ercond=i3.ercond) ///
     (c.ipv#i1.ercond=c.ipv#i2.ercond)(c.ipv#i1.ercond=c.ipv#i3.ercond)
* Wald Test of Omnibus main effect of group and interaction with IPV
contrast  i.ercond,        // DF=2 test of main effect at IPV=0
contrast  c.ipv#i.ercond,  // DF=2 test of interaction
* Cell means of condition per IPV
margins i.ercond, at(c.neutral=0 c.ipv=0) // cell means for IPV=0
margins i.ercond, at(c.neutral=0 c.ipv=1) // cell means for IPV=1
* Simple effects of condition per IPV
lincom i1.ercond*-1 + i2.ercond*1 + c.ipv#i1.ercond*0  + c.ipv#i2.ercond*0 // N/C: IPV=No
lincom i1.ercond*-1 + i3.ercond*1 + c.ipv#i1.ercond*0  + c.ipv#i3.ercond*0 // N/S: IPV=No
lincom i2.ercond*-1 + i3.ercond*1 + c.ipv#i2.ercond*0  + c.ipv#i3.ercond*0 // C/S: IPV=No
lincom i1.ercond*-1 + i2.ercond*1 + c.ipv#i1.ercond*-1 + c.ipv#i2.ercond*1 // N/C: IPV=Yes
lincom i1.ercond*-1 + i3.ercond*1 + c.ipv#i1.ercond*-1 + c.ipv#i3.ercond*1 // N/S: IPV=Yes
lincom i2.ercond*-1 + i3.ercond*1 + c.ipv#i2.ercond*-1 + c.ipv#i3.ercond*1 // C/S: IPV=Yes
* Simple effects of IPV per condition
lincom c.ipv + c.ipv#i1.ercond  // No/Yes IPV Difference: No Instruction
lincom c.ipv + c.ipv#i2.ercond  // No/Yes IPV Difference: Cognitive Reappraisal
lincom c.ipv + c.ipv#i3.ercond  // No/Yes IPV Difference: Suppression
* Specific interaction contrasts
contrast {c.ipv#i.ercond -1  1  0}, pveffects // No/Yes IPV by None/Cog Interaction
contrast {c.ipv#i.ercond -1  0  1}, pveffects // No/Yes IPV by None/Sup Interaction
contrast {c.ipv#i.ercond  0 -1  1}, pveffects // No/Yes IPV by Cog/Sup  Interaction
```

```
TITLE "SAS GLM of Aggressive Verbalizations Outcome using ML";
PROC GLIMMIX DATA=work.Example2 NOCLPRINT GRADIENT NAMELEN=100 METHOD=MSPL;
CLASS ERcond;
MODEL Aggr = Neutral IPV ERcond IPV*ERcond / SOLUTION CHISQ LINK=IDENTITY DIST=NORMAL;
CONTRAST "Wald Test of Model R2" Neutral 1, IPV 1, ERcond -1 1 0, ERcond -1 0 1,
                                 IPV*ERcond -1 1 0, IPV*ERcond -1 0 1 / CHISQ;
* Cell means and simple effects of condition per IPV;
LSMEANS ERcond / ILINK AT (Neutral IPV) = (0 0) CL; * Cell means for IPV=0;
LSMEANS ERcond / ILINK AT (Neutral IPV) = (0 1) CL; * Cell means for IPV=1;
* Simple effects of condition per IPV;
ESTIMATE "None/Cog Difference: IPV=No"   ERcond -1  1  0 IPV*ERcond  0  0  0;
ESTIMATE "None/Sup Difference: IPV=No"   ERcond -1  0  1 IPV*ERcond  0  0  0;
ESTIMATE "Cog/Sup  Difference: IPV=No"   ERcond  0 -1  1 IPV*ERcond  0  0  0;
ESTIMATE "None/Cog Difference: IPV=Yes"  ERcond -1  1  0 IPV*ERcond -1  1  0;
ESTIMATE "None/Sup Difference: IPV=Yes"  ERcond -1  0  1 IPV*ERcond -1  0  1;
ESTIMATE "Cog/Sup  Difference: IPV=Yes"  ERcond  0 -1  1 IPV*ERcond  0 -1  1;
* Simple effects of IPV per condition;
ESTIMATE "No/Yes IPV Difference: No Instruct"      IPV 1 IPV*ERcond  1  0  0;
ESTIMATE "No/Yes IPV Difference: Cog Reapp"        IPV 1 IPV*ERcond  0  1  0;
ESTIMATE "No/Yes IPV Difference: Suppression"      IPV 1 IPV*ERcond  0  0  1;
* Specific interaction contrasts;
ESTIMATE "No/Yes IPV by None/Cog Interaction"            IPV*ERcond -1  1  0;
ESTIMATE "No/Yes IPV by None/Sup Interaction"            IPV*ERcond -1  0  1;
ESTIMATE "No/Yes IPV by Cog/Sup  Interaction"            IPV*ERcond  0 -1  1;
RUN; TITLE;
```

SAS Output:

Fit Statistics

-2 Log Likelihood	1181.71
AIC (smaller is better)	1197.71
AICC (smaller is better)	1198.37
BIC (smaller is better)	1225.04
CAIC (smaller is better)	1233.04
HQIC (smaller is better)	1208.74
Pearson Chi-Square	2515.58
Pearson Chi-Square / DF	**11.18 → This indicates poor fit to the distribution (should be 1)**

Parameter Estimates

Effect	ERcond	Estimate	Standard Error	DF	t Value	Pr > |t|	Gradient	
Intercept		2.3272	0.4553	218	5.11	<.0001	-185E-17	Beta0
neutral		1.3297	0.8306	218	1.60	0.1109	1.67E-16	Beta1
ipv		1.9649	0.8251	218	2.38	0.0181	-197E-17	Beta2
ercond	1_None	0.2643	0.6496	218	0.41	0.6845	5.69E-16	Beta3
ercond	2_CogR	-1.5031	0.6483	218	-2.32	0.0213	1.01E-15	Beta4
ercond	3_Supp	0	.	.	.	.	.	
ipv*ercond	1_None	-1.6988	1.1910	218	-1.43	0.1552	4.16E-16	Beta5
ipv*ercond	2_CogR	-2.7720	1.2031	218	-2.30	0.0222	-189E-18	Beta6
ipv*ercond	3_Supp	0	.	.	.	.	.	
Scale		**11.1803**	**1.0541**	**.**	**.**	**.**	**8.95E-16**	**Resid Var**

Type III Tests of Fixed Effects

Effect	Num DF	Den DF	Chi-Square	F Value	Pr > ChiSq	Pr > F	
Neutral	1	218	2.56	2.56	0.1094	0.1109	
IPV	1	218	0.92	0.92	0.3380	0.3391	marginal over ERcond
ERcond	2	218	8.63	4.31	0.0134	0.0145	At IPV=0
IPV*ERcond	2	218	5.46	2.73	0.0653	0.0676	

p = .0676, seriously?

Contrasts

Label	Num DF	Den DF	Chi-Square	F Value	Pr > ChiSq	Pr > F
Wald Test of Model R2	6	218	31.47	5.24	<.0001	<.0001

ercond Least Squares Means

ERcond	neutral	ipv	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper
1_None	0.00	0.00	2.5915	0.4659	218	5.56	<.0001	0.05	1.6732	3.5098
2_CogR	0.00	0.00	0.8242	0.4636	218	1.78	0.0768	0.05	**-0.08944**	1.7378
3_Supp	0.00	0.00	2.3272	0.4553	218	5.11	<.0001	0.05	1.4299	3.2245
1_None	0.00	1.00	2.8577	0.7392	218	3.87	0.0001	0.05	1.4007	4.3146
2_CogR	0.00	1.00	0.01703	0.7523	218	0.02	0.9820	0.05	**-1.4656**	1.4997
3_Supp	0.00	1.00	4.2921	0.6904	218	6.22	<.0001	0.05	2.9314	5.6528

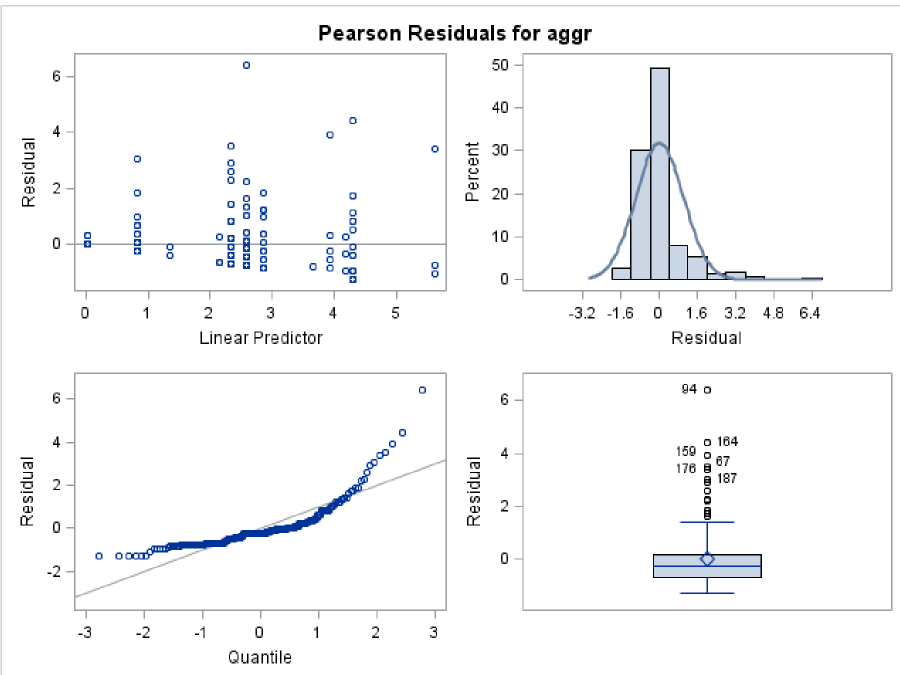
ercond Least Squares Means

ERcond	Mean	Standard Error Mean	Lower Mean	Upper Mean
1_None	2.5915	0.4659	1.6732	3.5098
2_CogR	0.8242	0.4636	-0.08944	1.7378
3_Supp	2.3272	0.4553	1.4299	3.2245
1_None	2.8577	0.7392	1.4007	4.3146
2_CogR	0.01703	0.7523	-1.4656	1.4997
3_Supp	4.2921	0.6904	2.9314	5.6528

Because we are using an identity link, the data-scale “means” are the same as the model-scale estimates.

Estimates

Label	Estimate	Standard Error	DF	t Value	Pr > |t|
None/Cog Difference: IPV=No	-1.7674	0.6497	218	-2.72	0.0071
None/Sup Difference: IPV=No	-0.2643	0.6496	218	-0.41	0.6845
Cog/Sup Difference: IPV=No	1.5031	0.6483	218	2.32	0.0213
None/Cog Difference: IPV=Yes	-2.8406	1.0453	218	-2.72	0.0071
None/Sup Difference: IPV=Yes	1.4345	0.9992	218	1.44	0.1526
Cog/Sup Difference: IPV=Yes	4.2751	1.0126	218	4.22	<.0001
No/Yes IPV Difference: No Instruct	0.2661	0.8631	218	0.31	0.7581
No/Yes IPV Difference: Cog Reapp	-0.8071	0.8777	218	-0.92	0.3588
No/Yes IPV Difference: Suppression	1.9649	0.8251	218	2.38	0.0181
No/Yes IPV by None/Cog Interaction	-1.0733	1.2303	218	-0.87	0.3840
No/Yes IPV by None/Sup Interaction	1.6988	1.1910	218	1.43	0.1552
No/Yes IPV by Cog/Sup Interaction	2.7720	1.2031	218	2.30	0.0222



What about that whole non-normal residuals thing? Yep, still an issue… in addition, the variance appears to grow with the mean.

What to do instead? We get a new model. Let's try using a log link (to keep the predicted counts positive) and a Poisson conditional distribution (where the conditional variance is the same as the conditional mean, which means it is non-constant).

Model Predicting Aggressive Verbalizations using Poisson Distribution and Log Link

$$Log[E(Aggr_i)] = \beta_0 + \beta_1(Neutral_i) + \beta_2(IPV_i) + \beta_3(SuppVsNone_i) + \beta_4(SuppVsCog_i) + \beta_5(IPV_i)(SuppVsNone_i) + \beta_6(IPV_i)(SuppVsCog_i)$$

```
display as result "STATA Poisson Model of Aggressive Verbalizations Outcome"
glm aggr c.neutral c.ipv##ib(last).ercond, ///
         link(log) family(poisson) // Poisson residuals
estat ic, n(225),
* Wald Test of Model R2 (not given by default in GLM)
test (c.neutral=0)(c.ipv=0)(i1.ercond=i2.ercond)(i1.ercond=i3.ercond) ///
     (c.ipv#i1.ercond=c.ipv#i2.ercond)(c.ipv#i1.ercond=c.ipv#i3.ercond)
* Wald Test of Omnibus main effect of group and interaction with IPV
contrast  i.ercond,        // DF=2 test of main effect at IPV=0
contrast  c.ipv#i.ercond,  // DF=2 test of interaction
* Cell means of condition per IPV
margins i.ercond, at(c.neutral=0 c.ipv=0) predict(xb) // log count cell means for IPV=0
margins i.ercond, at(c.neutral=0 c.ipv=1) predict(xb) // log count cell means for IPV=1
margins i.ercond, at(c.neutral=0 c.ipv=0) // count cell means for IPV=0
margins i.ercond, at(c.neutral=0 c.ipv=1) // count cell means for IPV=1
```

We are using STATA GLM to see how well our residuals fit the Poisson conditional distribution.

```
* Simple effects of condition per IPV
lincom i1.ercond*-1 + i2.ercond*1 + c.ipv#i1.ercond*0  + c.ipv#i2.ercond*0 // N/C: IPV=No
lincom i1.ercond*-1 + i3.ercond*1 + c.ipv#i1.ercond*0  + c.ipv#i3.ercond*0 // N/S: IPV=No
lincom i2.ercond*-1 + i3.ercond*1 + c.ipv#i2.ercond*0  + c.ipv#i3.ercond*0 // C/S: IPV=No
lincom i1.ercond*-1 + i2.ercond*1 + c.ipv#i1.ercond*-1 + c.ipv#i2.ercond*1 // N/C: IPV=Yes
lincom i1.ercond*-1 + i3.ercond*1 + c.ipv#i1.ercond*-1 + c.ipv#i3.ercond*1 // N/S: IPV=Yes
lincom i2.ercond*-1 + i3.ercond*1 + c.ipv#i2.ercond*-1 + c.ipv#i3.ercond*1 // C/S: IPV=Yes
* Simple effects of IPV per condition (in log count)
lincom c.ipv + c.ipv#i1.ercond  // No/Yes IPV Difference: No Instruction
lincom c.ipv + c.ipv#i2.ercond  // No/Yes IPV Difference: Cognitive Reappraisal
lincom c.ipv + c.ipv#i3.ercond  // No/Yes IPV Difference: Suppression
* Specific interaction contrasts (in log count)
contrast {c.ipv#i.ercond -1  1  0}, pveffects // No/Yes IPV by None/Cog Interaction
contrast {c.ipv#i.ercond -1  0  1}, pveffects // No/Yes IPV by None/Sup Interaction
contrast {c.ipv#i.ercond  0 -1  1}, pveffects // No/Yes IPV by Cog/Sup Interaction

display as result "STATA Poisson Model of Aggressive Verbalizations Outcome"
display as result "Request Incidence-Rate Ratios (via eform or irr)"
glm aggr c.neutral c.ipv##ib(last).ercond, eform ///
         link(log) family(poisson) // Poisson residuals
* Simple effects of condition per IPV
lincom i1.ercond*-1 + i2.ercond*1 + c.ipv#i1.ercond*0  + c.ipv#i2.ercond*0, irr // N/C: No
lincom i1.ercond*-1 + i3.ercond*1 + c.ipv#i1.ercond*0  + c.ipv#i3.ercond*0, irr // N/S: No
lincom i2.ercond*-1 + i3.ercond*1 + c.ipv#i2.ercond*0  + c.ipv#i3.ercond*0, irr // C/S: No
lincom i1.ercond*-1 + i2.ercond*1 + c.ipv#i1.ercond*-1 + c.ipv#i2.ercond*1, irr // N/C: Yes
lincom i1.ercond*-1 + i3.ercond*1 + c.ipv#i1.ercond*-1 + c.ipv#i3.ercond*1, irr // N/S: Yes
lincom i2.ercond*-1 + i3.ercond*1 + c.ipv#i2.ercond*-1 + c.ipv#i3.ercond*1, irr // C/S: Yes
* Simple effects of IPV per condition (in log count)
lincom c.ipv + c.ipv#i1.ercond, irr  // No/Yes IPV Difference: No Instruction
lincom c.ipv + c.ipv#i2.ercond, irr  // No/Yes IPV Difference: Cognitive Reappraisal
lincom c.ipv + c.ipv#i3.ercond, irr  // No/Yes IPV Difference: Suppression
* Specific interaction contrasts (in log count)
contrast {c.ipv#i.ercond -1  1  0}, pveffects eform // No/Yes IPV by None/Cog Interaction
contrast {c.ipv#i.ercond -1  0  1}, pveffects eform // No/Yes IPV by None/Sup Interaction
contrast {c.ipv#i.ercond  0 -1  1}, pveffects eform // No/Yes IPV by Cog/Sup Interaction
```

Partial STATA Output in Original Log Scale:

```
Generalized linear models                          No. of obs      =        225
Optimization     : ML                              Residual df     =        218
                                                   Scale parameter =          1
Deviance         =  782.8814534                    (1/df) Deviance =   3.591199
Pearson          =  1016.550794                    (1/df) Pearson  =   4.663077 → still not good
Variance function: V(u) = u                        [Poisson]
Link function    : g(u) = ln(u)                    [Log]
                                                   AIC             =    5.15093 → not usual AIC!
Log likelihood   = -572.4796749                    BIC             =  -397.8284 → not usual BIC!
------------------------------------------------------------------------------
        aggr |      Coef.  OIM Std. Err.    z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     neutral |   .4739779   .1365618     3.47   0.001     .2063216    .7416341
         ipv |   .5776772   .1329694     4.34   0.000     .3170621    .8382924
             |
      ercond |
     1_None  |   .0998564   .1227213     0.81   0.416    -.1406729    .3403857
     2_CogR  |  -.9672557   .1685875    -5.74   0.000    -1.297681   -.6368302
             |
ercond#c.ipv |
     1_None  |  -.4902856   .1998456    -2.45   0.014    -.8819758   -.0985954
     2_CogR  |  -2.410448   .6092926    -3.96   0.000    -3.604639   -1.216256
             |
       _cons |   .8440365   .0888265     9.50   0.000     .6699397    1.018133
------------------------------------------------------------------------------
-----------------------------------------------------------------------------
       Model |        Obs  ll(null)  ll(model)      df         AIC        BIC
-------------+---------------------------------------------------------------
           . |        225         .  -572.4797       7    1158.959   1182.872
-----------------------------------------------------------------------------
```

```
TITLE "SAS Poisson Model of Aggressive Verbalizations";
PROC GLIMMIX DATA=work.Example2 NOCLPRINT GRADIENT NAMELEN=100 METHOD=MSPL;
CLASS ERcond;
MODEL Aggr = Neutral IPV ERcond IPV*ERcond / SOLUTION CHISQ LINK=LOG DIST=POISSON;
CONTRAST "Wald Test of Model R2" Neutral 1, IPV 1, ERcond -1 1 0, ERcond -1 0 1,
                                  IPV*ERcond -1 1 0, IPV*ERcond -1 0 1 / CHISQ;
* Cell means and simple effects of condition per IPV;
LSMEANS ERcond / ILINK AT (Neutral IPV) = (0 0) CL; * Cell means for IPV=0;
LSMEANS ERcond / ILINK AT (Neutral IPV) = (0 1) CL; * Cell means for IPV=1;
* Simple effects of ERcond per IPV: EXP --> IRRs;
ESTIMATE "None/Cog Difference: IPV=No"   ERcond -1  1  0 IPV*ERcond  0  0  0 / EXP;
ESTIMATE "None/Sup Difference: IPV=No"   ERcond -1  0  1 IPV*ERcond  0  0  0 / EXP;
ESTIMATE "Cog/Sup  Difference: IPV=No"   ERcond  0 -1  1 IPV*ERcond  0  0  0 / EXP;
ESTIMATE "None/Cog Difference: IPV=Yes"  ERcond -1  1  0 IPV*ERcond -1  1  0 / EXP;
ESTIMATE "None/Sup Difference: IPV=Yes"  ERcond -1  0  1 IPV*ERcond -1  0  1 / EXP;
ESTIMATE "Cog/Sup  Difference: IPV=Yes"  ERcond  0 -1  1 IPV*ERcond  0 -1  1 / EXP;
* Simple effects of IPV per condition: EXP --> IRRs;
ESTIMATE "No/Yes IPV Difference: No Instruct"      IPV 1 IPV*ERcond  1  0  0 / EXP;
ESTIMATE "No/Yes IPV Difference: Cog Reapp"        IPV 1 IPV*ERcond  0  1  0 / EXP;
ESTIMATE "No/Yes IPV Difference: Suppression"      IPV 1 IPV*ERcond  0  0  1 / EXP;
* Specific interaction contrasts: EXP --> IRRs;
ESTIMATE "No/Yes IPV by None/Cog Interaction"            IPV*ERcond -1  1  0 / EXP;
ESTIMATE "No/Yes IPV by None/Sup Interaction"            IPV*ERcond -1  0  1 / EXP;
ESTIMATE "No/Yes IPV by Cog/Sup  Interaction"            IPV*ERcond  0 -1  1 / EXP;
RUN; TITLE;
```

SAS Output:

Fit Statistics

-2 Log Likelihood	1144.96
AIC (smaller is better)	1158.96
AICC (smaller is better)	1159.48
BIC (smaller is better)	1182.87
CAIC (smaller is better)	1189.87
HQIC (smaller is better)	1168.61
Pearson Chi-Square	1016.55
Pearson Chi-Square / DF	**4.66 → better, but still not low enough (1 is the goal)**

Parameter Estimates

Effect	ERcond	Estimate	Standard Error	DF	t Value	Pr > |t|	Gradient
Intercept		0.8440	0.08883	218	9.50	<.0001	6.237E-7
neutral		0.4740	0.1366	218	3.47	0.0006	5.648E-8
ipv		0.5777	0.1330	218	4.34	<.0001	2.813E-7
ercond	1_None	0.09986	0.1227	218	0.81	0.4167	1.127E-7
ercond	2_CogR	-0.9673	0.1686	218	-5.74	<.0001	4.551E-7
ercond	3_Supp	0	.	.	.	.	.
ipv*ercond	1_None	-0.4903	0.1998	218	-2.45	0.0149	2.617E-9
ipv*ercond	2_CogR	-2.4104	0.6093	218	-3.96	0.0001	2.684E-7
ipv*ercond	3_Supp	0	.	.	.	.	.

Type III Tests of Fixed Effects

Effect	Num DF	Den DF	Chi-Square	F Value	Pr > ChiSq	Pr > F
Neutral	1	218	12.05	12.05	0.0005	0.0006
IPV	1	218	3.45	3.45	0.0631	0.0645
ERcond	2	218	43.87	21.94	<.0001	<.0001
IPV*ERcond	2	218	19.29	9.64	<.0001	<.0001

Now that's more like it! ☺

Contrasts

Label	Num DF	Den DF	Chi-Square	F Value	Pr > ChiSq	Pr > F
Wald Test of Model R2	6	218	124.04	20.67	<.0001	<.0001

ercond Least Squares Means

ERcond	neutral	ipv	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper
1_None	0.00	0.00	0.9439	0.08560	218	11.03	<.0001	0.05	0.7752	1.1126
2_CogR	0.00	0.00	-0.1232	0.1437	218	-0.86	0.3922	0.05	-0.4065	0.1601
3_Supp	0.00	0.00	0.8440	0.08883	218	9.50	<.0001	0.05	0.6690	1.0191
1_None	0.00	1.00	1.0313	0.1283	218	8.04	<.0001	0.05	0.7784	1.2841
2_CogR	0.00	1.00	-1.9560	0.5777	218	-3.39	0.0008	0.05	-3.0946	-0.8174
3_Supp	0.00	1.00	1.4217	0.09998	218	14.22	<.0001	0.05	1.2247	1.6188

ercond Least Squares Means

ERcond	Mean	Standard Error Mean	Lower Mean	Upper Mean
1_None	2.5700	0.2200	2.1710	3.0423
2_CogR	0.8841	0.1271	0.6660	1.1736
3_Supp	2.3257	0.2066	1.9522	2.7707
1_None	2.8047	0.3598	2.1781	3.6115
2_CogR	0.1414	0.08170	0.04529	0.4416
3_Supp	4.1442	0.4143	3.4030	5.0468

Because we are using an log link, the data-scale “means” are the actual counts (after exponentiation of the model-scale log count estimates).

Estimates

Label	Estimate	Standard Error	DF	t Value	Pr > |t|	Exponentiated Estimate → IRR
None/Cog Difference: IPV=No	-1.0671	0.1654	218	-6.45	<.0001	0.3440
None/Sup Difference: IPV=No	-0.09986	0.1227	218	-0.81	0.4167	0.9050
Cog/Sup Difference: IPV=No	0.9673	0.1686	218	5.74	<.0001	2.6307
None/Cog Difference: IPV=Yes	-2.9873	0.5908	218	-5.06	<.0001	0.05042
None/Sup Difference: IPV=Yes	0.3904	0.1581	218	2.47	0.0143	1.4776
Cog/Sup Difference: IPV=Yes	3.3777	0.5854	218	5.77	<.0001	29.3034
No/Yes IPV Difference: No Instruct	0.08739	0.1505	218	0.58	0.5621	1.0913
No/Yes IPV Difference: Cog Reapp	-1.8328	0.5948	218	-3.08	0.0023	0.1600
No/Yes IPV Difference: Suppression	0.5777	0.1330	218	4.34	<.0001	1.7819
No/Yes IPV by None/Cog Interaction	-1.9202	0.6135	218	-3.13	0.0020	0.1466
No/Yes IPV by None/Sup Interaction	0.4903	0.1998	218	2.45	0.0149	1.6328
No/Yes IPV by Cog/Sup Interaction	2.4104	0.6093	218	3.96	0.0001	11.1389

The Poisson distribution has only one parameter—the mean, which is supposed to also be the variance. In count data it is often more reasonable to allow the variance to differ from the mean (usually to be greater, known as “over-dispersion”). There are two ways to do this: add a constant amount of additional variance (“NB1”) or to allow the variance to change along with the mean (“NB2”)—let’s try the latter here.

Model Predicting Aggressive Verbalizations using Negative Binomial Distribution and Log Link

```
display as result "STATA Negative Binomial Model of Aggressive Verbalizations Outcome"
nbreg aggr c.neutral c.ipv##ib(last).ercond, // Run this to get LRT of scale factor
```

STATA Output:

```
Negative binomial regression                    Number of obs     =        225
                                                LR chi2(6)        =      44.51
Dispersion     = mean                           Prob > chi2       =     0.0000
Log likelihood = -408.97314                     Pseudo R2         =     0.0516
------------------------------------------------------------------------------
        aggr |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     neutral |   .4492786   .3634238     1.24   0.216    -.2630189    1.161576
         ipv |   .5670645   .3391233     1.67   0.094    -.0976049    1.231734
             |
      ercond |
```

```
     1_None |   .0783514   .2760152     0.28   0.777    -.4626285    .6193313
     2_CogR |  -.9582916   .2963431    -3.23   0.001    -1.539113   -.3774697
            |
erccond#c.ipv |
     1_None |  -.4378567   .4940955    -0.89   0.376    -1.406266    .5305527
     2_CogR |  -2.429861   .7596435    -3.20   0.001    -3.918735   -.9409874
            |
      _cons |   .8512255   .1926943     4.42   0.000     .4735517    1.228899
-------------+----------------------------------------------------------------
    /lnalpha |   .4567591   .1525152                      .1578348    .7556833
-------------+----------------------------------------------------------------
       alpha |   1.578948   .2408136                      1.170973    2.129066 → is "k" dispersion
------------------------------------------------------------------------------
LR test of alpha=0: chibar2(01) = 327.01              Prob >= chibar2 = 0.000 → NegBin is better
```

Negative binomial is better than Poisson, but is it "good enough" yet?

```
display as result "STATA Negative Binomial Model of Aggressive Verbalizations Outcome"
glm aggr c.neutral c.ipv##ib(last).ercond, ///
         link(log) family(nbinomial ml) // Negative binomial residuals
estat ic, n(225),
* Wald Test of Model R2 (not given by default in GLM)
test (c.neutral=0)(c.ipv=0)(i1.ercond=i2.ercond)(i1.ercond=i3.ercond) ///
     (c.ipv#i1.ercond=c.ipv#i2.ercond)(c.ipv#i1.ercond=c.ipv#i3.ercond)
* Wald Test of Omnibus main effect of group and interaction with IPV
contrast  i.ercond,        // DF=2 test of main effect at IPV=0
contrast  c.ipv#i.ercond,  // DF=2 test of interaction
* Cell means of condition per IPV
margins i.ercond, at(c.neutral=0 c.ipv=0) predict(xb) // log count cell means for IPV=0
margins i.ercond, at(c.neutral=0 c.ipv=1) predict(xb) // log count cell means for IPV=1
margins i.ercond, at(c.neutral=0 c.ipv=0) // count cell means for IPV=0
margins i.ercond, at(c.neutral=0 c.ipv=1) // count cell means for IPV=1
* Simple effects of condition per IPV
lincom i1.ercond*-1 + i2.ercond*1 + c.ipv#i1.ercond*0  + c.ipv#i2.ercond*0 // N/C: IPV=No
lincom i1.ercond*-1 + i3.ercond*1 + c.ipv#i1.ercond*0  + c.ipv#i3.ercond*0 // N/S: IPV=No
lincom i2.ercond*-1 + i3.ercond*1 + c.ipv#i2.ercond*0  + c.ipv#i3.ercond*0 // C/S: IPV=No
lincom i1.ercond*-1 + i2.ercond*1 + c.ipv#i1.ercond*-1 + c.ipv#i2.ercond*1 // N/C: IPV=Yes
lincom i1.ercond*-1 + i3.ercond*1 + c.ipv#i1.ercond*-1 + c.ipv#i3.ercond*1 // N/S: IPV=Yes
lincom i2.ercond*-1 + i3.ercond*1 + c.ipv#i2.ercond*-1 + c.ipv#i3.ercond*1 // C/S: IPV=Yes
* Simple effects of IPV per condition (in log count)
lincom c.ipv + c.ipv#i1.ercond  // No/Yes IPV Difference: No Instruction
lincom c.ipv + c.ipv#i2.ercond  // No/Yes IPV Difference: Cognitive Reappraisal
lincom c.ipv + c.ipv#i3.ercond  // No/Yes IPV Difference: Suppression
* Specific interaction contrasts (in log count)
contrast {c.ipv#i.ercond -1  1  0}, pveffects // No/Yes IPV by None/Cog Interaction
contrast {c.ipv#i.ercond -1  0  1}, pveffects // No/Yes IPV by None/Sup Interaction
contrast {c.ipv#i.ercond  0 -1  1}, pveffects // No/Yes IPV by Cog/Sup Interaction

display as result "STATA Negative Binomial Model of Aggressive Verbalizations Outcome"
display as result "Request Incidence-Rate Ratios (eform or irr)"
glm aggr c.neutral c.ipv##ib(last).ercond, eform ///
         link(log) family(nbinomial ml) // Negative binomial residuals
* Simple effects of condition per IPV
lincom i1.ercond*-1 + i2.ercond*1 + c.ipv#i1.ercond*0  + c.ipv#i2.ercond*0, irr // N/C: No
lincom i1.ercond*-1 + i3.ercond*1 + c.ipv#i1.ercond*0  + c.ipv#i3.ercond*0, irr // N/S: No
lincom i2.ercond*-1 + i3.ercond*1 + c.ipv#i2.ercond*0  + c.ipv#i3.ercond*0, irr // C/S: No
lincom i1.ercond*-1 + i2.ercond*1 + c.ipv#i1.ercond*-1 + c.ipv#i2.ercond*1, irr // N/C: Yes
lincom i1.ercond*-1 + i3.ercond*1 + c.ipv#i1.ercond*-1 + c.ipv#i3.ercond*1, irr // N/S: Yes
lincom i2.ercond*-1 + i3.ercond*1 + c.ipv#i2.ercond*-1 + c.ipv#i3.ercond*1, irr // C/S: Yes
* Simple effects of IPV per condition (in log count)
lincom c.ipv + c.ipv#i1.ercond, irr  // No/Yes IPV Difference: No Instruction
lincom c.ipv + c.ipv#i2.ercond, irr  // No/Yes IPV Difference: Cognitive Reappraisal
lincom c.ipv + c.ipv#i3.ercond, irr  // No/Yes IPV Difference: Suppression
* Specific interaction contrasts (in log count)
contrast {c.ipv#i.ercond -1  1  0}, pveffects eform // No/Yes IPV by None/Cog Interaction
contrast {c.ipv#i.ercond -1  0  1}, pveffects eform // No/Yes IPV by None/Sup Interaction
contrast {c.ipv#i.ercond  0 -1  1}, pveffects eform // No/Yes IPV by Cog/Sup Interaction
```

We are using STATA GLM to see how well our residuals fit the negative binomial conditional distribution.

```
TITLE "SAS Negative Binomial Model of Aggressive Verbalizations";
PROC GLIMMIX DATA=work.Example2 NOCLPRINT NAMELEN=100 METHOD=MSPL PLOTS=(PEARSONPANEL);
CLASS ERcond;
MODEL Aggr = Neutral IPV ERcond IPV*ERcond / CHISQ SOLUTION LINK=LOG DIST=NEGBIN;
CONTRAST "Wald Test of Model R2" Neutral 1, IPV 1, ERcond -1 1 0, ERcond -1 0 1,
                                 IPV*ERcond -1 1 0, IPV*ERcond -1 0 1 / CHISQ;
* Cell means and simple effects of condition per IPV;
LSMEANS ERcond / ILINK AT (Neutral IPV) = (0 0) CL; * Cell means for IPV=0;
LSMEANS ERcond / ILINK AT (Neutral IPV) = (0 1) CL; * Cell means for IPV=1;
* Simple effects of ERcond per IPV: EXP --> IRRs;
ESTIMATE "None/Cog Difference: IPV=No"   ERcond -1  1  0 IPV*ERcond  0  0  0 / EXP;
ESTIMATE "None/Sup Difference: IPV=No"   ERcond -1  0  1 IPV*ERcond  0  0  0 / EXP;
ESTIMATE "Cog/Sup  Difference: IPV=No"   ERcond  0 -1  1 IPV*ERcond  0  0  0 / EXP;
ESTIMATE "None/Cog Difference: IPV=Yes"  ERcond -1  1  0 IPV*ERcond -1  1  0 / EXP;
ESTIMATE "None/Sup Difference: IPV=Yes"  ERcond -1  0  1 IPV*ERcond -1  0  1 / EXP;
ESTIMATE "Cog/Sup  Difference: IPV=Yes"  ERcond  0 -1  1 IPV*ERcond  0 -1  1 / EXP;
* Simple effects of IPV per condition: EXP --> IRRs;
ESTIMATE "No/Yes IPV Difference: No Instruct"      IPV 1 IPV*ERcond  1  0  0 / EXP;
ESTIMATE "No/Yes IPV Difference: Cog Reapp"        IPV 1 IPV*ERcond  0  1  0 / EXP;
ESTIMATE "No/Yes IPV Difference: Suppression"      IPV 1 IPV*ERcond  0  0  1 / EXP;
* Specific interaction contrasts: EXP --> IRRs;
ESTIMATE "No/Yes IPV by None/Cog Interaction"            IPV*ERcond -1  1  0 / EXP;
ESTIMATE "No/Yes IPV by None/Sup Interaction"            IPV*ERcond -1  0  1 / EXP;
ESTIMATE "No/Yes IPV by Cog/Sup  Interaction"            IPV*ERcond  0 -1  1 / EXP;
OUTPUT OUT=work.NegBinPred PREDICTED(ILINK)=PredCount; * Save predicted counts to dataset;
RUN; TITLE;
PROC CORR DATA=work.NegBinPred; VAR Aggr; WITH PredCount; RUN; * Get corr of pred & Aggr;
```

SAS Output:

Fit Statistics

-2 Log Likelihood	817.95
AIC (smaller is better)	833.95
AICC (smaller is better)	834.61
BIC (smaller is better)	861.28
CAIC (smaller is better)	869.28
HQIC (smaller is better)	844.98
Pearson Chi-Square	247.87
Pearson Chi-Square / DF	**1.10**

Poisson model −2LL =1144.96
−2ΔLL(df = 1) = 1146.96 – 817.95 = 329.01, p = 1.57911E-73

So the model fits significantly better from adding a "dispersion" (scale) parameter that allows the variance to exceed the mean: $Variance = \mu + (1.5789\mu^2)$ (as shown in STATA GLM output)

The 1.10 means the fit to the distribution is close to the data! ☺

Parameter Estimates

Effect	ERcond	Estimate	Standard Error	DF	t Value	Pr > |t|	Gradient
Intercept		0.8512	0.1927	218	4.42	<.0001	-7.22E-8
neutral		0.4493	0.3634	218	1.24	0.2177	-2.89E-8
ipv		0.5671	0.3391	218	1.67	0.0959	-2.77E-8
ercond	1_None	0.07835	0.2760	218	0.28	0.7768	-4.13E-8
ercond	2_CogR	-0.9583	0.2963	218	-3.23	0.0014	7.168E-9
ercond	3_Supp	0	.	.	.	.	.
ipv*ercond	1_None	-0.4379	0.4941	218	-0.89	0.3765	1.266E-9
ipv*ercond	2_CogR	-2.4299	0.7596	218	-3.20	0.0016	-748E-12
ipv*ercond	3_Supp	0	.	.	.	.	.
Scale		1.5789	0.2408	.	.	.	-4.38E-6 → k dispersion

Type III Tests of Fixed Effects

Effect	Num DF	Den DF	Chi-Square	F Value	Pr > ChiSq	Pr > F
neutral	1	218	1.53	1.53	0.2164	0.2177
ipv	1	218	1.91	1.91	0.1668	0.1682
ercond	2	218	14.53	7.26	0.0007	0.0009
ipv*ercond	2	218	10.24	5.12	0.0060	0.0067 → Still good!

In STATA the scale factor is also given as log(scale), along with an LRT for its significance against 0.

Contrasts

Label	Num DF	Den DF	Chi-Square	F Value	Pr > ChiSq	Pr > F
Wald Test of Model R2	6	218	42.46	7.08	<.0001	<.0001

ercond Least Squares Means (would become panel A of Figure 1 in results)

ERcond	neutral	ipv	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper
1_None	0.00	0.00	0.9296	0.1983	218	4.69	<.0001	0.05	0.5387	1.3204
2_CogR	0.00	0.00	-0.1071	0.2254	218	-0.47	0.6353	0.05	-0.5513	0.3372
3_Supp	0.00	0.00	0.8512	0.1927	218	4.42	<.0001	0.05	0.4714	1.2310
1_None	0.00	1.00	1.0588	0.3043	218	3.48	0.0006	0.05	0.4591	1.6585
2_CogR	0.00	1.00	-1.9699	0.6469	218	-3.05	0.0026	0.05	-3.2448	-0.6949
3_Supp	0.00	1.00	1.4183	0.2796	218	5.07	<.0001	0.05	0.8673	1.9693

ercond Least Squares Means (would become panel B of Figure 1 in results)

ERcond	Mean	Standard Error Mean	Lower Mean	Upper Mean
1_None	2.5334	0.5024	1.7138	3.7450
2_CogR	0.8985	0.2025	0.5762	1.4010
3_Supp	2.3425	0.4514	1.6023	3.4247
1_None	2.8829	0.8772	1.5826	5.2513
2_CogR	0.1395	0.09022	0.03898	0.4991
3_Supp	4.1301	1.1547	2.3804	7.1658

```
Pearson Correlation, N = 225
                 aggr
PredCount    .35771 → R-square=.128
Mu           <.0001

Stata pseudo-Rsquare=.06, btw
```

Estimates (would become Table 1 in results)

Label	Estimate	Standard Error	DF	t Value	Pr > |t|	Exponentiated Estimate
None/Cog Difference: IPV=No	-1.0366	0.2966	218	-3.49	0.0006	0.3546
None/Sup Difference: IPV=No	-0.07835	0.2760	218	-0.28	0.7768	0.9246
Cog/Sup Difference: IPV=No	0.9583	0.2963	218	3.23	0.0014	2.6072
None/Cog Difference: IPV=Yes	-3.0286	0.7113	218	-4.26	<.0001	0.04838
None/Sup Difference: IPV=Yes	0.3595	0.4082	218	0.88	0.3794	1.4326
Cog/Sup Difference: IPV=Yes	3.3882	0.6998	218	4.84	<.0001	29.6112
No/Yes IPV Difference: No Instruct	0.1292	0.3578	218	0.36	0.7184	1.1379
No/Yes IPV Difference: Cog Reapp	-1.8628	0.6830	218	-2.73	0.0069	0.1552
No/Yes IPV Difference: Suppression	0.5671	0.3391	218	1.67	0.0959	1.7631
No/Yes IPV by None/Cog Interaction	-1.9920	0.7716	218	-2.58	0.0105	0.1364
No/Yes IPV by None/Sup Interaction	0.4379	0.4941	218	0.89	0.3765	1.5494
No/Yes IPV by Cog/Sup Interaction	2.4299	0.7596	218	3.20	0.0016	11.3573

Given the large amount of zero values, it will also be helpful to test if the models adequately address them—let's compare with Zero-Inflated Poisson and Zero-Inflated Negative Binomial:

```
display as result "STATA Zero-Inflated Poisson Model of Aggressive Verbalizations Outcome"
zip aggr c.neutral c.ipv##ib(last).ercond, inflate(_cons), // inflate=zero-inflation model
estat ic, n(225)

TITLE "SAS Zero-Inflated Poisson Model of Aggressive Verbalizations";
PROC GENMOD DATA=work.Example2 NAMELEN=100;
CLASS ERcond;
MODEL Aggr = Neutral IPV ERcond IPV*ERcond / LINK=LOG DIST=ZIP;
ZEROMODEL / LINK=LOGIT; RUN; TITLE;
```

In STATA, we have to switch to ZIP and ZINB because GLM does not have zero-inflated (ZI) models. In SAS, we have to switch to GENDMOD because GLIMMIX does not have ZI models.

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance		978.1948	
Scaled Deviance		978.1948	
Pearson Chi-Square	217	449.1678	**2.0699 → Worse than the NB model**
Scaled Pearson X2	217	449.1678	2.0699

```
Log Likelihood                          69.0499
Full Log Likelihood                   -489.0974
AIC (smaller is better)                994.1948 → worse than for NB model
AICC (smaller is better)               994.8615
BIC (smaller is better)               1021.5236 → worse than for NB model

             Analysis Of Maximum Likelihood Zero Inflation Parameter Estimates
                                Standard     Wald 95% Confidence          Wald
Parameter    DF     Estimate       Error           Limits             Chi-Square    Pr > ChiSq
Intercept     1      -0.5461      0.1650     -0.8696      -0.2227          10.95        0.0009
-0.5641 gives the logit of the probability of being an extra 0 if mean = variance in Poisson...
```

```
*** STATA BLEW UP!!
*display as result "Zero-Inflated Negative Binomial of Aggressive Verbalizations Outcome"
*zinb aggr c.neutral c.ipv##ib(last).ercond, inflate(_cons), //inflate=zero-inflation model
*estat ic, n(225)

TITLE "SAS Zero-Inflated Negative Binomial Model of Aggressive Verbalizations";
PROC GENMOD DATA=work.Example2 NAMELEN=100;
CLASS ERcond;
MODEL Aggr = Neutral IPV ERcond IPV*ERcond / LINK=LOG DIST=ZINB;
ZEROMODEL / LINK=LOGIT; RUN; TITLE;
```

```
Algorithm converged. → but wait for it...
             Criteria For Assessing Goodness Of Fit
Criterion                     DF           Value        Value/DF
Deviance                                817.9463
Scaled Deviance                         817.9463
Pearson Chi-Square           217        247.8707          1.1423
Scaled Pearson X2            217        247.8707          1.1423 → Comparable to NB model
Log Likelihood                         -408.9731
Full Log Likelihood                    -408.9731
AIC (smaller is better)                 835.9463 → worse than NB
AICC (smaller is better)                836.7835
BIC (smaller is better)                 866.6912 → worse than NM

             Analysis Of Maximum Likelihood Zero Inflation Parameter Estimates
                                Standard     Wald 95% Confidence          Wald
Parameter    DF     Estimate       Error           Limits             Chi-Square    Pr > ChiSq
Intercept     1     -22.2405    39404.29    -77253.2     77208.75           0.00        0.9995
```

The logit of being an "extra 0" = −22 (crazy SE means no good)! This is probability =.0000000002789 of being an "extra 0". So there are no extra 0's in this distribution not already predicted by the negative binomial.

One last test—does the negative binomial scale parameter need to differ by group? This is called a <u>Heterogeneous Negative Binomial</u> model. It is not available directly in SAS, but it can be programmed in NLMIXED, as I found <u>here</u> by Robin High:

```
display as result "STATA Heterogeneous Negative Binomial Model of Aggressive Verbalization"
gnbreg aggr c.neutral c.ipv##ib(last).ercond, lnalpha(c.ipv##ib(last).ercond),
estat ic, n(225)
```

```
Generalized negative binomial regression        Number of obs     =        225
                                                LR chi2(6)        =      25.52
                                                Prob > chi2       =     0.0003
Log likelihood = -404.92151                     Pseudo R2         =     0.0305
------------------------------------------------------------------------------
        aggr |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
aggr         |
     neutral |   .4321594   .2910113     1.49   0.138    -.1382123    1.002531
         ipv |   .5941631    .329864     1.80   0.072    -.0523585    1.240685
```

```
             |
      ercond |
      1_None |   .1207165   .2837524     0.43   0.671    -.4354279    .6768609
      2_CogR |  -.9690053   .3418598    -2.83   0.005    -1.639038   -.2989724
             |
ercond#c.ipv |
      1_None |  -.5270671   .4493779    -1.17   0.241    -1.407832    .3536974
      2_CogR |  -2.415329   .7186262    -3.36   0.001    -3.823811   -1.006847
             |
       _cons |   .8408582   .1989622     4.23   0.000     .4508995    1.230817
-------------+----------------------------------------------------------------
lnalpha      |
     neutral |  -.6806454   .5386007    -1.26   0.206    -1.736283    .3749927
         ipv |   -.159143   .4744156    -0.34   0.737     -1.08898    .7706945
             |
      ercond |
      1_None |  -.0317307   .4017125    -0.08   0.937    -.8190728    .7556113
      2_CogR |   .6128926   .4748512     1.29   0.197    -.3177987    1.543584
             |
ercond#c.ipv |
      1_None |  -.5459967   .7596402    -0.72   0.472    -2.034864    .9428706
      2_CogR |  -12.29375   45.56926    -0.27   0.787    -101.6079    77.02036 → Uh-oh
             |
       _cons |   .5520306   .2852233     1.94   0.053    -.0069968    1.111058
-----------------------------------------------------------------------------
       Model |        Obs  ll(null)  ll(model)      df         AIC        BIC
-------------+---------------------------------------------------------------
           . |        225 -417.6803  -404.9215      14     837.843   885.6684
-----------------------------------------------------------------------------
```

```
TITLE "SAS Heterogeneous Negative Binomial Model of Aggressive Verbalizations";
PROC NLMIXED DATA=work.Example2 METHOD=GAUSS TECH=QUANEW GCONV=1e-12;
* Must list all parms to be estimated here with start values;
* Bs = fixed effects predicting the mean;
* Ds = fixed effects predicting the dispersion (scale factor);
PARMS B0 0.85 B1neutral  0.46 B2IPV  0.59 B3SvN  0.12 B4SvC -0.97 B5IPVxSvN -0.53 B6IPVxSvC -2.42
      D0 0.54 D1neutral -0.68 D2IPV -0.16 D3SvN -0.30 D4SvC  0.63 D5IPVxSvN -0.55 D6IPVxSvC -12.29;
 * Linear predictor for the mean;
   LogCount = B0 + B1neutral*neutral + B2IPV*IPV + B3SvN*(ERcond=1) + B4SvC*(ERcond=2)
              + B5IPVxSvN*IPV*(ERcond=1) + B6IPVxSvC*IPV*(ERcond=2);
   ExpCount = EXP(LogCount); * Inverse link;
 * Linear predictor for the dispersion scale parameter;
   LogScale = D0 + D1neutral*neutral + D2IPV*IPV + D3SvN*(ERcond=1) + D4SvC*(ERcond=2)
              + D5IPVxSvN*IPV*(ERcond=1) + D6IPVxSvC*IPV*(ERcond=2);
   ExpScale=EXP(LogScale); * Inverse Link;
 * Log-likelihood for NB-1 (constant) or NB-2 (multiplicative);
   * NB-1: LL = Aggr*LOG(ExpScale) - (Aggr+(1/(ExpScale/ExpCount)))*LOG(1+ExpScale)
           + lgamma(Aggr+(1/(ExpScale/ExpCount))) - lgamma(1/(ExpScale/ExpCount)) - lgamma(Aggr+1);
   * NB-2; LL = Aggr*LOG(ExpScale*ExpCount) - (Aggr+(1/ExpScale))*LOG(1+(ExpScale*ExpCount))
           + lgamma(Aggr+(1/ExpScale)) - lgamma(1/ExpScale) - lgamma(Aggr+1);
MODEL Aggr ~ GENERAL(LL);
RUN;
```

NOTE: GCONV convergence criterion satisfied.

Fit Statistics

-2 Log Likelihood	809.8
AIC (smaller is better)	837.8
AICC (smaller is better)	839.8
BIC (smaller is better)	885.7

Parameter Estimates

Parameter	Estimate	Standard Error	DF	t Value	Pr > |t|	95% Confidence Limits		Gradient
B0	0.8409	0.1990	225	4.23	<.0001	0.4488	1.2330	0.000016
B1neutral	0.4322	0.2910	225	1.49	0.1389	-0.1413	1.0057	6.632E-6
B2IPV	0.5941	0.3299	225	1.80	0.0730	-0.05590	1.2441	-0.00001
B3SvN	0.1207	0.2838	225	0.43	0.6709	-0.4385	0.6799	-8.67E-6
B4SvC	-0.9690	0.3419	225	-2.83	0.0050	-1.6427	-0.2953	6.969E-8
B5IPVxSvN	-0.5270	0.4494	225	-1.17	0.2422	-1.4125	0.3585	0.000022
B6IPVxSvC	-2.4158	0.7186	225	-3.36	0.0009	-3.8319	-0.9997	-0.00002

```
D0              0.5520      0.2852   225     1.94    0.0542   -0.01002    1.1141    -6.78E-6
D1neutral      -0.6807      0.5386   225    -1.26    0.2076    -1.7420    0.3807    -0.00002
D2IPV          -0.1591      0.4744   225    -0.34    0.7376    -1.0940    0.7757    -2.65E-6
D3SvN         -0.03173      0.4017   225    -0.08    0.9371    -0.8233    0.7599    -3.97E-6
D4SvC           0.6129      0.4749   225     1.29    0.1981    -0.3228    1.5486     -3.5E-7
D5IPVxSvN      -0.5460      0.7596   225    -0.72    0.4730    -2.0429    0.9509    2.884E-6
D6IPVxSvC     -12.2900      339.44   225    -0.04    0.9711    -681.17    656.59    2.377E-6 Uh-oh!
```

After re-running the model removing the interaction term in predicting the scale factor, none of the effects predicting different scale factors by group are significant and the information criteria are higher (worse), so this indicates the original negative binomial with a constant scale factor is likely to be sufficient.

Sample results section using SAS output:

We examined the extent to which how the count of aggressive verbalizations in the experimental condition differed across three strategy conditions (none, cognitive reappraisal, or suppression) as a function of whether participants had a history of intimate partner violence (IPV; no, yes) while controlling for the main effect of number of aggressive verbalizations in a neutral condition. We estimated generalized linear models using maximum likelihood using SAS GLIMMIX, such that all fixed effects were tested using residual denominator degrees of freedom (i.e., as is the case in traditional linear regression). Effect sizes are provided using incident-rate ratios (IIR), which are exponentiated slope coefficients interpreted similarly to odds ratios, in which IRR values between 0 and 1 indicate negative effects, 1 indicates no effect, and values > 1 indicate positive effects. SAS ESTIMATE and LSMEANS statements were used to request simple slopes and model-predicted outcomes.

Before examining the results, we first examined the fit of the conditional distribution to the model residuals. As expected given the highly skewed observed count distribution, a model specifying an identity link function and normal residuals (i.e., a standard analysis of covariance) did not fit well (Pearson $\chi^2/DF\ = 11.18$) and resulted in confidence intervals for the cell means that included negative count values. An alternative model specifying a log link function and Poisson conditional distribution (in which the conditional mean and variance are the same) appeared to fit the observed distribution better (Pearson $\chi^2/DF\ = 4.66$). However, the conditional variance significantly exceeded the conditional mean, as indicated by a significant likelihood ratio test for a model specifying a negative binomial distribution instead (i.e., that included a scale factor to allow over-dispersion as a quadratic function of the mean, NB2), $-2\Delta LL(1) = 329.01$, $p < .0001$. Adding a zero-inflation parameter did not improve model fit, indicating that the observed 0 values were adequately captured within the negative binomial distribution (Pearson $\chi^2/DF\ = 1.10$). Finally, we examined the potential for group differences in the log of the dispersion scale factor using the same linear predictor as for the log count, but no main effects or interactions were significant, suggesting the original negative binomial with a single scale factor is likely to be sufficient.

The overall model explained a significant amount of variance in aggressive verbalizations, $F(6, 218) = 7.08$, $p <$ .0001, The correlation between the predicted and actual counts was .358 ($R^2 = .128$); dispersion parameter = 1.579. The main effect of aggressive verbalizations during the neutral condition was nonsignificant, indicating that the log count of the outcome increased nonsignificantly by 0.45 for every verbalization in the neutral condition. As expected, there was a significant interaction between strategy condition and history of IPA, $F(2, 218) = 5.12$, $p = .006$. Figure 1 depicts the adjusted cell means for the log counts in panel A, and the expected counts in panel B (holding neutral condition verbalizations constant at 0). Table 1 provides simple slopes and slope differences within the interaction. For persons with or without a history of IPV, aggressive verbalizations were significantly lower when using a cognitive reappraisal strategy than when using no strategy or a suppression strategy. However, these benefits of a cognitive reappraisal strategy effects were significantly stronger in persons with a history of IPV relative to persons without a history of IPV. In addition, while there were no significant IPV group differences when using no strategy or a cognitive reappraisal strategy, the number of aggressive verbalizations was significantly higher in persons with than without a history of IPV when using a suppression strategy. This IPV group difference was significantly larger for participants using a suppression condition than those using a cognitive reappraisal strategy, but not larger for those using no strategy.