Background:
In probit or logistic regressions, one cannot 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 nlcum. 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 nlcum 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 \textit{normd}) and for dummy independent variable (use of function \textit{norm})

ii. calculate marginal effects – hand calculation

iii. calculate marginal effects – use of \texttt{dprobit}

iv. calculate marginal effects – use of \texttt{mfx} command

v. calculate marginal effects – use of \texttt{nlcom}

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

n. Logistic regression
   i. calculate marginal effects – hand calculation
   ii. calculate marginal effects – use of \texttt{mfx} command
   iii. calculate effect using \texttt{nlcom}
   iv. calculate interaction effect using \texttt{nlcom} – using Dr. Norton’s method

\textbf{Odds versus probability:}

\textbf{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.

\textbf{Probability:} Probability is the expected number of flu patients divided by the total number of patients.

\textbf{Relationship:}

\[
\text{Odds} = \frac{\text{probability}}{(1 - \text{probability})}. = \frac{\text{Probability}}{1 - \text{Probability}}
\]

\textit{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{\text{odds}}{1 + \text{odds}}

**Example:**
If the odds are 10:1 then the probability = \frac{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:**
\logit(p) = \beta_0 + \beta_1 \text{ old } _\text{old} \\
\logit(p) = \beta_0 + \beta_1 \text{ old } _\text{old}

```
. logit a1c_test old_old
```

|       | Coef.    | Std. Err. | z    | P>|z|  | [95% Conf. Interval] |
|-------|----------|-----------|------|------|----------------------|
| a1c_test | -.0422966 | .0102205  | -4.14| 0.000| -.0623285  -.0222648 |
| _cons  | .8999483  | .0063666  | 141.20| 0.000| .88647  .9114266 |

Std. Err for odds ratios is not meaningful.

```
. logit
```

|       | Coef. | Std. Err. | z    | P>|z|  | [95% Conf. Interval] |
|-------|-------|-----------|------|------|----------------------|
| old_old | -.0422966 | .0102205  | -4.14| 0.000| -.0623285  -.0222648 |
| _cons  | .8999483  | .0063666  | 141.20| 0.000| .88647  .9114266 |

When old _old = 1, the risk of A1c test is
\logit(p_1) = \beta_0 + \beta_1

When old _old = 0 the risk of A1c test is
\logit(p_0) = \beta_0

Take the difference:
\logit(p_1) - \logit(p_0) = ([\beta_0 + \beta_1] - \beta_0) = \beta_1
Interaction Terms Vs. Interaction Effects in Logistic and Probit Regression  
CRMportals Inc.,  
---------------------------------------------------------------------------------------------------------------------  
Copyrights – 2006 CRMportals Inc., 4

Odds ratio:

\[
\ln \left( \frac{p_1/(1-p_1)}{p_0/(1-p_0)} \right) = \ln(OR) = \beta_i
\]

Model with interaction

Let us fit the following model with interaction:

\[
\text{logit}(p) = \beta_0 + \beta_1 \text{old_old} + \beta_2 \text{endo_vis} + \beta_3 \text{old_old} \ast \text{endo_vis} \quad \text{(Interaction)}
\]

\[
\ln \left( \frac{p}{1-p} \right) = \beta_0 + \beta_1 \text{old_old} + \beta_2 \text{endo_vis} + \beta_3 \text{old_old} \ast \text{endo_vis}
\]

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

```stata
logistic alc_test old_old endo_vis oldXendo
```

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” cannot 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.

```stata
. adjust, by (old_old endo_vis) exp
```
Interaction Terms Vs. Interaction Effects in Logistic and Probit Regression

CRMportals Inc.,

---------------------------------------------------------------------------------------------------------------------
Copyrights – 2006 CRMportals Inc., 5

<table>
<thead>
<tr>
<th>Endocrinologist</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age &gt;= 75</td>
</tr>
<tr>
<td>0</td>
</tr>
<tr>
<td>-----------------</td>
</tr>
<tr>
<td>0</td>
</tr>
<tr>
<td>1</td>
</tr>
</tbody>
</table>

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                          Pseudo R2       =     0.0064
Prob > chi2     =     0.0000                          Log likelihood = -116985.08
------------------------------------------------------------------------------
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
------------------------------------------------------------------------------
Interaction Terms Vs. Interaction Effects in Logistic and Probit Regression

CRMportals Inc.,

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 ($\beta_1, \beta_3$) are gone because old_old = 0 and the interaction term becomes zero.

Then take the differences:

$$\log it(p_1) - \log it(p_0) = [\beta_0 + \beta_1 + \beta_2 + \beta_3] - [\beta_0 + \beta_2]$$

$$\log it(p_1) - \log it(p_0) = \beta_1 + \beta_3$$

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”

$$\exp(\beta_1 + \beta_3) = \text{odds ratio} = \exp (-.0396509 + .0652091) = 1.0258876$$

. 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

|   | Odds Ratio | Std. Err. | z   | P>|z| | [95% Conf. Interval] |
|---|------------|-----------|-----|------|----------------------|
|   | 1.025888   | 0.0279809 | 0.94| 0.349| 0.9724863 1.082221   |

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

<table>
<thead>
<tr>
<th>Old_old = 1</th>
<th>Endo_vis = 1</th>
<th>Endo_vis = 0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_0 + \beta_1 + \beta_2 + \beta_3$</td>
<td>$\beta_0 + \beta_1$</td>
<td></td>
</tr>
</tbody>
</table>
Old_old = 0 | $\beta_0 + \beta_2$ | $\beta_0$

For example, the odds of A1c test among old_old and with endo_vis = 0 is: $\exp(\beta_0 + \beta_1)$

**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 $\text{Odds} = \left[ \frac{p}{1-p} \right]$ then $p = \left[ \frac{\text{odds}}{1+\text{odds}} \right]

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

$$\frac{\exp(\beta_0 + \beta_1)}{1 + (\exp(\beta_0 + \beta_1))}$$

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

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

\[
\exp(\beta_0 + \beta_1 + \beta_2 + \beta_3)
\]

\[
. \text{display } \exp(.8109787+(-.0396509) + .501553 + .0652091)
3.8117557
\]

\[
1 + \exp((\beta_0 + \beta_1 + \beta_2 + \beta_3)
\]

\[
. \text{display } 1 + (\exp(.8109787+(-.0396509) + .501553 + .0652091))
4.8117557
\]

Numerator/Denominator:

\[
. \text{display } 3.8117557/4.8117557
.79217565
\]

Using \textit{nlcom} – risk difference

\[
. \text{logit a1c_test old_old}
\]

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

Logit estimates

\[
\begin{align*}
\text{a1c_test} & \mid \text{Coef.} \quad \text{Std. Err.} \quad z \quad \text{P>|z|} \quad \text{[95% Conf. Interval]} \\
\text{old_old} & \mid -0.0422966 \quad 0.0102205 \quad -4.14 \quad 0.000 \quad -0.0623285 \quad -0.0222648 \\
\text{cons} & \mid 0.8989483 \quad 0.0063666 \quad 141.20 \quad 0.000 \quad 0.88647 \quad 0.9114266
\end{align*}
\]

\[
P_1 - P_0 = \frac{1}{1 + \exp(-\beta_0 - \beta_1)} - \frac{1}{1 + \exp(-\beta_0)}
\]

\[
. \text{nlcom } 1/(1+\exp(-_b[old_old] - _b[cons])) - 1/(1+\exp(-_b[cons]))
\]

\[
._nl_1: 1/(1+\exp(-_b[old_old] - _b[cons])) - 1/(1+\exp(-_b[cons]))
\]

\[
\begin{align*}
\text{alc_test} & \mid \text{Coef.} \quad \text{Std. Err.} \quad z \quad \text{P>|z|} \quad \text{[95% Conf. Interval]} \\
._nl_1 & \mid -0.0087727 \quad 0.002124 \quad -4.13 \quad 0.000 \quad -0.0129356 \quad -0.0046098
\end{align*}
\]

\[
\text{cs a1c_test old_old}
\]

| Age >= 75 | \\
| Exp | Unexp | Total |
|-----------------+-----------------+----------|
| Cases | 52487 | 85288 | 137775 |
| Noncases | 22285 | 34712 | 56997 |
| Total | 74772 | 120000 | 194772 |

Copyrights – 2006 CRMportals Inc., 8
Interaction Terms Vs. Interaction Effects in Logistic and Probit Regression

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

<table>
<thead>
<tr>
<th>Risk</th>
<th>.7019606</th>
<th>.7107333</th>
<th>.7073655</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Point estimate</td>
<td>[95% Conf. Interval]</td>
<td></td>
</tr>
<tr>
<td>Risk difference</td>
<td>-.0087727</td>
<td>-.0129356</td>
<td>-.0046098</td>
</tr>
<tr>
<td>Risk ratio</td>
<td>.9876568</td>
<td>.9818441</td>
<td>.9935039</td>
</tr>
<tr>
<td>Prev. frac. ex.</td>
<td>.0123432</td>
<td>.0064961</td>
<td>.0181559</td>
</tr>
<tr>
<td>Prev. frac. pop</td>
<td>.0047385</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

+-----------------------------------------------
   chi2(1) = 17.13  Pr>chi2 = 0.0000

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

\[ \frac{\partial \text{Prob}(y_i = 1)}{\partial x_k} = \frac{\partial \Phi}{\partial x_k} = \phi(x_i \beta) \times \beta_k \]

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 \( x_i \beta \).

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 \( dxg = -.0006 \approx -.001 \); 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  
---------+------------------------------------------------------
dxg |  -.0006066   .0004859    -1.25   0.212   1.68771  
---------+------------------------------------------------------
obs. P |   .7073655
pred. P |   .7073668  (at x-bar)
------------------------------------------------------------------------------
z and P>|z| are the test of the underlying coefficient being 0
```

use nlcom
Interaction Terms Vs. Interaction Effects in Logistic and Probit Regression

CRMportals Inc.,

. probit alc_test dxg
Iteration 0: log likelihood = -117738.45
Iteration 1: log likelihood = -117737.67
Iteration 2: log likelihood = -117737.67

Probit estimates

|                | 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]*1.6877118093987+_b[_cons]) * _b[dxg]

. probit alc_test old_old dxg
Iteration 0: log likelihood = -117738.45
Iteration 1: log likelihood = -117729.41
Iteration 2: log likelihood = -117729.41

Probit estimates

|                | 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: \( \text{Prob}[y_i = 1 | x_i, D_i = 1] = \Phi(x_i \beta + \delta) \)
If  \( D_i = 0 \) then: \( \text{Prob}[y_i = 1 | x_i, D_i = 0] = \Phi(x_i \beta) \)
The impact effect for gender is then given by the differences between the two CDF values:
\[ \Delta = \Phi(x_i \beta + \delta) - \Phi(x_i \beta) \]

. probit alc_test old_old dxg
Iteration 0: log likelihood = -117738.45
Iteration 1: log likelihood = -117729.41
Iteration 2: log likelihood = -117729.41

Probit estimates

|                | Coef.   | Std. Err. |      z | P>|z| | [95% Conf. Interval] |
|----------------|---------|-----------|-------|------|---------------------|
| alc_test       | 1.6877118093987 | 0.0038349 | 143.08 | 0.000 | .5411706 .5562029   |

Copyrights – 2006 CRMportals Inc.,

11
Old-old Impact: What is the effect of old_old on the probability of A1c test?

a) Get mean of dxg

\[ \text{sum dxg} \]

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>dxg</td>
<td>19472</td>
<td>1.687711</td>
<td>2.108394</td>
<td>.068</td>
<td>25.829</td>
</tr>
</tbody>
</table>

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

\[ \text{display .5577377 + (-.0013964 \times 1.69) + (-.0250912)} \]
\[ .53028658 \]

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

\[ \text{display .5577377 + (-.0013964 \times 1.69)} \]
\[ .55537778 \]

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

\[ \text{display norm(.53028658) - norm(.55537778)} \]
\[ -.00863848 \]

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

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

```stata
.probit alc_test old_old dxg
```

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

Probit estimates

| alc_test | Coef.   | Std. Err. | z      | P>|z| | [95% Conf. Interval] |
|----------|---------|-----------|--------|-----|---------------------|
| old_old  | -0.0250912 | 0.0061722 | -4.07  | 0.000 | -0.0371885 to -0.0129939 |
| dxg      | -0.0013964 | 0.0014168 | -0.99  | 0.324 | -0.0041732 to 0.0013805 |
| _cons    | 0.5577377  | 0.0044369 | 125.70 | 0.000 | 0.5490415 to 0.566434 |

. mfx compute
Marginal effects after probit

\[ y = \Pr(\text{a1c_test}) \] (predict)
\[ = 0.70738065 \]

| variable | dy/dx   | Std. Err. | z    | P>|z| | [95% C.I.]   | X |
|----------|---------|-----------|------|------|----------------|---|
| old_old* | -0.0086385 | 0.00213   | -4.06 | 0.000 | -0.01281 - 0.004467 | 0.383895 |
| dxg      | -0.00048 | 0.00049 | -0.99 | 0.324 | -0.001435 - 0.000475 | 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  | -0.0250912 | 0.0061722 | -4.07 | 0.000 | -0.0371885 - 0.0129939 |
| dxg      | -0.0013964 | 0.0014168 | -0.99 | 0.324 | -0.0041732 - 0.0013805 |
| _cons    | 0.5577377 | 0.0044369 | 125.70 | 0.000 | 0.5490415 - 0.566434 |
```

```
.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]*1.6877118093987 + _b[old_old]*1 + _b[_cons]) - norm(_b[dxg]*1.6877118093987 + _b[old_old]*0 + _b[_cons])

| a1c_test | Coef. | Std. Err. | z    | P>|z| | [95% Conf. Interval] |
|----------|-------|-----------|------|------|----------------------|---|
| nl_1     | -0.0086385 | 0.0021282 | -4.06 | 0.000 | -0.0128097 - 0.0044672 |
```

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

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

\[
\frac{\Delta^2 F(u)}{\Delta x_1 \Delta x_2} = \Phi(\beta_0 + \beta_1 + \beta_2 + \beta_3 + \beta_4 \cdot dxgmean) - \Phi(\beta_0 + \beta_1 + \beta_4 \cdot dxgmean)
- \Phi(\beta_0 + \beta_2 + \beta_4 \cdot dxgmean) + \Phi(\beta_0 + \beta_4 \cdot dxgmean)
\]

.quickly sum dxg
.local dxgmean = r(mean)

.local xb1 /*
> /*/ _b[old_old] /*
> */ _b[eno_vis] /*
> */ + _b[oldXendo] /*
> */ + _b[dxg]*dxgmean' /*
> */ + _b[_cons] /*
Interaction Terms Vs. Interaction Effects in Logistic and Probit Regression

CRMportals Inc.,

---

. 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')

_\text{nl}_1: \text{norm}(\_b[old\_old] + \_b[endo\_vis] + \_b[oldXendo] + \_b[dxg]1.67281001328588 + \_b[_cons]) - \text{norm}(\_b[old\_old] + \_b[dxg]1.67281001328588 + \_b[_cons]) - \text{norm}(\_b[endo\_vis] + \_b[dxg]1.67281001328588 + \_b[_cons]) + \text{norm}(\_b[dxg]1.67281001328588 + \_b[_cons])

---

\begin{tabular}{lccccc}
\hline
\text{a1c\_test} & \text{Coef.} & \text{Std. Err.} & \text{z} & \text{P>|z|} & [95\% \text{Conf. Interval}] \\
\hline
_\text{nl}_1 & -0.0064721 & 0.0221576 & -0.29 & 0.770 & -0.0499002 & 0.036956 \\
\hline
\end{tabular}

\textbf{Interpretation:}

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

\textbf{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

\begin{itemize}
\item Number of obs = 10000
\item LR chi2(4) = 99.01
\item Prob > chi2 = 0.0000
\end{itemize}

Log likelihood = -5996.8906

\begin{tabular}{lccccc}
\hline
\text{a1c\_test} & \text{Coef.} & \text{Std. Err.} & \text{z} & \text{P>|z|} & [95\% \text{Conf. Interval}] \\
\hline
\text{old\_old} & 0.0171063 & 0.0298301 & 0.57 & 0.566 & -0.0413595 & 0.0755722 \\
\text{endo\_vis} & 0.3584812 & 0.0445311 & 8.05 & 0.000 & 0.2712019 & 0.4457606 \\
\text{oldXendo} & -0.0185596 & 0.0753691 & -0.25 & 0.805 & -0.1662804 & 0.1291611 \\
\text{dxg} & -0.0025473 & 0.006227 & -0.41 & 0.682 & -0.014752 & 0.0096574 \\
\text{_cons} & 0.4820616 & 0.0208704 & 23.10 & 0.000 & 0.4411563 & 0.5229668 \\
\hline
\end{tabular}

Copyrights – 2006 CRMportals Inc., 15
Interaction Terms Vs. Interaction Effects in Logistic and Probit Regression
CRMportals Inc.,

---

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

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>_probit_ie</td>
<td>10000</td>
<td>-.006472</td>
<td>.0000176</td>
<td>-.0066553</td>
<td>-.0064586</td>
</tr>
<tr>
<td>_probit_se</td>
<td>10000</td>
<td>.0221575</td>
<td>.0000908</td>
<td>.0220888</td>
<td>.0231249</td>
</tr>
<tr>
<td>_probit_z</td>
<td>10000</td>
<td>-.292094</td>
<td>.000398</td>
<td>-.292395</td>
<td>-.2877969</td>
</tr>
</tbody>
</table>
```

LOGISTIC REGRESSION – MARGINAL EFFECTS

\[
\text{prob}(y_i = 1) = \frac{\exp(x_i\beta)}{1 + \exp(x_i\beta)} \quad \text{and} \quad 1 - \text{prob}(y_i = 1) = \frac{1}{1 + \exp(x_i\beta)}
\]

Continuous variable:

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

\[
\frac{\partial \text{Prob}(y_i = 1)}{\partial x_k} = \frac{\partial F}{\partial x_k} = \frac{\exp(x_i\beta)}{1 + \exp(x_i\beta)} * \frac{\exp(x_i\beta)}{1 + \exp(x_i\beta)} * \beta_k
\]

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

\[
\frac{\partial \text{Prob}(y_i = 1)}{\partial x_k} = \text{P}(1 - \text{P})\beta_k = \frac{1}{1 + \exp(x_i\beta)} * \frac{1}{1 + \exp(x_i\beta)} * \beta_k
\]

```
. logit alc_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

| alc_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  |
```

Copyrights – 2006 CRMportals Inc.,
Interaction Terms Vs. Interaction Effects in Logistic and Probit Regression
CRMportals Inc.,

.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      | -0.0006114 | .00049    | -1.25 | 0.210 | (-0.001566, 0.000344) | 1.68771 |

Hand Calculation:

a) Get mean of dxg

.sum dxg

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>dxg</td>
<td>194772</td>
<td>1.687711</td>
<td>2.108394</td>
<td>.068</td>
<td>25.829</td>
</tr>
</tbody>
</table>

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

.display exp(-((-0.0029539 * 1.687711) + .8876166))
.41369294

c) Evaluate logistic CDF at this mean and take exponent

.display exp((-0.0029539 * 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]*1.68771118093987+_b[_cons])))) * _b[dxg]

| alc_test | Coef. | Std. Err. | z    | P>|z| | 95% Conf. Interval |
|----------|-------|-----------|------|------|-------------------|
| _nl_1    | -.0006114 | .0004873 | -1.25 | 0.210 | (-0.0015665, .0003436) |

Dummy variable – old_old

Copyrights – 2006 CRMportals Inc., 17
Interaction Terms Vs. Interaction Effects in Logistic and Probit Regression

CRMportals Inc.,

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 | old_old = 1, dxg = 1.6877) = \frac{1}{1 + \exp(-\beta_0 + \beta_1(1.6877) + \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 | old_old = 0, dxg = 1.6877) = \frac{1}{1 + \exp(-\beta_0 + \beta_1(1.68))}
```

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

Copyrights – 2006 CRMportals Inc., 18
Interaction Terms Vs. Interaction Effects in Logistic and Probit Regression
CRMportals Inc.,

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

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 alc_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

------------------------------------------------------------------
          alc_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(alc_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\} /*

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

> */ + _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]*1.6728100132858 + 
8 + _b[_cons]))) - 1/(1+(exp(-(_b[old_old] + _b[dxg]*1.6728100132858 + 
_b[_cons]))) - 1/(1+(exp(-(_b[endo_vis] + _b[dxg]*1.6728100132858 + _b[_cons]))) - 1/(1+(exp(-(_b[dxg]*1.672

REFERENCE:
