Visual Interpretation of interaction effects in Generalized Linear Mixed Model with Proc PLM from a randomized controlled trial

Yingbo Ma, Division of HIV and STD Programs, LA County Department of Public Health, Los Angeles, CA

Kwa Sey, Division of HIV and STD Programs, LA County Department of Public Health, Los Angeles, CA

Abstract:

Visualizing interaction effects may enable us to gain an intuitive understanding of relationships among the variables in the model. The PLM (post-linear modeling) procedure in conjunction with the STORE statement, can restore the model information from SAS/STAT linear modeling procedures (coefficients matrix) and use it to perform advanced postfitting tasks without rerunning the original model. This paper showcases how we have used Proc PLM to depict and explore the interaction between categorical variables and continuous variables in a randomized control trial of a behavioral intervention conducted in Los Angeles among young men who have sex with men. Participants were randomized to either the intervention or control group in 22 clusters with 24 participants per cluster. Condomless anal sex (CAS) within the past 3 months, measured as a binary outcome, was recorded at baseline, and at two follow-ups. A Generalized Linear Mixed Model (GLIMMIX procedure in this example) was used to fit the data. Interaction terms were used to test whether the impact of the Intervention on CAS was different in the intervention and control groups (categorical variable) with time (continuous variable). To visually interpret the interaction, we used the slicefit option with EFFECTPLOT statement in the PLM procedure to display the changes in CAS with time (linear and quadratic function) by separate lines for each level (Intervention vs. Control). The EFFECTPLOT statement and many more statements in PROC PLM have improved our ability to display and interpret significant interactions in the linear mixed model.

The data used in this paper haven't been published . The abstract is submitted and considered only for WUSS conference.

Introduction:

The data used in the abstract was from a randomized control trial to assess the impact of the intervention on condomless anal sex with a man and/or transgender woman. 528 men enrolled in the study. A randomized controlled trial was conducted using a switching-replications study design. Written, informed consent was obtained for study participation and surveys were conducted using audio computer-assisted self-interviews (ACASI). After completion of the baseline survey, participants were randomized to either the intervention or wait-list control group using a 1:1 ratio with a block size of 24, resulting in 22 cohorts. Intervention participants completed the intervention within a month of completing their baseline survey. Wait-list control participants were scheduled to participate in the intervention after completion of their 6 month follow up interview. Below is an example how the data set is constructed:

Study ID	arm	cluster_no	Condomless anal sex	timem
1001	Intervention	1	1	0
1001	Intervention	1	0	4
1002	Control	1	1	0
1022	Control	1	1	0
1022	Control	1	0	3
1023	Control	1	1	0
1023	Control	1	1	3

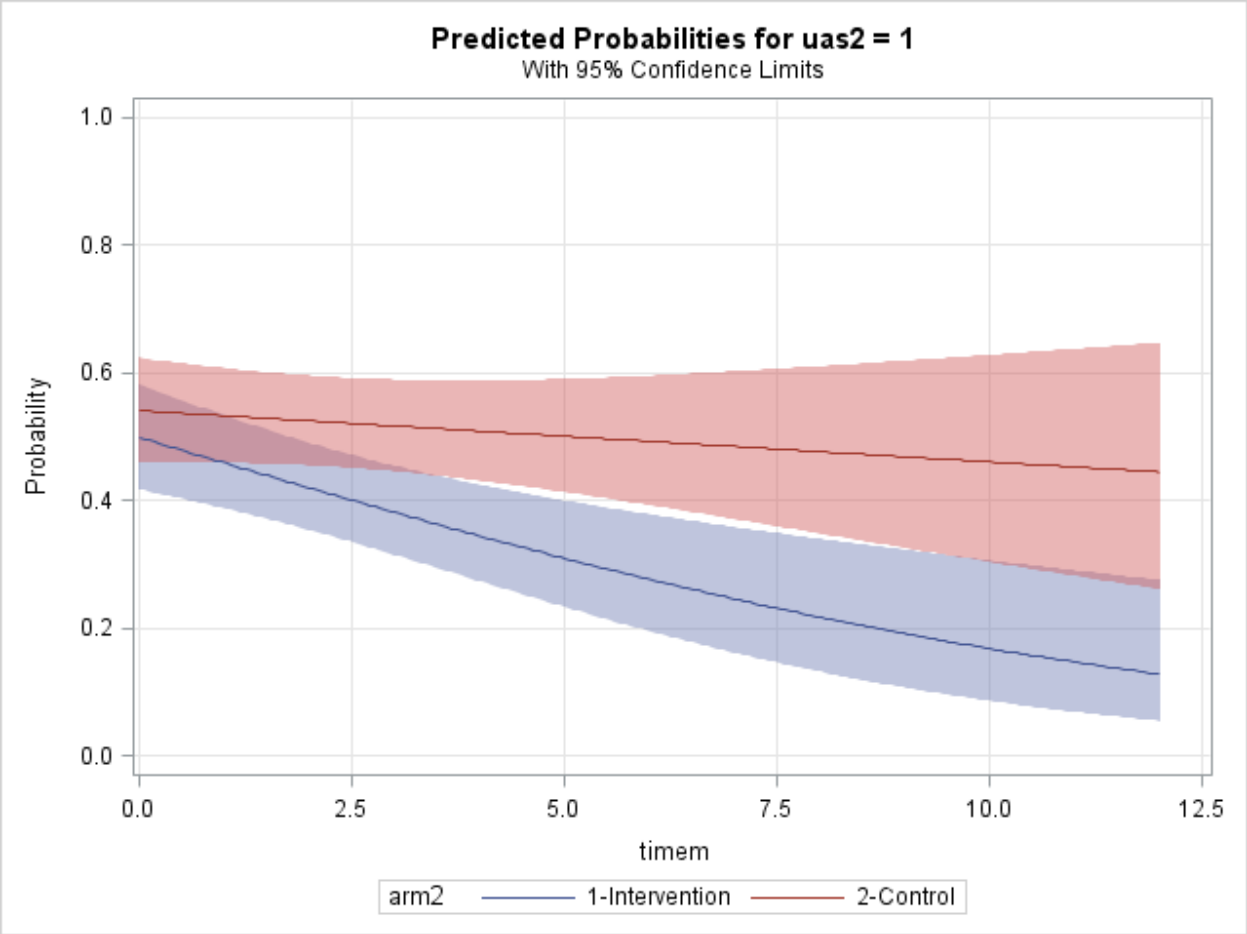
1024	Intervention	1	1	0
1024	Intervention	1	1	3
1025	Control	2	1	0
1025	Control	2	1	4
1026	Intervention	2	1	0
1026	Intervention	2	0	5
1027	Intervention	3	1	0
1027	Intervention	3	1	6
1083	Intervention	4	0	0
1083	Intervention	4	0	3
1084	Intervention	4	1	0
1084	Intervention	4	0	3
1085	Control	4	0	0
1085	Control	4	0	4
1086	Intervention	4	1	0
1086	Intervention	4	0	3
1087	Intervention	5	1	0
1087	Intervention	5	0	4
1088	Control	4	0	0
1088	Control	4	1	3

GLIMMIX procedure fits statistical models to data with correlations or nonconstant variability and where the response is not necessarily normally distributed. The correlations can arise from repeated observation of the same sampling units. In this study, we model binary data with a logit link with categorical and continuous variables. The following code fits a GLMM to the data including baseline and the first followup. Proc PLM with EFFECTPLOT statement is used visualize the interactions. The slicefit option in EFFECTPLOT statement displays the outcome (CAS) against a continuous predictor (Time) on the X-axis, with separate lines by a categorical predictor (Intervention vs. Control). We also estimate the simple slopes in a continuous-by-categorical interaction.

```
title'CAS COMPARING BASELINE TO FIRST FOLLOWUP';
proc glimmix data=test.mlms_all plots=residualpanel(conditional marginal);;
      class id arm cluster_no ;
      model cas (event='1') = arm|timem / solution dist=binary link=logit;
      random cluster_no id;
      store catcont;

run;

*graph simple slopes;
proc plm source=catcont;
effectplot slicefit (x=timem sliceby=arm ) / clm;
run;
```



The slope differences are evident by the change in time for comparisons between Intervention group and Control group (***p*** = 0.03). UAS decreases with time passing by in both intervention and control group. The slope for intervention is steeper.

We also calculate simple slopes and test differences in simple slopes.

```
*calculate and compare the simple slopes(effect)
proc plm restore = catcont;
      estimate 'timem slope, arm=1 intervention' timem 1 timem*arm 1 0,
       'timem slope, arm=2 control' timem 1 timem*arm 0 1 / e;
      estimate 'diff slopes, arm1 vs arm2' timem*arm -1 1/ e;
run;
```

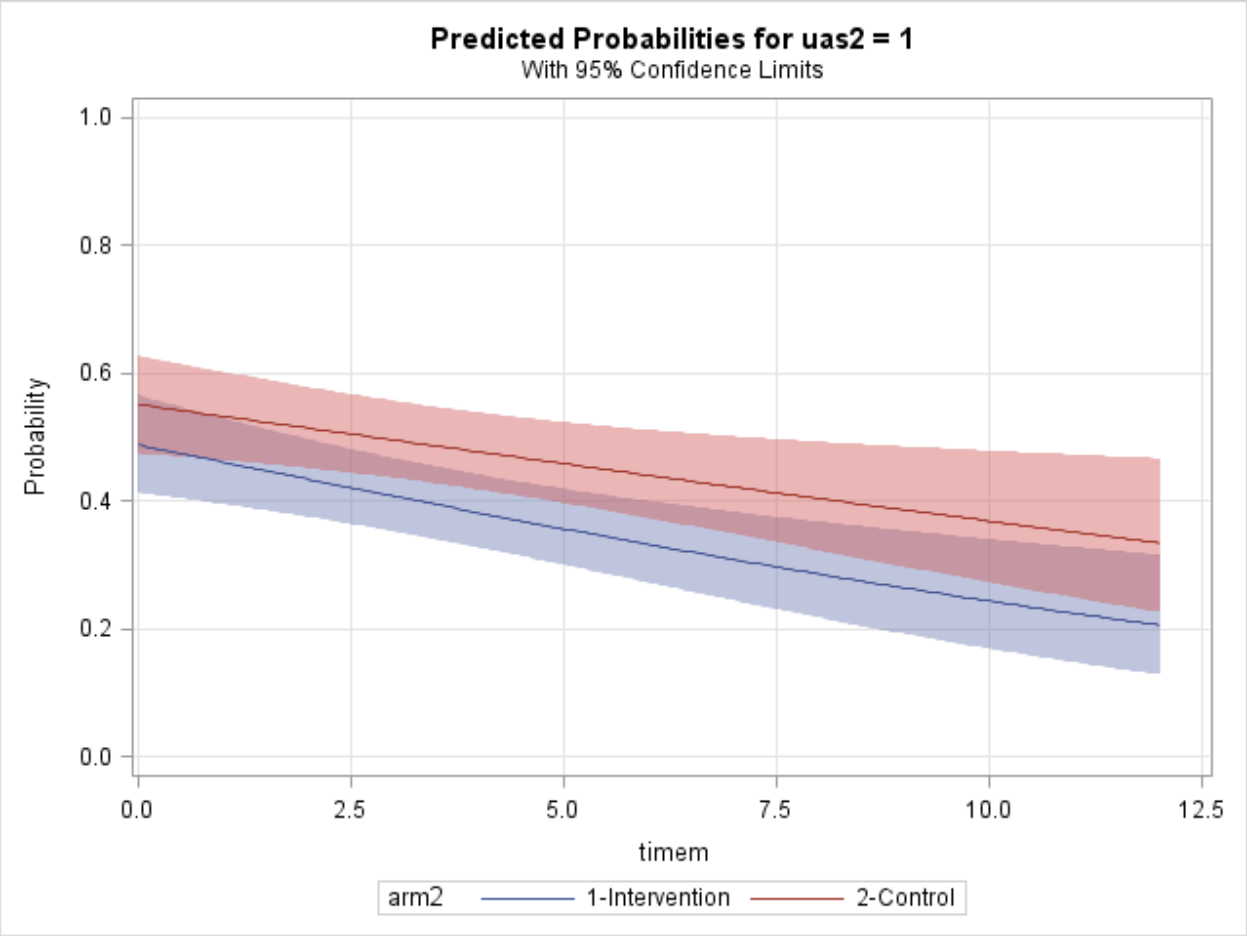
Estimates					
Label	Estimate	Standard Error	DF	t Value	Pr > |t|
Time slope for arm=1 intervention	**-0.1594**	0.04594	385	-3.47	0.0006
Time slope for arm=2 control	**-0.03217**	0.04045	385	-0.8	0.4269

Slopes in Arm1-intervention and arm2-contorl do differ.

Estimate					
Label	Estimate	Standard Error	DF	t Value	Pr > |t|

Diff slopes, arm1 vs arm2	0.1272	0.06114	385	2.08	**0.038**

Howeve, when all follow-up data were included in the model. The interaction was no longer significant.



Quadratic effect : Models a curvilinear relationship between time and UAS .

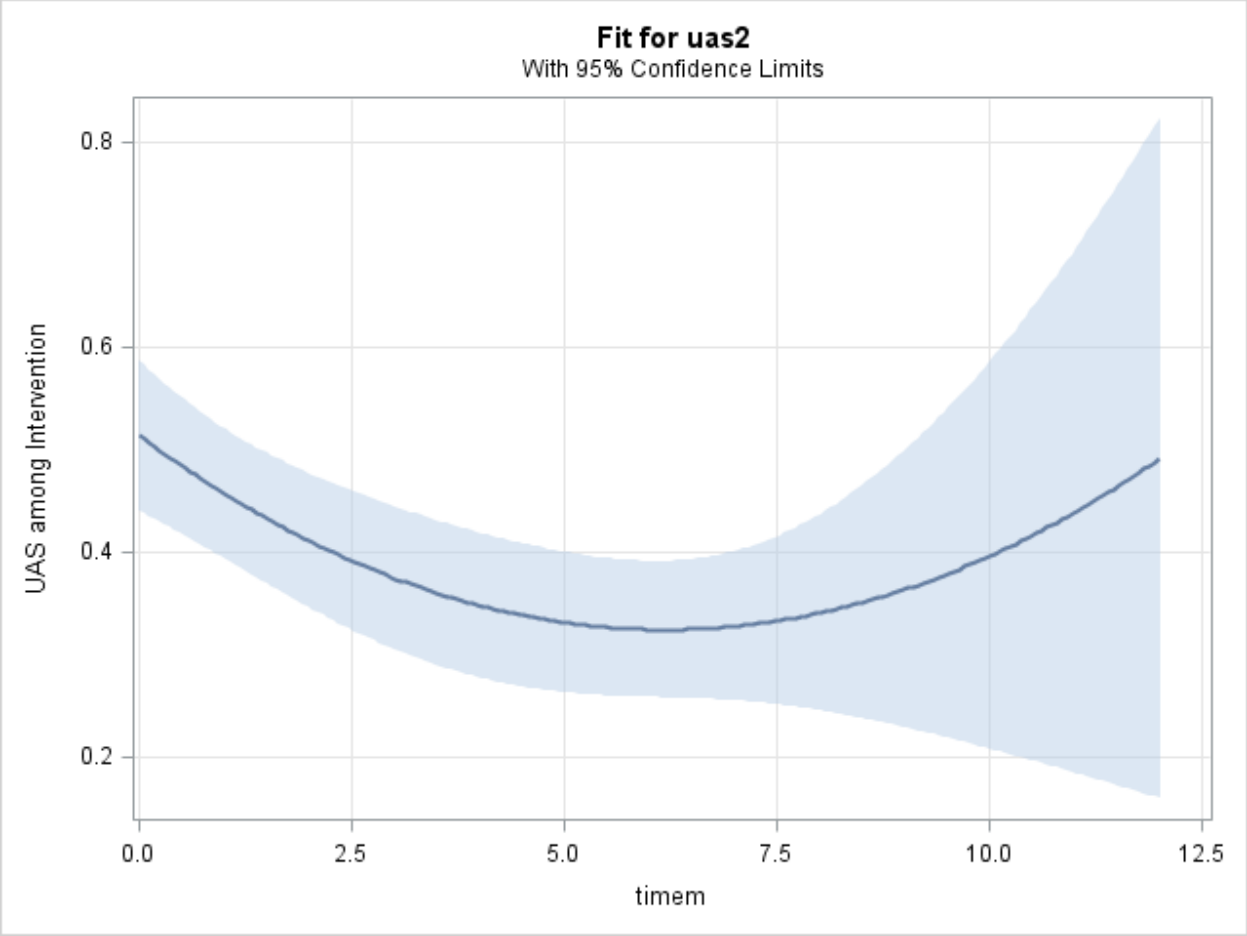
Among Intervention group:

```
/*quadratic effect of time on uas*/
proc glimmix data=test.mlms_all3;
where CAS="Intervention";
   class id  Cluster_no ;
   model CAS (event='1') = timem|timem / solution ;
   random Cluter_no id;
   store quad;
run;
*graph quadratic  slopes;
*The "fit" type of effectplot is made for plotting the outcome vs a single
continuous predictor
proc plm restore=quad;
effectplot fit (x=timem);
```

```
run;
```

- Positive sign indicates that slope becomes more positive as time increases (U-shaped curve) Increasing returns on increasing time in months.

Solutions for Fixed Effects					
Effect	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**
Intercept	0.5133	0.03756	21	13.67	<.0001
timem	-0.06107	0.02006	346	-3.04	0.003
timem*timem	**0.004935**	0.002584	346	1.91	0.057



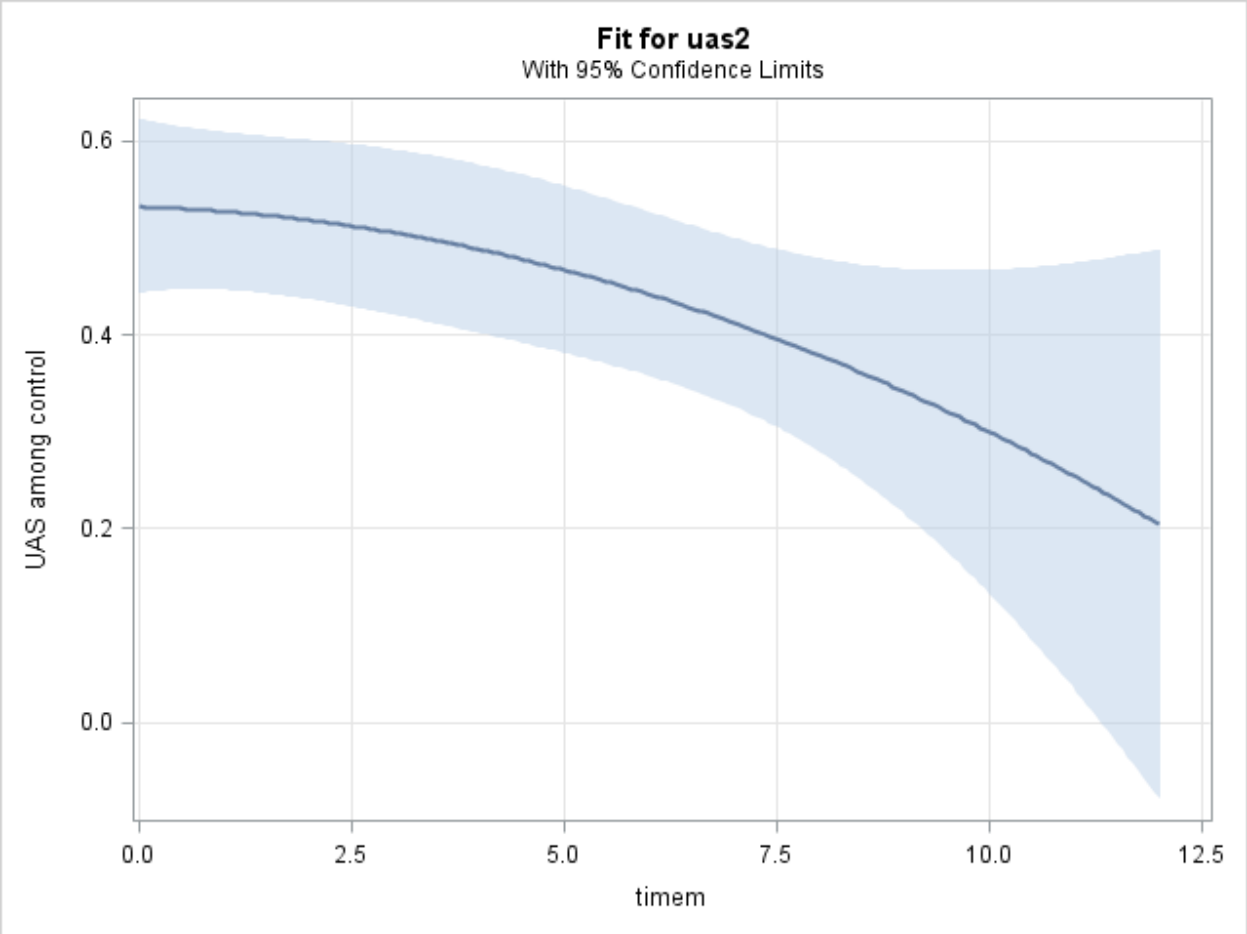
Among Control groups:

```
proc glimmix data=test.mlms_all3;
where CAS='Control';
```

```
   class id  Cluster_no ;
   model CAS (event='1') = timem|timem / solution ;
   random Cluter_no id;
   store quad;
run;
*graph quadratic  slopes;
proc plm restore=quad;
effectplot fit (x=timem);
run;
```

Negative sign indicates that slope becomes more negative as time increases (inverted U-shaped curve).

Solutions for Fixed Effects					
Effect	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**
Intercept	0.5322	0.04586	21	11.6	<.0001
timem	-0.00281	0.01877	325	-0.15	0.881
timem*timem	**-0.00205**	0.002278	325	-0.9	0.37



Include a quadratic time effect in the GLMM:

```
data test.mlms_all2;
set test.mlms_all;
timesq=timem*timem;/*quadratic time*/
label timesq="Time Squared";
run;
proc glimmix data=test.mlms_all2;
     class id arm cluster_no ;
     model CAS (event='1') = arm|timem timesq arm*timesq/ s dist=binary
link=logit chisq;
     random cluster_no id;
     store catcont;
run;

*graph simple slopes;
proc plm source=catcont;
effectplot slicefit (x=timem sliceby=arm) / clm;
run;
```