Interaction Terms Vs. Interaction Effects in Logistic and Probit Regression
CRMportals Inc.,

Background:
In probit or logistic regressions, one can not base statistical inferences based on simply looking at the co-efficient and statistical significance of the interaction terms (Ai et al., 2003).

A basic introduction on what is meant by interaction effect is explained in http://glimo.vub.ac.be/downloads/interaction.htm (What is interaction effect?), and some detailed introduction on interaction is provided in A Primer on Interaction Effects in Multiple Linear Regression (http://www.unc.edu/~preacher/interact/interactions.htm). Interaction effects in CART type model is given in, Correlation and Interaction Effects with Random Forests (http://www.goldenhelix.com/correlation_interaction.html). For interaction effect in designed experiments and specifically in factorial models, see G.E. Box, W.G. Hunter, and J.S. Hunter, Statistics for Experimenters

A nice introduction by Norton and Ai (see references) who did pioneering work on "computational aspects of interaction effects for non-linear models" is http://www.academyhealth.org/2004/ppt/norton2.ppt.

With interaction terms, one has to be very careful when interpreting any of the terms involved in the interaction. This write-up examines the models with interactions and applies Dr. Norton's method to arrive at the size, standard errors and significance of the interaction terms. However, Dr. Norton's program is not able to handle 194,000 observations; it took approximately 11 hours to estimate 75,000 observations for a model with 1 interaction (old_old, endo_vis, old_old*endo_vis) and 1 continuous variable. Therefore, we looked for alternatives using nlcom. This write-up examines comparisons of interest in the presence of interaction terms, using STATA 8.2.

Some tutorials:
The paper is organized as follows:

a. Difference between probability and odds
b. ***logistic*** command in STATA gives odds ratios
c. ***logit*** command in STATA gives estimates
d. difficulties interpreting main effects when the model has interaction terms
e. use of STATA command to get the odds of the combinations of old_old and endocrinologist visits ([1,1], [1,0], [0,1], [0,0])
f. use of these cells to get the odds ratio given in the output and not given in the output
g. use of lincom in STATA to estimate specific cell
h. use of probabilities to do comparisons
i. use of nlcom to estimate risk difference
j. probit regression
k. Interpretation of probit co-efficients
l. Converting probit co-efficients to change in probabilities for easy interpretation

---

i. continuous independent variable ***(***use of function ***normd)*** and for dummy independent variable **(**use of function **norm)**
ii. calculate marginal effects – hand calculation
iii. calculate marginal effects – use of dprobit
iv. calculate marginal effects – use of mfx command
v. calculate marginal effects – use of nlcom

m. Probit regression with interaction effects (for 10,000 observations)
   i. Calculate interaction effect using nlcom
   ii. Using Dr.Norton's ineff program

n. Logistic regression
   i. calculate marginal effects – hand calculation
   ii. calcualte marginal effects – use of mfx command
   iii. calculate effect using nlcom
   iv. calculate interaction effect using nlcom – using Dr. Norton's method

Odds versus probability:

Odds: The ratio of the probability of a patient catching flu to the probability not catching the flu.

For example, if the odds of having allergy this season are 20:1 (read "twenty to one"). The sizes of the numbers on either side of the colon represent the relative chances of not catching flu (on the left) and catching flu (on the right). In other words, what you are told is that the chance of not catching flu is 20 times as great as the chance of having allergy.

Note that odds of 10:1 are not the same as a probability of 1/10.

If an event has a probability of 1/10, then the probability of the event not happening is 9/10. So the chance of the event not happening is nine times as great as the chance of the event happening; the odds are 9:1.

Probability: Probability is the expected number of flu patients divided by the total number of patients.

<u>Relationship:</u>

Odds = probability divided by (1 – probability). $= \frac{\Pr obabilty}{1 - probability}$

Example:
If an event has a probability of 1/10, then the probability of the event not happening is 9/10. So the chance of the event not happening is nine times as great as the chance of the event happening; the odds are 9:1.

Probability = odds divided by (1 + odds) = $\frac{odds}{1+odds}$

Example:
If the odds are 10:1 then the probability = 1/11

In this case we assume that there are 11 likely outcomes and events not happening is 10 and event happening is 1. So the probability of the even happening = 1 / 11.

Simple Model:

$$\log it(p) = \beta_0 + \beta_1\, old\_old \quad or \ln\left[\frac{\hat{p}}{1-\hat{p}}\right] = \beta_0 + \beta_1\, old\_old$$

```
. logistic a1c_test old_old

Logistic regression                               Number of obs   =     194772
                                                  LR chi2(1)      =      17.10
                                                  Prob > chi2     =     0.0000
Log likelihood =  -117729.9                       Pseudo R2       =     0.0001

------------------------------------------------------------------------------
    a1c_test | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     old_old |   .9585854   .0097972    -4.14   0.000     .9395742    .9779813
------------------------------------------------------------------------------
```

Std. Err for odds ratios is not meaningful.

```
. logit

Logit estimates                                   Number of obs   =     194772
                                                  LR chi2(1)      =      17.10
                                                  Prob > chi2     =     0.0000
Log likelihood =  -117729.9                       Pseudo R2       =     0.0001

------------------------------------------------------------------------------
    a1c_test |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     old_old |  -.0422966   .0102205    -4.14   0.000    -.0623285   -.0222648
       _cons |   .8989483   .0063666   141.20   0.000       .88647    .9114266
------------------------------------------------------------------------------
```

When old_old = 1, the risk of A1c test is

$$\log it(\hat{p}_1) = \beta_0 + \beta_1$$

When old_old = 0 the risk of A1c test is

$$\log it(\hat{p}_0) = \beta_0$$

Take the difference:

$$\log it(\hat{p}_1) - \log it(\hat{p}_0) = ([\beta_0 + \beta_1] - \beta_0) = \beta_1$$

Odds ratio:

$$\ln\left[\frac{\hat{p}_1/(1-\hat{p}_1)}{\hat{p}_0/(1-\hat{p}_0)}\right] = \ln(OR) = \beta_1$$

Model with interaction

Let us fit the following model with interaction:

$$\log it(p) = \beta_0 + \beta_1\, old\_old + \beta_2\, endo\_vis + \beta_3\, old\_old * endo\_vis \;\; (Interaction)$$

$$\ln\left[\frac{p}{1-p}\right] = \beta_0 + \beta_1 old\_old + \beta_2 endo\_vis + \beta_3 old\_old * endo\_vis$$

Given below are the odds ratios produced by the logistic regression in STATA. Now we can see that one can not look at the interaction term alone and interpret the results.

```
logistic a1c_test old_old endo_vis oldXendo

Logistic regression                               Number of obs   =     194772
                                                  LR chi2(3)      =    1506.73
                                                  Prob > chi2     =     0.0000
Log likelihood = -116985.08                       Pseudo R2       =     0.0064

------------------------------------------------------------------------------
    a1c_test | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     old_old |   .9611249   .0106487    -3.58   0.000     .9404788    .9822243
    endo_vis |   1.651284    .028952    28.61   0.000     1.595503    1.709015
    oldXendo |   1.067382   .0314229     2.22   0.027     1.007538    1.130781
------------------------------------------------------------------------------
```

With interaction terms, one has to be very careful when interpreting any of the terms involved in the interaction. For example, in the above model "endo_vis" can not be interpreted as the overall comparison of endocrinologist visit to "no endocrinologist visit," because this term is part of an interaction. It is the effect of endocrinologist visit when the "other" terms in the interaction term are at the reference values (ie. when old_old = 0). Similarly, the "old_old" cannot be interpreted as the overall comparison of "old_old" to "young-old". It is the effect of "old-old" when "other" terms in the interaction term is at the reference value (ie. endo_vis = 0).

To help in the interpretation of the odds ratios, let's obtain the odds of receiving an A1c-test for each of the 4 cells formed by this 2 x 2 design using the **adjust** command.

```
. adjust, by (old_old endo_vis) exp
```

```
          | Endocronologist
Age >= 75 |       0        1
----------+-----------------
        0 | 2.25011  3.71557
        1 | 2.16264  3.81176
----------------------------
```

1) The odds ratio for "old_old" represents the odds ratio of old_old when there is no endocrinologist visit is = 0.9611. (Note: The odds ratio for the old_old, when endocrinologist visit = 0 can be read directly from the output which is 0.9611 (0.94, 0.98) because the interaction term and endocrinologist visit drop out). Interpretation: When there is no endocrinologist visit, the odds of a **old_old** having an A1c test is .96 times that of an young_old.

```
. display 2.16264/2.25011
.96112
```

2) the odds ratio "endo_vis" is the odds ratio formed by comparing an endocrinologist to no endocrinologist visit for young_old (because this is the reference group for old_old). (Note: The odds ratio for the endocrinologist, old_old = 0 can be read directly from the output which is 1.65 (1.60, 1.71) because the interaction term and endocrinologist visit drop out).

```
. display 3.71557/2.25011
1.65128
```

3) the odds ratio old_old seeing an endocrinologist compared to an young-old seeing an endocrinologist (not given in the logistic estimates)

```
. display 3.81176/3.71557
1.02588
```

Using logit estimates to do comparisons:

```
Logit estimates                                   Number of obs   =     194772
                                                  LR chi2(3)      =    1506.73
                                                  Prob > chi2     =     0.0000
Log likelihood = -116985.08                       Pseudo R2       =     0.0064

------------------------------------------------------------------------------
    a1c_test |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     old_old |  -.0396509   .0110794    -3.58   0.000    -.0613662   -.0179356
    endo_vis |    .501553    .017533    28.61   0.000     .4671888    .5359171
    oldXendo |   .0652091   .0294392     2.22   0.027     .0075093    .1229089
       _cons |   .8109787   .0069608   116.51   0.000     .7973358    .8246216
------------------------------------------------------------------------------
```

---

a) risk of A1c test with old_old =1 given endocrinologist visit =1

$$\log it(p_1) = \beta_0 + \beta_1 + \beta_2 + \beta_3$$

(b) risk of A1c test with old_old =0 given endocrinologist visit =1

$$\log it(p_0) = \beta_0 + \beta_2$$

The terms (β_1, β_3) are gone because old_old = 0 and the interaction term becomes zero.

Then take the differences:

$$\log it(p_1) - \log it(p_o) = [\beta_0 + \beta_1 + \beta_2 + \beta_3] - [\beta_0 + \beta_2]$$
$$\log it(p_1) - \log it(p_o) = \beta_1 + \beta_3$$

If we represent logit as ln (p/1-p) then

$$\ln\left[\frac{p_1}{1-p_1}\right] - \ln\left[\frac{p_0}{1-p_0}\right] = [\beta_0 + \beta_1 + \beta_2 + \beta_3] - [\beta_0 + \beta_2] = \beta_1 + \beta_3$$

These are the co-efficients for "old_old" and "old_old*endo_vis"
exp($\beta_1 + \beta_3$) = odds ratio = exp (-.0396509 + .0652091) = 1.0258876 .

```
. display exp(-.0396509 + .0652091)
1.0258876
```

Use of lincom:

One can use STATA's commands to produce this: Variance is calculated by lincom using matrix algebra.

```
. lincom old_old + oldXendo, or

 ( 1)  old_old + oldXendo = 0

------------------------------------------------------------------------------
    a1c_test | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         (1) |   1.025888   .0279809     0.94   0.349     .9724863    1.082221
------------------------------------------------------------------------------
```

We can use the following table of ln odds for the cross classification of old_old and endo_vis

	Endo_vis = 1	Endo_vis = 0
Old_old = 1	$\beta_0 + \beta_1 + \beta_2 + \beta_3$	$\beta_0 + \beta_1$

Old_old = 0	$\beta_0 + \beta_2$	β_0

For example, the odds of A1c test among old_old and with endo_vis = 0 is: exp (β0 + β1)

Results Summary in terms of odds ratios:

a) The association between HbA1c test and old_old = 0.9611 among those not seeing an endocrinologist
b) The association between HbA1c test and old_old = 1.0258 among those seeing an endocrinologist

Presenting estimates – Predicted Probabilities

As stated earlier, with interaction terms, co-efficients of variables that are involved in interactions do not have a straightforward interpretation. One way to interpret these models with interactions may be through predicted probabilities. If we write out the non-linear combinations of interest, STATA's nlcom will produce the point estimates and confidence intervals.

Comparisons with Probabilities:
Use the simple relationship between odds and risk.
If Odds = $\left[\frac{p}{1-p}\right]$ then p = $\left[\frac{odds}{1+odds}\right]$

Estimate change in probability of receiving A1c test for old_old when endocrinologist visit = 0:

$$\left[\frac{\exp(\beta_0+\beta_1)}{1+(\exp(\beta_0+\beta_1))}\right]$$

exp ($\beta_0 + \beta_1$) = 2.1626

```
. display exp(.8109787+(-.0396509))
2.1626359
```

1+ (exp (($\beta_0 + \beta_1$)

```
. display 1 + (exp(.8109787+(-.0396509)))
3.1626359
```

Numerator/Denominator:

```
display 2.1626359/3.1626359
.68380805
```

In the same way estimate change in probability receiving A1c test for old_old when endocrinologist visit = 1:

Exp ($\beta_0 + \beta_1 + \beta_2 + \beta_3$)

```
. display exp(.8109787+( -.0396509) +  .501553 + .0652091)
3.8117557
```

1+ exp (($\beta_0 + \beta_1 + \beta_2 + \beta_3$)

```
. display 1 + (exp(.8109787+( -.0396509) +  .501553 + .0652091))
4.8117557
```

Numerator/Denominator:

```
. display 3.8117557/4.8117557
.79217565
```

Using nlcom – risk difference

```
. logit a1c_test old_old

Iteration 0:   log likelihood = -117738.45
Iteration 1:   log likelihood =  -117729.9
Iteration 2:   log likelihood =  -117729.9

Logit estimates                                   Number of obs   =     194772
                                                  LR chi2(1)      =      17.10
                                                  Prob > chi2     =     0.0000
Log likelihood =  -117729.9                       Pseudo R2       =     0.0001

------------------------------------------------------------------------------
    a1c_test |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     old_old |  -.0422966   .0102205    -4.14   0.000    -.0623285   -.0222648
       _cons |   .8989483   .0063666   141.20   0.000       .88647    .9114266
------------------------------------------------------------------------------
```

$$p_1 - p_0 = \frac{1}{1+\exp(-\beta_0 - \beta_1)} - \frac{1}{1+\exp(-\beta_0)}$$

```
. nlcom 1/(1+exp(-_b[old_old] - _b[_cons])) - 1/(1+exp(- _b[_cons] ))

       _nl_1:  1/(1+exp(-_b[old_old] - _b[_cons])) - 1/(1+exp(- _b[_cons] ))

------------------------------------------------------------------------------
    a1c_test |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _nl_1 |  -.0087727    .002124    -4.13   0.000    -.0129356   -.0046098
------------------------------------------------------------------------------
```

```
cs a1c_test old_old

                 | Age >= 75              |
                 |   Exposed   Unexposed  |      Total
-----------------+------------------------+----------
           Cases |     52487       85288  |    137775
        Noncases |     22285       34712  |     56997
-----------------+------------------------+----------
           Total |     74772      120000  |    194772
```

```
                 |                        |
            Risk |  .7019606    .7107333  |  .7073655
                 |                        |
                 |      Point estimate    |  [95% Conf. Interval]
                 |------------------------+----------------------
 Risk difference |        -.0087727       | -.0129356   -.0046098
      Risk ratio |         .9876568       |  .9818441    .9935039
 Prev. frac. ex. |         .0123432       |  .0064961    .0181559
 Prev. frac. pop |         .0047385       |
                 +-----------------------------------------------
                               chi2(1) =    17.13  Pr>chi2 = 0.0000
```

It is probably useful to tabulate results as follows and then calculate predicted probabilities rather than odds.

	Old_old	Endo_vis	Cardio_vis	OldoldXendo	OldoldXCardio	Log-likelihood
1	X					
2	X	X				
3	X	X	X			
4	X	X	X	X		
6	X	X	X	X	X	

PROBIT REGRESSION

Probit Coefficients – Continuous variable (dxg):

```
. probit a1c_test dxg

Iteration 0:   log likelihood = -117738.45
Iteration 1:   log likelihood = -117737.67
Iteration 2:   log likelihood = -117737.67

Probit estimates                                  Number of obs   =     194772
                                                  LR chi2(1)      =       1.56
                                                  Prob > chi2     =     0.2120
Log likelihood = -117737.67                       Pseudo R2       =     0.0000

------------------------------------------------------------------------------
    a1c_test |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         dxg |  -.0017647   .0014136    -1.25   0.212    -.0045353    .0010059
       _cons |   .5486867   .0038349   143.08   0.000     .5411706    .5562029
------------------------------------------------------------------------------
```

Interpretation: The co-efficient for dxg (-.0017647) represents the effect of an infinitesimal change in **x** on the standardized probit index. If **dxg** is changed by an infinitesimal (or small) amount, the standardized probit index decreases, on average, by 0.001 of a standard deviation

Marginal Effects:

$$\frac{\partial \text{Prob}(y_i = 1)}{\partial x_k} = \frac{\partial \Phi}{\partial x_k} = \phi(\mathbf{x}_i' \boldsymbol{\beta}) \times \beta_k$$

where $\phi(\cdot)$ denotes the probability density function for the standard normal. The probability density function gives the height of the curve at the relevant index value $\mathbf{x}_i' \boldsymbol{\beta}$.

What is the effect of a small change in dxg on the probability of A1c test?

a) Get mean of dxg

```
. sum dxg
    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
         dxg |    194772    1.687711    2.108394       .068     25.829
```

b) Evaluate mean standardized probit index at this mean

```
. display .5486867 + (-.0017647)*1.687711
.5457084
```

c) Find the height of the standardized normal curve at this point using the pdf table entries and use this to translate the probit coefficient into a probability effect

```
. display normd(.5457084)*-.0017647
-.00060662
```

So marginal effect of dxg = -.0006 ~= -.001; This implies that an infinitesimally small change in x ***decreases*** the probability of receiving hba1c test by **0.1%** at the average.

Check your hand calculation by dprobit (canned routine in STATA)

```
. dprobit a1c_test dxg

Iteration 0:   log likelihood = -117738.45
Iteration 1:   log likelihood = -117737.67
Iteration 2:   log likelihood = -117737.67

Probit estimates                                  Number of obs = 194772
                                                  LR chi2(1)    =   1.56
                                                  Prob > chi2   = 0.2120
Log likelihood = -117737.67                       Pseudo R2     = 0.0000

------------------------------------------------------------------------------
a1c_test |      dF/dx   Std. Err.      z    P>|z|     x-bar  [    95% C.I.   ]
---------+--------------------------------------------------------------------
     dxg |  -.0006066   .0004859    -1.25   0.212   1.68771  -.001559  .000346
---------+--------------------------------------------------------------------
  obs. P |   .7073655
 pred. P |   .7073668  (at x-bar)
------------------------------------------------------------------------------
    z and P>|z| are the test of the underlying coefficient being 0
```

use nlcom

```
. probit a1c_test dxg

Iteration 0:   log likelihood = -117738.45
Iteration 1:   log likelihood = -117737.67
Iteration 2:   log likelihood = -117737.67

Probit estimates                                  Number of obs   =     194772
                                                  LR chi2(1)      =       1.56
                                                  Prob > chi2     =     0.2120
Log likelihood = -117737.67                       Pseudo R2       =     0.0000

------------------------------------------------------------------------------
    a1c_test |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         dxg |  -.0017647   .0014136    -1.25   0.212    -.0045353    .0010059
       _cons |   .5486867   .0038349   143.08   0.000     .5411706    .5562029
------------------------------------------------------------------------------

. quietly sum dxg
. local dxgmean = r(mean)
. local xb _b[dxg]*`dxgmean'+_b[_cons]
. nlcom normd(`xb') * _b[dxg]

       _nl_1:  normd(_b[dxg]*1.68771118093987+_b[_cons]) * _b[dxg]

------------------------------------------------------------------------------
    a1c_test |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _nl_1 |  -.0006066   .0004859    -1.25   0.212     -.001559    .0003458
------------------------------------------------------------------------------
```

Marginal effects – dummy variable (old_old):

For a dummy variable, it makes no sense to compute a derivative.
If $D_i = 1$ then: $\text{Prob}[y_i = 1 | \mathbf{x}_i, D_i = 1] = \Phi(\mathbf{x}_i' \boldsymbol{\beta} + \delta)$
If $D_i = 0$ then: $\text{Prob}[y_i = 1 | \mathbf{x}_i, D_i = 0] = \Phi(\mathbf{x}_i' \boldsymbol{\beta})$
The impact effect for gender is then given by the differences between the two CDF values:
$\Delta = \Phi(\mathbf{x}_i' \boldsymbol{\beta} + \delta) - \Phi(\mathbf{x}_i' \boldsymbol{\beta})$

```
. probit a1c_test old_old dxg

Iteration 0:   log likelihood = -117738.45
Iteration 1:   log likelihood = -117729.41
Iteration 2:   log likelihood = -117729.41

Probit estimates                                  Number of obs   =     194772
                                                  LR chi2(2)      =      18.08
                                                  Prob > chi2     =     0.0001
Log likelihood = -117729.41                       Pseudo R2       =     0.0001

------------------------------------------------------------------------------
    a1c_test |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
```

```
     old_old |  -.0250912   .0061722    -4.07   0.000    -.0371885   -.0129939
         dxg |  -.0013964   .0014168    -0.99   0.324    -.0041732    .0013805
       _cons |   .5577377   .0044369   125.70   0.000     .5490415     .566434
```

Old-old Impact: What is the effect of old_old on the probability of A1c test?

a) Get mean of dxg

```
. sum dxg
    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
         dxg |    194772    1.687711    2.108394       .068     25.829
```

b) Evaluate mean standardized probit index at this mean and at old_old = 1

```
. display .5577377 + (-.0013964 *1.69) + (-.0250912)
.53028658
```

c) Evaluate mean standardized probit index at this mean and at old_old = 0

```
. display .5577377 + (-.0013964 *1.69)
.55537778
```

d) Find difference between the two CDF values (Notice the use of ***norm*** rather than normd)

```
. display norm(.53028658) - norm(.55537778)
-.00863848
```

Being an old_old decreases the probability of testing (holing comorbidity at the sample mean level) by .86 percentage points.

Check your hand calculation by using mfx compute command (canned routine in STATA)

```
. probit a1c_test old_old dxg

Iteration 0:   log likelihood = -117738.45
Iteration 1:   log likelihood = -117729.41
Iteration 2:   log likelihood = -117729.41

Probit estimates                                  Number of obs   =     194772
                                                  LR chi2(2)      =      18.08
                                                  Prob > chi2     =     0.0001
Log likelihood = -117729.41                       Pseudo R2       =     0.0001

------------------------------------------------------------------------------
    a1c_test |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     old_old |  -.0250912   .0061722    -4.07   0.000    -.0371885   -.0129939
         dxg |  -.0013964   .0014168    -0.99   0.324    -.0041732    .0013805
       _cons |   .5577377   .0044369   125.70   0.000     .5490415     .566434
------------------------------------------------------------------------------

. mfx compute
```

```
Marginal effects after probit
      y  = Pr(a1c_test) (predict)
         =  .70738065
------------------------------------------------------------------------------
variable |      dy/dx    Std. Err.     z    P>|z|  [    95% C.I.   ]      X
---------+--------------------------------------------------------------------
 old_old*|  -.0086385      .00213   -4.06   0.000   -.01281 -.004467   .383895
     dxg |    -.00048      .00049   -0.99   0.324  -.001435  .000475   1.68771
------------------------------------------------------------------------------
(*) dy/dx is for discrete change of dummy variable from 0 to 1
```

Use nlcom

```
. probit a1c_test old_old dxg

Iteration 0:   log likelihood = -117738.45
Iteration 1:   log likelihood = -117729.41
Iteration 2:   log likelihood = -117729.41

Probit estimates                                  Number of obs   =     194772
                                                  LR chi2(2)      =      18.08
                                                  Prob > chi2     =     0.0001
Log likelihood = -117729.41                       Pseudo R2       =     0.0001

------------------------------------------------------------------------------
    a1c_test |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     old_old |  -.0250912   .0061722    -4.07   0.000    -.0371885   -.0129939
         dxg |  -.0013964   .0014168    -0.99   0.324    -.0041732    .0013805
       _cons |   .5577377   .0044369   125.70   0.000     .5490415     .566434

. quietly sum dxg
. local dxgmean = r(mean)
. local xb1  _b[dxg]*`dxgmean'+_b[old_old]*1 + _b[_cons]
. local xb0  _b[dxg]*`dxgmean'+_b[old_old]*0 + _b[_cons]
. nlcom norm(`xb1')  - norm(`xb0')

       _nl_1:  norm(_b[dxg]*1.68771118093987+_b[old_old]*1 + _b[_cons]) -
norm(_b[dxg]*1.68771118093987+_b[old_old]*0 +
>  _b[_cons])
------------------------------------------------------------------------------
    a1c_test |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _nl_1 |  -.0086385   .0021282    -4.06   0.000    -.0128097   -.0044672
```

PROBIT REGRESSION with Interaction Effects

```
. probit a1c_test old_old endo_vis oldXendo dxg

Iteration 0:   log likelihood = -6046.3976
Iteration 1:   log likelihood = -5996.9948
Iteration 2:   log likelihood = -5996.8906
Iteration 3:   log likelihood = -5996.8906

Probit estimates                                  Number of obs   =      10000
                                                  LR chi2(4)      =      99.01
                                                  Prob > chi2     =     0.0000
```

```
Log likelihood = -5996.8906                        Pseudo R2       =     0.0082

------------------------------------------------------------------------------
    a1c_test |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     old_old |   .0171063   .0298301     0.57   0.566    -.0413595    .0755722
    endo_vis |   .3584812   .0445311     8.05   0.000     .2712019    .4457606
    oldXendo |  -.0185596   .0753691    -0.25   0.805    -.1662804    .1291611
         dxg |  -.0025473    .006227    -0.41   0.682     -.014752    .0096574
       _cons |   .4820616   .0208704    23.10   0.000     .4411563    .5229668
------------------------------------------------------------------------------

. mfx compute

Marginal effects after probit
      y  = Pr(a1c_test) (predict)
         =  .70912011
------------------------------------------------------------------------------
variable |      dy/dx    Std. Err.     z    P>|z|  [    95% C.I.   ]      X
---------+--------------------------------------------------------------------
 old_old*|   .0058573       .0102    0.57   0.566  -.014127  .025841     .3816
endo_vis*|   .1144986      .01287    8.90   0.000   .089275  .139722     .1888
oldXendo*|  -.0063902      .02606   -0.25   0.806  -.057473  .044693     .0643
     dxg |  -.0008732      .00213   -0.41   0.682  -.005057   .00331   1.67281
------------------------------------------------------------------------------
(*) dy/dx is for discrete change of dummy variable from 0 to 1
```

Use the formula and get correct marginal effects

Think of all the possible contrasts and evaluate the estimated equation for

1) for Old_old = 1 and endo_vis = 1 (xb1)
2) for old_old = 1 and endo_vis = 0 (xb2)
3) for old_old = 0 and endo_vis = 1 (xb3)
4) for old_old = 0 and endo_vis = 0 (xb4)
5) calculate mean of dxg
6) evaluate the following formula using nlcom

$$\left[\frac{\Delta^2 F(u)}{\Delta x_1 \Delta x_2}\right] = \Phi(\beta_o + \beta_1 + \beta_2 + \beta_3 + \beta_4 * dxgmean) - \Phi(\beta_o + \beta_1 + \beta_4 * dxgmean)$$

$$- \Phi(\beta_o + \beta_2 + \beta_4 * dxgmean) + \Phi(\beta_o + \beta_4 * dxgmean)$$

```
.quietly sum dxg
. local dxgmean = r(mean)

. local xb1 /*
> */    _b[old_old]     /*
> */ +  _b[endo_vis]    /*
> */ +  _b[oldXendo]    /*
> */ +  _b[dxg]*`dxgmean' /*
> */ +  _b[_cons]
```

```
. local xb2 /*
> */    _b[old_old]       /*
> */ +  _b[dxg]*`dxgmean' /*
> */ +  _b[_cons]

.
. local xb3 /*
> */    _b[endo_vis]      /*
> */ +  _b[dxg]*`dxgmean' /*
> */ +  _b[_cons]

.
. local xb4 /*
> */    _b[dxg]*`dxgmean' /*
> */ +  _b[_cons]

. nlcom  norm(`xb1') - norm(`xb2') - norm(`xb3') + norm(`xb4')

_nl_1:  norm(_b[old_old] + _b[endo_vis] + _b[oldXendo] + _b[dxg]*1.672810001328588
>  + _b[_cons]) - norm(_b[old_old] + _b[dxg]*1.672810001328588 + _b[_cons]) -
> norm(_b[endo_vis] + _b[dxg]*1.672810001328588 + _b[_cons]) +
> norm(_b[dxg]*1.672810001328588 + _b[_cons])

------------------------------------------------------------------------------
    a1c_test |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _nl_1 |  -.0064721   .0221576    -0.29   0.770    -.0499002     .036956
------------------------------------------------------------------------------
```

Interpretation:

The interaction effect is negative and insignificant. In our case, all the approaches to estimate marginal effect give similar results.

Check with Dr. Nortons's inteff program

```
. probit a1c_test old_old endo_vis oldXendo dxg

Iteration 0:   log likelihood = -6046.3976
Iteration 1:   log likelihood = -5996.9948
Iteration 2:   log likelihood = -5996.8906
Iteration 3:   log likelihood = -5996.8906

Probit estimates                                  Number of obs   =      10000
                                                  LR chi2(4)      =      99.01
                                                  Prob > chi2     =     0.0000
Log likelihood = -5996.8906                       Pseudo R2       =     0.0082

------------------------------------------------------------------------------
    a1c_test |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     old_old |   .0171063   .0298301     0.57   0.566    -.0413595    .0755722
    endo_vis |   .3584812   .0445311     8.05   0.000     .2712019    .4457606
    oldXendo |  -.0185596   .0753691    -0.25   0.805    -.1662804    .1291611
         dxg |  -.0025473    .006227    -0.41   0.682     -.014752    .0096574
       _cons |   .4820616   .0208704    23.10   0.000     .4411563    .5229668
------------------------------------------------------------------------------
```

```
. inteff a1c_test old_old endo_vis oldXendo dxg ,
Probit with two dummy variables interacted

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
  _probit_ie |     10000    -.006472    .0000176  -.0066553  -.0064586
  _probit_se |     10000    .0221575    .0000908   .0220888   .0231249
   _probit_z |     10000    -.292094     .000398   -.292395  -.2877969
```

LOGISTIC REGRESSION – MARGINAL EFFECTS

$$\text{prob}(y_i = 1) = \frac{\exp(\mathbf{x}_i'\boldsymbol{\beta})}{1+\exp(\mathbf{x}_i'\boldsymbol{\beta})} \quad \text{and} \quad 1-\text{prob}(y_i = 1) = \frac{1}{1+\exp(\mathbf{x}_i'\boldsymbol{\beta})}$$

Continuous variable:

The effect of a small change in the independent variable on the log odds ratio of the event occurring.

$$\frac{\partial \text{Prob}(y_i = 1)}{\partial x_k} = \frac{\partial F}{\partial x_k} = \frac{\exp(\mathbf{x}_i'\boldsymbol{\beta})}{1+\exp(\mathbf{x}_i'\boldsymbol{\beta})} * \frac{\exp(\mathbf{x}_i'\boldsymbol{\beta})}{1+\exp(\mathbf{x}_i'\boldsymbol{\beta})} * \beta_k$$

The marginal effect is then simply the gradient of the logistic CDF at this mean value.
It can also be represented by

$$\frac{\partial \text{Prob}(y_i = 1)}{\partial x_k} = P_i \times (1-P)_i \times \beta_k = \frac{1}{1+\exp-(\mathbf{x}_i'\boldsymbol{\beta})} * \frac{1}{1+\exp(\mathbf{x}_i'\boldsymbol{\beta})} * \beta_k$$

```
. logit a1c_test dxg

Iteration 0:   log likelihood = -117738.45
Iteration 1:   log likelihood = -117737.66
Iteration 2:   log likelihood = -117737.66

Logit estimates                                   Number of obs   =     194772
                                                  LR chi2(1)      =       1.57
                                                  Prob > chi2     =     0.2101
Log likelihood = -117737.66                       Pseudo R2       =     0.0000

------------------------------------------------------------------------------
    a1c_test |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         dxg |  -.0029539    .002354    -1.25   0.210    -.0075676    .0016599
       _cons |   .8876166   .0063787   139.15   0.000     .8751145    .9001187
------------------------------------------------------------------------------
```



```
. mfx compute

Marginal effects after logit
      y  = Pr(a1c_test) (predict)
         =  .70736719
------------------------------------------------------------------------------
variable |      dy/dx    Std. Err.     z    P>|z|  [    95% C.I.   ]      X
---------+--------------------------------------------------------------------
     dxg |  -.0006114      .00049   -1.25   0.210  -.001566  .000344   1.68771
------------------------------------------------------------------------------
```

Hand Calculation:

a) Get mean of dxg

```
. sum dxg
 Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
      dxg |    194772    1.687711    2.108394       .068     25.829
```

b) Evaluate logistic CDF at this mean and take *exponent* of the negative of this

```
. display exp(-((-.0029539 *1.687711) + .8876166))
.41369294
```

c) Evaluate logistic CDF at this mean and take *exponent*

```
. display exp((-.0029539 *1.687711) + .8876166)
2.4172518
```

d) Multiply : 1/(1+4136) * 1/1+2.4172) and the co-efficient of the dxg variable

```
. display (1/(1+.41369294)) * (1/(1+2.4172518)) * -.0029539
-.00061145
```

With nlcom:

```
. quietly sum dxg
. local dxgmean = r(mean)
. local xb _b[dxg]*`dxgmean'+_b[_cons]
. nlcom (1/(1+exp(-(`xb')))) * (1/(1+exp(`xb'))) * _b[dxg]

       _nl_1:  (1/(1+exp(-(_b[dxg]*1.68771118093987+_b[_cons])))) *
(1/(1+exp(_b[dxg]*1.6877111
> 8093987+_b[_cons]))) * _b[dxg]

------------------------------------------------------------------------------
    a1c_test |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _nl_1 |  -.0006114   .0004873    -1.25   0.210    -.0015665    .0003436
------------------------------------------------------------------------------
```

Dummy variable – old_old

```
. logit a1c_test dxg old_old

Iteration 0:   log likelihood = -117738.45
Iteration 1:   log likelihood =  -117729.4
Iteration 2:   log likelihood =  -117729.4

Logit estimates                                   Number of obs   =     194772
                                                  LR chi2(2)      =      18.10
                                                  Prob > chi2     =     0.0001
Log likelihood =  -117729.4                       Pseudo R2       =     0.0001

------------------------------------------------------------------------------
    a1c_test |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         dxg |  -.0023518     .00236    -1.00   0.319    -.0069772    .0022737
     old_old |  -.0416555   .0102408    -4.07   0.000    -.0617271   -.0215839
       _cons |   .9026764   .0073869   122.20   0.000     .8881983    .9171545
------------------------------------------------------------------------------

. mfx compute

Marginal effects after logit
      y  = Pr(a1c_test) (predict)
         =  .70738471
------------------------------------------------------------------------------
variable |      dy/dx    Std. Err.     z    P>|z|  [    95% C.I.   ]      X
---------+--------------------------------------------------------------------
     dxg |  -.0004868      .00049   -1.00   0.319  -.001444  .000471   1.68771
 old_old*|  -.0086395      .00213   -4.06   0.000   -.01281 -.004469   .383895
------------------------------------------------------------------------------
(*) dy/dx is for discrete change of dummy variable from 0 to 1
```

Hand Calculation:

a) Get mean of dxg

```
. sum dxg
    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
         dxg |    194772    1.687711    2.108394       .068     25.829
```

b) Evaluate function when old_old = 1

$$P(Y=1 \mid old\_old, dxg = 1.6877) = \frac{1}{1+\exp-(\beta_0 + \beta_1(1.68) + \beta_2(1))}$$

```
.display exp(-( .9026764 + (-.0023518*1.687711) + (-.0416555)))
.42441151
. display 1/(1+.42441151)
.70204431
```

c) Evaluate function when old_old = 0

$$P(Y=1 \mid old\_old, dxg = 1.6877) = \frac{1}{1+\exp-(\beta_0 + \beta_1(1.68))}$$

```
. display exp(-( .9026764 + (-.0023518*1.687711)))
```

```
.4070956
. display 1/(1+.4070956)
.71068377
```

d) The difference between the two values is the difference in the probability of receiving hba1c test because of age.

```
. display .70204431-.71068377
-.00863946
```

With nlcom:

```
. quietly sum dxg

. local dxgmean = r(mean)

. local xb0 _b[_cons]+(_b[dxg]*`dxgmean')

. local xb1 _b[_cons]+(_b[dxg]*`dxgmean')+_b[old_old]

. nlcom 1/(1+exp(-(`xb1'))) - 1/(1+exp(-(`xb0')))

       _nl_1:  1/(1+exp(-(_b[_cons]+(_b[dxg]*1.68771118093987)+_b[old_old]))) -
1/(1+exp(-(_b[_
> cons]+(_b[dxg]*1.68771118093987))))

------------------------------------------------------------------------------
    a1c_test |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _nl_1 |  -.0086395   .0021281    -4.06   0.000    -.0128104   -.0044685
------------------------------------------------------------------------------
```

LOGISTIC REGRESSION with Interaction Effects

Use the formula and get correct marginal effects

$$\left[\frac{\Delta^2 F(u)}{\Delta x_1 \Delta x_2}\right] = \left[\frac{1}{1+\exp-(\beta_o+\beta_1+\beta_2+\beta_3+\beta_4 * dxgmean)}\right] -$$
$$\left[\frac{1}{1+\exp-(\beta_o+\beta_1+\beta_4 * dxgmean)}\right] - \left[\frac{1}{1+\exp-(\beta_o+\beta_2+\beta_4 * dxgmean)}\right] +$$
$$\left[\frac{1}{1+\exp-(\beta_o+\beta_4 * dxgmean)}\right]$$

Think of all the possible contrasts and evaluate the estimated equation for

1) for Old_old = 1 and endo_vis = 1 (xb1)
2) for old_old = 1 and endo_vis = 0 (xb2)
3) for old_old = 0 and endo_vis = 1 (xb3)
4) for old_old = 0 and endo_vis = 0 (xb4)

---

5) calculate mean of dxg
6) evaluate the following formula using nlcom

```
. logit a1c_test old_old endo_vis oldXendo dxg

Iteration 0:   log likelihood = -6046.3976
Iteration 1:   log likelihood = -5997.3365
Iteration 2:   log likelihood = -5996.8874
Iteration 3:   log likelihood = -5996.8873

Logit estimates                                   Number of obs   =      10000
                                                  LR chi2(4)      =      99.02
                                                  Prob > chi2     =     0.0000
Log likelihood = -5996.8873                       Pseudo R2       =     0.0082

------------------------------------------------------------------------------
    a1c_test |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     old_old |   .0281896   .0491501     0.57   0.566    -.0681429    .1245221
    endo_vis |    .606646   .0770566     7.87   0.000     .4556177    .7576742
    oldXendo |  -.0309183   .1305416    -0.24   0.813    -.2867751    .2249385
         dxg |  -.0043481   .0104154    -0.42   0.676    -.0247619    .0160658
       _cons |   .7776468   .0344863    22.55   0.000     .7100549    .8452387
------------------------------------------------------------------------------

. mfx compute

Marginal effects after logit
      y  = Pr(a1c_test) (predict)
         =  .70964843
------------------------------------------------------------------------------
variable |      dy/dx    Std. Err.     z    P>|z|  [    95% C.I.   ]      X
---------+--------------------------------------------------------------------
 old_old*|   .0058002      .01009    0.57   0.565  -.013978  .025578     .3816
endo_vis*|   .1144238      .01281    8.93   0.000   .089309  .139538     .1888
oldXendo*|  -.0064064       .0272   -0.24   0.814  -.059716  .046903     .0643
     dxg |  -.0008959      .00215   -0.42   0.676  -.005102   .00331   1.67281
------------------------------------------------------------------------------
(*) dy/dx is for discrete change of dummy variable from 0 to 1

. *-----------------------------------
. * nlcom to get differences in p
. * Old-old
. *-----------------------------------
. quietly sum dxg

. local dxgmean = r(mean)

. local xb1 /*
> */    _b[old_old]     /*
> */ +  _b[endo_vis]    /*
> */ +  _b[oldXendo]    /*
> */ +  _b[dxg]*`dxgmean' /*
> */ +  _b[_cons]

.
. local xb2 /*
>  */    _b[old_old]      /*
>  */ +  _b[dxg]*`dxgmean' /*
```

```
>  */ +  _b[_cons]

.
. local xb3 /*
>  */    _b[endo_vis]      /*
>  */ +  _b[dxg]*`dxgmean' /*
>  */ +  _b[_cons]

.
.  local xb4 /*
>  */    _b[dxg]*`dxgmean' /*
>  */ +  _b[_cons]

.
.  nlcom  1/(1+(exp(-(`xb1')))) - 1/(1+(exp(-(`xb2'))))  - 1/(1+(exp(-(`xb3')))) +
1/(1+(exp(-(
> `xb4'))))

       _nl_1:  1/(1+(exp(-(_b[old_old] + _b[endo_vis] + _b[oldXendo] +
_b[dxg]*1.67281000132858
> 8 + _b[_cons])))) - 1/(1+(exp(-(_b[old_old] + _b[dxg]*1.672810001328588 +
_b[_cons])))) - 1/(
> 1+(exp(-(_b[endo_vis] + _b[dxg]*1.672810001328588 + _b[_cons])))) + 1/(1+(exp(-
(_b[dxg]*1.672
> 810001328588 + _b[_cons]))))

------------------------------------------------------------------------------
    a1c_test |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _nl_1 |  -.0065047    .022156    -0.29   0.769    -.0499296    .0369201
```

REFERENCES:

Ai C, Norton EC: Interaction Term in Logit and Probit Models. *Economic Letters* 80:123-129.

Norton EC, Ai. C: Computing interaction effects and standard errors in logit and probit models *The Stata Journal*, 2004, 4(2):103-116