Using Stata's Margins Command to Estimate and Interpret Adjusted Predictions and Marginal Effects

Richard Williams
rwilliam@ND.Edu
http://www.nd.edu/~rwilliam/
University of Notre Dame
Last revised January 14, 2017

Motivation for Paper

- Many journals place a strong emphasis on the sign and statistical significance of effects – but often there is very little emphasis on the substantive and practical significance
- Unlike scholars in some other fields, most Sociologists seem to know little about things like marginal effects or adjusted predictions, let alone use them in their work
- Many users of Stata seem to have been reluctant to adopt the margins command.
  - The manual entry is long, the options are daunting, the output is sometimes unintelligible, and the advantages over older and simpler commands like adjust and mfx are not always understood

- This presentation therefore tries to do the following
  - Briefly explain what adjusted predictions and marginal effects are, and how they can contribute to the interpretation of results
  - Explain what factor variables (introduced in Stata 11) are, and why their use is often critical for obtaining correct results
  - Explain some of the different approaches to adjusted predictions and marginal effects, and the pros and cons of each:
    - APMs (Adjusted Predictions at the Means)
    - AAPs (Average Adjusted Predictions)
    - APRs (Adjusted Predictions at Representative values)
    - MEMs (Marginal Effects at the Means)
    - AMEs (Average Marginal Effects)
    - MERs (Marginal Effects at Representative values)

Adjusted Predictions - New margins versus the old adjust

```
. version 11.1
. webuse nhanes2f, clear
. keep if !missing(diabetes, black, female, age, age2, agegrp)
(2 observations deleted)
. label variable age2 "age squared"
. * Compute the variables we will need
. tab1 agegrp, gen(agegrp)
. gen femage = female*age
. label variable femage "female * age interaction"
. sum diabetes black female age age2 femage, separator(6)

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
    diabetes |     10335    .0482825     .214373          0          1
       black |     10335    .1050798    .3066711          0          1
      female |     10335    .5250121    .4993982          0          1
         age |     10335    47.56584    17.21752         20         74
        age2 |     10335    2558.924    1616.804        400       5476
      femage |     10335    25.05031    26.91168          0         74
```

Model 1: Basic Model

```
. logit diabetes black female age , nolog

Logistic regression                               Number of obs   =      10335
                                                  LR chi2(3)      =     374.17
                                                  Prob > chi2     =     0.0000
Log likelihood = -1811.9828                       Pseudo R2       =     0.0936

------------------------------------------------------------------------------
    diabetes |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       black |   .7179046   .1268061     5.66   0.000     .4693691      .96644
      female |   .1545569   .0942982     1.64   0.101    -.0302642    .3393779
         age |   .0594654   .0037333    15.93   0.000     .0521484    .0667825
       _cons |  -6.405437   .2372224   -27.00   0.000    -6.870384    -5.94049
------------------------------------------------------------------------------
```

- Among other things, the results show that getting older is bad for your health – but just how bad is it???
- Adjusted predictions (aka predictive margins) can make these results more tangible.
- With adjusted predictions, you specify values for each of the independent variables in the model, and then compute the probability of the event occurring for an individual who has those values.
- So, for example, we will use the adjust command to compute the probability that an “average” 20 year old will have diabetes and compare it to the probability that an “average” 70 year old will.

```
. adjust age = 20 black female, pr

------------------------------------------------------------------------------
     Dependent variable: diabetes     Equation: diabetes     Command: logit
 Covariates set to mean: black = .10507983, female = .52501209
 Covariate set to value: age = 20
------------------------------------------------------------------------------

----------------------
      All |         pr
----------+-----------
          |    .006308
----------------------
     Key:  pr  =  Probability

. adjust age = 70 black female, pr

------------------------------------------------------------------------------
     Dependent variable: diabetes     Equation: diabetes     Command: logit
 Covariates set to mean: black = .10507983, female = .52501209
 Covariate set to value: age = 70
------------------------------------------------------------------------------

----------------------
      All |         pr
----------+-----------
          |    .110438
----------------------
     Key:  pr  =  Probability
```

- The results show that a 20 year old has less than a 1 percent chance of having diabetes, while an otherwise-comparable 70 year old has an 11 percent chance.

- But what does “average” mean? In this case, we used the common, but not universal, practice of using the mean values for the other independent variables (female, black) that are in the model.

- The margins command easily (in fact more easily) produces the same results

```
. margins, at(age=(20 70)) atmeans vsquish

Adjusted predictions                            Number of obs   =      10335
Model VCE    : OIM

Expression   : Pr(diabetes), predict()
1._at        : black           =    .1050798 (mean)
               female          =    .5250121 (mean)
               age             =          20
2._at        : black           =    .1050798 (mean)
               female          =    .5250121 (mean)
               age             =          70

------------------------------------------------------------------------------
             |            Delta-method
             |     Margin   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         _at |
          1  |   .0063084   .0009888     6.38   0.000     .0043703    .0082465
          2  |   .1104379    .005868    18.82   0.000     .0989369     .121939
------------------------------------------------------------------------------
```

Factor variables

- So far, we have not used factor variables (or even explained what they are)
- The previous problems were addressed equally well with both older Stata commands and the newer margins command
- We will now show how margin's ability to use factor variables makes it much more powerful and accurate than its predecessors

Model 2: Squared term added

```
. quietly logit diabetes black female age age2, nolog
. adjust age = 70 black female age2, pr

------------------------------------------------------------------------------
     Dependent variable: diabetes     Equation: diabetes     Command: logit
 Covariates set to mean: black = .10507983, female = .52501209, age2 = 2558.9238
 Covariate set to value: age = 70
------------------------------------------------------------------------------

----------------------
      All |         pr
----------+-----------
          |    .373211
----------------------
     Key:  pr  =  Probability
```

- In this model, adjust reports a much higher predicted probability of diabetes than before – 37 percent as opposed to 11 percent!
- But, luckily, adjust is wrong. Because it does not know that age and age2 are related, it uses the mean value of age2 in its calculations, rather than the correct value of 70 squared.
- While there are ways to fix this, using the margins command and factor variables is a safer solution.
  - The use of factor variables tells margins that age and age^2 are not independent of each other and it does the calculations accordingly.
  - In this case it leads to a much smaller (and also correct) estimate of 10.3 percent.

```
. quietly logit diabetes i.black i.female age c.age#c.age, nolog
. margins, at(age = 70) atmeans

Adjusted predictions                            Number of obs   =      10335
Model VCE    : OIM

Expression   : Pr(diabetes), predict()
at           : 0.black         =    .8949202 (mean)
               1.black         =    .1050798 (mean)
               0.female        =    .4749879 (mean)
               1.female        =    .5250121 (mean)
               age             =          70

------------------------------------------------------------------------------
             |            Delta-method
             |     Margin   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _cons |   .1029814   .0063178    16.30   0.000     .0905988     .115364
------------------------------------------------------------------------------
```

- The i.black and i.female notation tells Stata that black and female are categorical variables rather than continuous. As the Stata 11 User Manual explains (section 11.4.3.1), “i.group is called a factor variable…When you type i.group, it forms the indicators for the unique values of group.”
- The # (pronounced cross) operator is used for interactions.
  - The use of # implies the i. prefix, i.e. unless you indicate otherwise Stata will assume that the variables on both sides of the # operator are categorical and will compute interaction terms accordingly.
  - Hence, we use the c. notation to override the default and tell Stata that age is a continuous variable.
  - So, c.age#c.age tells Stata to include age^2 in the model; we do not want or need to compute the variable separately.
  - By doing it this way, Stata knows that if age = 70, then age^2 = 4900, and it hence computes the predicted values correctly.

Model 3: Interaction Term

```
. quietly logit diabetes black female age femage, nolog
. * Although not obvious, adjust gets it wrong
. adjust female = 0 black age femage, pr

-------------------------------------------------------------------------------------
     Dependent variable: diabetes     Equation: diabetes     Command: logit
 Covariates set to mean: black = .10507983, age = 47.565844, femage = 25.050314
 Covariate set to value: female = 0
-------------------------------------------------------------------------------------
----------------------
      All |         pr
----------+-----------
          |    .015345
----------------------
     Key:  pr  =  Probability
```

- Once again, adjust gets it wrong
- If female = 0, femage must also equal zero
- But adjust does not know that, so it uses the average value of femage instead.
- Margins (when used with factor variables) does know that the different components of the interaction term are related, and does the calculation right.

```
. quietly logit diabetes i.black i.female age i.female#c.age, nolog
. margins female, atmeans grand

Adjusted predictions                            Number of obs   =      10335
Model VCE    : OIM

Expression   : Pr(diabetes), predict()
at           : 0.black         =    .8949202 (mean)
               1.black         =    .1050798 (mean)
               0.female        =    .4749879 (mean)
               1.female        =    .5250121 (mean)
               age             =    47.56584 (mean)

------------------------------------------------------------------------------
             |            Delta-method
             |     Margin   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      female |
          0  |   .0250225   .0027872     8.98   0.000     .0195597    .0304854
          1  |   .0372713   .0029632    12.58   0.000     .0314635    .0430791
             |
       _cons |   .0308641   .0020865    14.79   0.000     .0267746    .0349537
------------------------------------------------------------------------------
```

Model 4: Multiple dummies

```
. quietly logit diabetes black female agegrp2 agegrp3 agegrp4 agegrp5 agegrp6
. adjust agegrp6 = 1 black female agegrp2 agegrp3 agegrp4 agegrp5, pr

------------------------------------------------------------------------------------------
     Dependent variable: diabetes     Equation: diabetes     Command: logit
 Covariates set to mean: black = .10507983, female = .52501209, agegrp2 = .15674891,
agegrp3 = .12278665, agegrp4 = .12472182, agegrp5 = .27595549
 Covariate set to value: agegrp6 = 1
------------------------------------------------------------------------------------------

----------------------
      All |         pr
----------+-----------
          |    .320956
----------------------
     Key:  pr  =  Probability
```

- More depressing news for old people: now adjust says they have a 32 percent chance of having diabetes

- But once again adjust is wrong: If you are in the oldest age group, you can’t also have partial membership in some other age category. 0, not the means, is the correct value to use for the other age variables when computing probabilities.

- Margins (with factor variables) realizes this and does it right again.

```
. quietly logit diabetes i.black i.female i.agegrp, nolog
. margins agegrp, atmeans grand

Adjusted predictions                            Number of obs   =      10335
Model VCE    : OIM

Expression   : Pr(diabetes), predict()
at           : 0.black         =    .8949202 (mean)
               1.black         =    .1050798 (mean)
               0.female        =    .4749879 (mean)
               1.female        =    .5250121 (mean)
               1.agegrp        =    .2244799 (mean)
               2.agegrp        =    .1567489 (mean)
               3.agegrp        =    .1227866 (mean)
               4.agegrp        =    .1247218 (mean)
               5.agegrp        =    .2759555 (mean)
               6.agegrp        =    .0953072 (mean)

------------------------------------------------------------------------------
             |            Delta-method
             |     Margin   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      agegrp |
          1  |   .0061598   .0015891     3.88   0.000     .0030453    .0092744
          2  |   .0124985    .002717     4.60   0.000     .0071733    .0178238
          3  |   .0323541   .0049292     6.56   0.000     .0226932    .0420151
          4  |   .0541518   .0062521     8.66   0.000      .041898    .0664056
          5  |    .082505   .0051629    15.98   0.000     .0723859     .092624
          6  |   .1106978    .009985    11.09   0.000     .0911276     .130268
             |
       _cons |   .0303728   .0022281    13.63   0.000     .0260059    .0347398
------------------------------------------------------------------------------
```

Different Types of Adjusted Predictions

- There are at least three common approaches for computing adjusted predictions
  - APMs (Adjusted Predictions at the Means).
    - All of the examples so far have used this
  - AAPs (Average Adjusted Predictions)
  - APRs (Adjusted Predictions at Representative values)

- For convenience, we will explain and illustrate each of these approaches as we discuss the corresponding ways of computing marginal effects

Marginal Effects

- As Cameron & Trivedi note (p. 333), "An ME [marginal effect], or partial effect, most often measures the effect on the conditional mean of y of a change in one of the regressors, say X_k. In the linear regression model, the ME equals the relevant slope coefficient, greatly simplifying analysis. For nonlinear models, this is no longer the case, leading to remarkably many different methods for calculating MEs."

- Marginal effects are popular in some disciplines (e.g. Economics) because they often provide a good approximation to the amount of change in Y that will be produced by a 1-unit change in X_k. With binary dependent variables, they offer some of the same advantages that the Linear Probability Model (LPM) does – they give you a single number that expresses the effect of a variable on P(Y=1).

- Personally, I find marginal effects for categorical independent variables easier to understand and also more useful than marginal effects for continuous variables
- The ME for categorical variables shows how P(Y=1) changes as the categorical variable changes from 0 to 1, after controlling in some way for the other variables in the model.
  - With a dichotomous independent variable, the marginal effect is the difference in the adjusted predictions for the two groups, e.g. for blacks and whites.
- There are different ways of controlling for the other variables in the model. We will illustrate how they work for both Adjusted Predictions & Marginal Effects.

```
. * Back to basic model
. logit diabetes i.black i.female age , nolog

Logistic regression                               Number of obs   =      10335
                                                  LR chi2(3)      =     374.17
                                                  Prob > chi2     =     0.0000
Log likelihood = -1811.9828                       Pseudo R2       =     0.0936

------------------------------------------------------------------------------
    diabetes |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     1.black |   .7179046   .1268061     5.66   0.000     .4693691      .96644
    1.female |   .1545569   .0942982     1.64   0.101    -.0302642    .3393779
         age |   .0594654   .0037333    15.93   0.000     .0521484    .0667825
       _cons |  -6.405437   .2372224   -27.00   0.000    -6.870384    -5.94049
------------------------------------------------------------------------------
```

APMs - Adjusted Predictions at the Means

```
. margins black female, atmeans

Adjusted predictions                            Number of obs   =      10335
Model VCE    : OIM

Expression   : Pr(diabetes), predict()
at           : 0.black         =    .8949202 (mean)
               1.black         =    .1050798 (mean)
               0.female        =    .4749879 (mean)
               1.female        =    .5250121 (mean)
               age             =    47.56584 (mean)

------------------------------------------------------------------------------
             |            Delta-method
             |     Margin   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       black |
           0 |   .0294328   .0020089    14.65   0.000     .0254955    .0333702
           1 |   .0585321   .0067984     8.61   0.000     .0452076    .0718566
             |
      female |
           0 |   .0292703   .0024257    12.07   0.000      .024516    .0340245
           1 |   .0339962   .0025912    13.12   0.000     .0289175    .0390748
------------------------------------------------------------------------------
```

MEMs – Marginal Effects at the Means

```
. * MEMs - Marginal effects at the means
. margins, dydx(black female) atmeans

Conditional marginal effects                      Number of obs   =      10335
Model VCE    : OIM

Expression   : Pr(diabetes), predict()
dy/dx w.r.t. : 1.black 1.female
at           : 0.black         =    .8949202 (mean)
               1.black         =    .1050798 (mean)
               0.female        =    .4749879 (mean)
               1.female        =    .5250121 (mean)
               age             =    47.56584 (mean)

------------------------------------------------------------------------------
             |            Delta-method
             |      dy/dx   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     1.black |   .0290993   .0066198     4.40   0.000     .0161246    .0420739
    1.female |   .0047259   .0028785     1.64   0.101    -.0009158    .0103677
------------------------------------------------------------------------------
Note: dy/dx for factor levels is the discrete change from the base level.
```

- The results tell us that, if you had two otherwise-average individuals, one white, one black, the black's probability of having diabetes would be 2.9 percentage points higher (Black APM = .0585, white APM = .0294, MEM = .0585 - .0294 = .029).

- And what do we mean by average? With APMs & MEMs, average is defined as having the mean value for the other independent variables in the model, i.e. 47.57 years old, 10.5 percent black, and 52.5 percent female.

- So, if we didn't have the margins command, we could compute the APMs and the MEM for race as follows. Just plug in the values for the coefficients from the logistic regression and the mean values for the variables other than race.

```
. * Replicate results for black without using margins
. scalar female_mean = .5250121
. scalar age_mean = 47.56584
. scalar wlogodds = _b[1.black]*0 + _b[1.female]*female_mean + _b[age]*age_mean + _b[_cons]
. scalar wodds = exp(wlogodds)
. scalar wapm = wodds/(1 + wodds)
. di "White APM = " wapm
White APM = .02943284

. scalar blogodds = _b[1.black]*1 + _b[1.female]*female_mean + _b[age]*age_mean + _b[_cons]
. scalar bodds = exp(blogodds)
. scalar bapm = bodds/(1 + bodds)
. di "Black APM = " bapm
Black APM = .05853209

. di "MEM for black = " bapm - wapm
MEM for black = .02909925
```

- MEMs are easy to explain. They have been widely used. Indeed, for a long time, MEMs were the only option with Stata, because that is all the old mfx command supported.

- But, many do not like MEMs. While there are people who are 47.57 years old, there is nobody who is 10.5 percent black or 52.5 percent female.

- Further, the means are only one of many possible sets of values that could be used – and a set of values that no real person could actually have seems troublesome.

- For these and other reasons, many researchers prefer AAPs & AMEs.

AAPs - Average Adjusted Predictions

```
. * Average Adjusted Predictions (AAPs)
. margins black female

Predictive margins                                      Number of obs   =      10335
Model VCE    : OIM

Expression   : Pr(diabetes), predict()

------------------------------------------------------------------------------
             |            Delta-method
             |     Margin   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       black |
          0  |   .0443248   .0020991    21.12   0.000     .0402107    .0484389
          1  |    .084417   .0084484     9.99   0.000     .0678585    .1009756
             |
      female |
          0  |   .0446799   .0029119    15.34   0.000     .0389726    .0503871
          1  |   .0514786    .002926    17.59   0.000     .0457436    .0572135
------------------------------------------------------------------------------
```

AMEs – Average Marginal Effects

```
. margins, dydx(black female)

Average marginal effects                        Number of obs   =      10335
Model VCE    : OIM

Expression   : Pr(diabetes), predict()
dy/dx w.r.t. : 1.black 1.female

------------------------------------------------------------------------------
             |            Delta-method
             |      dy/dx   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     1.black |   .0400922   .0087055     4.61   0.000     .0230297    .0571547
    1.female |   .0067987   .0041282     1.65   0.100    -.0012924    .0148898
------------------------------------------------------------------------------
Note: dy/dx for factor levels is the discrete change from the base level.
```

- Intuitively, the AME for black is computed as follows:
  - Go to the first case. Treat that person as though s/he were white, regardless of what the person's race actually is. Leave all other independent variable values as is. Compute the probability this person (if he or she were white) would have diabetes
  - Now do the same thing, this time treating the person as though they were black.
  - The difference in the two probabilities just computed is the marginal effect for that case
  - Repeat the process for every case in the sample
  - Compute the average of all the marginal effects you have computed. This gives you the AME for black.

```
. * Replicate AME for black without using margins
. clonevar xblack = black
. quietly logit diabetes i.xblack i.female age, nolog
. margins, dydx(xblack)

Average marginal effects                        Number of obs   =      10335
Model VCE    : OIM

Expression   : Pr(diabetes), predict()
dy/dx w.r.t. : 1.xblack

------------------------------------------------------------------------------
             |            Delta-method
             |      dy/dx   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    1.xblack |   .0400922   .0087055     4.61   0.000     .0230297    .0571547
------------------------------------------------------------------------------
Note: dy/dx for factor levels is the discrete change from the base level.

. replace xblack = 0
. predict adjpredwhite
. replace xblack = 1
. predict adjpredblack
. gen meblack = adjpredblack - adjpredwhite
. sum adjpredwhite adjpredblack meblack

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
adjpredwhite |     10335    .0443248    .0362422    .005399   .1358214
adjpredblack |     10335     .084417    .0663927   .0110063   .2436938
     meblack |     10335    .0400922    .0301892   .0056073   .1078724
```

- In effect, you are comparing two hypothetical populations – one all white, one all black – that have the exact same values on the other independent variables in the model.
- Since the only difference between these two populations is their race, race must be the cause of the differences in their likelihood of diabetes.
- Many people like the fact that all of the data is being used, not just the means, and feel that this leads to superior estimates.
- Others, however, are not convinced that treating men as though they are women, and women as though they are men, really is a better way of computing marginal effects.

- The biggest problem with both of the last two approaches, however, may be that they only produce a single estimate of the marginal effect. However “average” is defined, averages can obscure difference in effects across cases.

- In reality, the effect that variables like race have on the probability of success varies with the characteristics of the person, e.g. racial differences could be much greater for older people than for younger.

- If we really only want a single number for the effect of race, we might as well just estimate an OLS regression, as OLS coefficients and AMEs are often very similar to each other.

- APRs (Adjusted Predictions at Representative values) & MERs (Marginal Effects at Representative Values) may therefore often be a superior alternative.

- APRs/MERs can be both intuitively meaningful, while showing how the effects of variables vary by other characteristics of the individual.

- With APRs/MERs, you choose ranges of values for one or more variables, and then see how the marginal effects differ across that range.

APRs – Adjusted Predictions at Representative values

```
. * APRs - Adjusted Predictions at Representative Values (Race Only)
. margins black, at(age=(20 30 40 50 60 70)) vsquish

Predictive margins                                      Number of obs   =      10335
Model VCE    : OIM

Expression   : Pr(diabetes), predict()
1._at        : age             =          20
2._at        : age             =          30
3._at        : age             =          40
4._at        : age             =          50
5._at        : age             =          60
6._at        : age             =          70

------------------------------------------------------------------------------
             |            Delta-method
             |     Margin   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   _at#black |
        1 0  |   .0058698   .0009307     6.31   0.000     .0040457    .0076938
        1 1  |   .0119597   .0021942     5.45   0.000     .0076592    .0162602
        2 0  |   .0105876   .0013063     8.11   0.000     .0080273    .0131479
        2 1  |    .021466   .0033237     6.46   0.000     .0149517    .0279804
        3 0  |   .0190245   .0017157    11.09   0.000     .0156619    .0223871
        3 1  |   .0382346   .0049857     7.67   0.000     .0284628    .0480065
        4 0  |   .0339524   .0021105    16.09   0.000     .0298159    .0380889
        4 1  |   .0671983   .0075517     8.90   0.000     .0523972    .0819994
        5 0  |   .0598751   .0028793    20.79   0.000     .0542318    .0655184
        5 1  |   .1154567   .0118357     9.75   0.000     .0922591    .1386544
        6 0  |   .1034603   .0057763    17.91   0.000     .0921388    .1147817
        6 1  |   .1912405    .019025    10.05   0.000     .1539522    .2285289
------------------------------------------------------------------------------
```

MERs – Marginal Effects at Representative values

```
. margins, dydx(black female) at(age=(20 30 40 50 60 70)) vsquish

Average marginal effects                        Number of obs   =      10335
Model VCE    : OIM

Expression   : Pr(diabetes), predict()
dy/dx w.r.t. : 1.black 1.female
1._at        : age             =          20
2._at        : age             =          30
3._at        : age             =          40
4._at        : age             =          50
5._at        : age             =          60
6._at        : age             =          70

------------------------------------------------------------------------------
             |            Delta-method
             |      dy/dx   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
1.black      |
         _at |
          1  |   .0060899   .0016303     3.74   0.000     .0028946    .0092852
          2  |   .0108784   .0027129     4.01   0.000     .0055612    .0161956
          3  |   .0192101   .0045185     4.25   0.000     .0103541    .0280662
          4  |   .0332459   .0074944     4.44   0.000      .018557    .0479347
          5  |   .0555816   .0121843     4.56   0.000     .0317008    .0794625
          6  |   .0877803   .0187859     4.67   0.000     .0509606    .1245999
-------------+----------------------------------------------------------------
1.female     |
         _at |
          1  |   .0009933   .0006215     1.60   0.110    -.0002248    .0022114
          2  |     .00178   .0010993     1.62   0.105    -.0003746    .0039345
          3  |    .003161   .0019339     1.63   0.102    -.0006294    .0069514
          4  |   .0055253   .0033615     1.64   0.100     -.001063    .0121137
          5  |   .0093981   .0057063     1.65   0.100     -.001786    .0205821
          6  |   .0152754   .0092827     1.65   0.100    -.0029184    .0334692
------------------------------------------------------------------------------
Note: dy/dx for factor levels is the discrete change from the base level.
```

- Earlier, the AME for black was 4 percent, i.e. on average blacks' probability of having diabetes is 4 percentage points higher than it is for whites.

- But, when we estimate marginal effects for different ages, we see that the effect of black differs greatly by age. It is less than 1 percentage point for 20 year olds and almost 9 percentage points for those aged 70.

- Similarly, while the AME for gender was only 0.6 percent, at different ages the effect is much smaller or much higher than that.

- In a large model, it may be cumbersome to specify representative values for every variable, but you can do so for those of greatest interest.
  - For other variables you have to set them to their means, or use average adjusted predictions, or use some other approach.

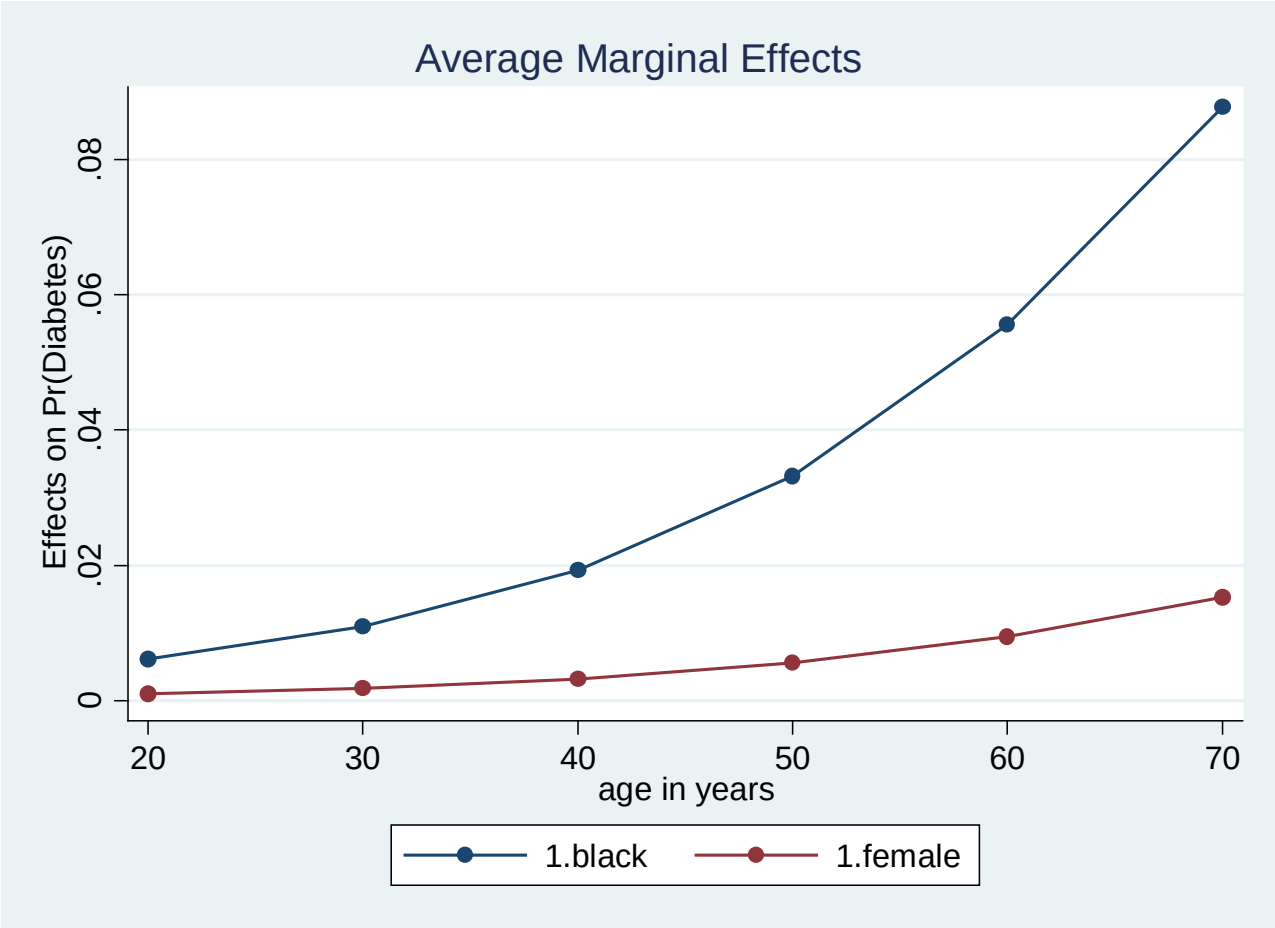
Graphing results

- The output from the margins command can be very difficult to read. It can be like looking at a 5 dimensional crosstab where none of the variables have value labels
- The marginsplot command introduced in Stata 12 makes it easy to create a visual display of results.

```
. quietly logit diabetes i.black i.female age, nolog

. quietly margins, dydx(black female) at(age=(20 30 40 50 60 70)) vsquish

. marginsplot, noci
```


Average Marginal Effects
Effects on Pr(Diabetes)
0
.02
.04
.06
.08
20
30
40
50
60
70
age in years
1.black
1.female

A more complicated example

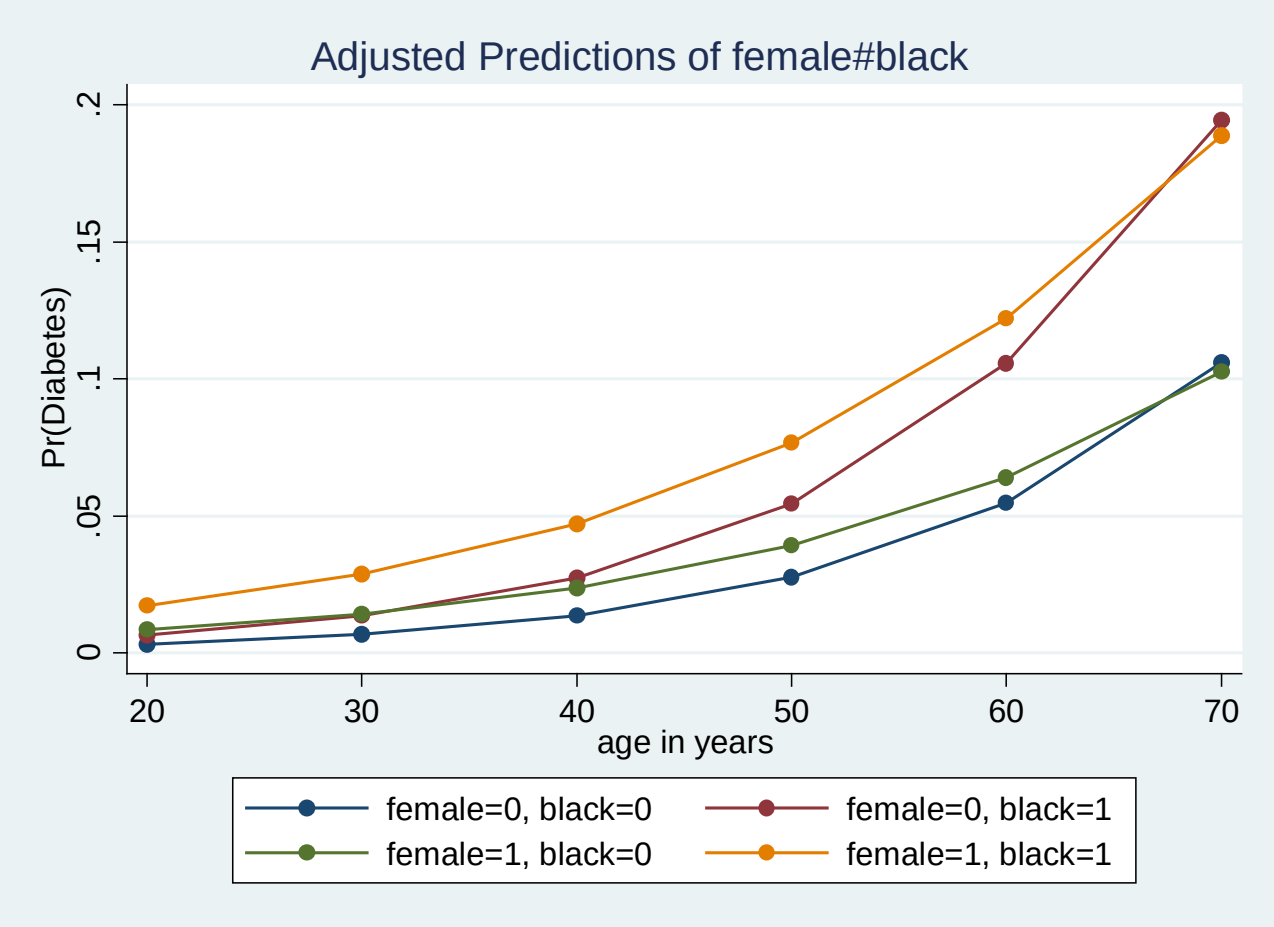
```
. quietly logit diabetes i.black i.female age i.female#c.age, nolog
. margins female#black, at(age=(20 30 40 50 60 70)) vsquish

Adjusted predictions                            Number of obs   =      10335
Model VCE    : OIM

Expression   : Pr(diabetes), predict()
1._at        : age             =          20
2._at        : age             =          30
3._at        : age             =          40
4._at        : age             =          50
5._at        : age             =          60
6._at        : age             =          70

------------------------------------------------------------------------------
                 |            Delta-method
                 |     Margin   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----------------+------------------------------------------------------------
_at#female#black |
          1 0 0  |    .003304      .0009     3.67   0.000       .00154    .0050681
          1 0 1  |    .006706   .0019396     3.46   0.001     .0029044    .0105076
          1 1 0  |   .0085838    .001651     5.20   0.000      .005348    .0118196
          1 1 1  |   .0173275   .0036582     4.74   0.000     .0101576    .0244974
          2 0 0  |   .0067332   .0014265     4.72   0.000     .0039372    .0095292
          2 0 1  |   .0136177   .0031728     4.29   0.000     .0073991    .0198362
          2 1 0  |   .0143006   .0021297     6.71   0.000     .0101264    .0184747
          2 1 1  |    .028699   .0049808     5.76   0.000     .0189368    .0384613
          3 0 0  |   .0136725   .0020998     6.51   0.000     .0095569    .0177881
          3 0 1  |   .0274562   .0049771     5.52   0.000     .0177013     .037211
          3 1 0  |   .0237336   .0025735     9.22   0.000     .0186896    .0287776
          3 1 1  |   .0471751   .0066696     7.07   0.000     .0341029    .0602473
          4 0 0  |   .0275651   .0028037     9.83   0.000       .02207    .0330603
          4 0 1  |   .0545794   .0075901     7.19   0.000     .0397031    .0694557
          4 1 0  |   .0391418   .0029532    13.25   0.000     .0333537    .0449299
          4 1 1  |   .0766076   .0090659     8.45   0.000     .0588388    .0943764
          5 0 0  |   .0547899   .0038691    14.16   0.000     .0472066    .0623733
          5 0 1  |   .1055879   .0121232     8.71   0.000     .0818269    .1293489
          5 1 0  |   .0638985   .0039287    16.26   0.000     .0561983    .0715986
          5 1 1  |   .1220509   .0131903     9.25   0.000     .0961985    .1479034
          6 0 0  |   .1059731   .0085641    12.37   0.000     .0891878    .1227584
          6 0 1  |   .1944623   .0217445     8.94   0.000     .1518439    .2370807
          6 1 0  |   .1026408   .0075849    13.53   0.000     .0877747    .1175069
          6 1 1  |   .1889354   .0206727     9.14   0.000     .1484176    .2294532
------------------------------------------------------------------------------

. marginsplot, noci
```


Adjusted Predictions of female#black
Pr(Diabetes)
0
.05
.1
.15
.2
20
30
40
50
60
70
age in years
female=0, black=0
female=0, black=1
female=1, black=0
female=1, black=1

Marginal effects of interaction terms

- People often ask what the marginal effect of an interaction term is. Stata's margins command replies: there isn't one. You just have the marginal effects of the component terms. The value of the interaction term can't change independently of the values of the component terms, so you can't estimate a separate effect for the interaction.

```
. quietly logit diabetes i.black i.female age i.female#c.age, nolog
. margins, dydx(*)

Average marginal effects                        Number of obs   =      10335
Model VCE    : OIM

Expression   : Pr(diabetes), predict()
dy/dx w.r.t. : 1.black 1.female age

------------------------------------------------------------------------------
             |            Delta-method
             |      dy/dx   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     1.black |   .0396176   .0086693     4.57   0.000      .022626    .0566092
    1.female |   .0067791   .0041302     1.64   0.101     -.001316    .0148743
         age |   .0026632   .0001904    13.99   0.000     .0022901    .0030364
------------------------------------------------------------------------------
Note: dy/dx for factor levels is the discrete change from the base level.
```

- For more on marginal effects and interactions, See Vince Wiggins' excellent discussion at

  http://www.stata.com/statalist/archive/2013-01/msg00293.html

A few other points

- Margins would also give the wrong answers if you did not use factor variables. You should use margins because older commands, like adjust and mfx, do not support the use of factor variables

- Margins supports the use of the svy: prefix with svyset data. Some older commands, like adjust, do not.

- With older versions of Stata, margins is, unfortunately, more difficult to use with multiple-outcome commands like ologit or mlogit. But this is also true of many older commands like adjust. Stata 14 made it much easier to use margins with multiple outcome commands.

- In the past the xi: prefix was used instead of factor variables. In most cases, *do not use xi: anymore*. The output from xi: looks horrible. More critically, the xi: prefix will cause the same problems that computing dummy variables yourself does, i.e. margins will not know how variables are inter-related.

- Long & Freese's spost13 commands were rewritten to take advantage of margins. Commands like mtable and mchange basically make it easy to execute several margins commands at once and to format the output. From within Stata type `findit spost13_ado`. Their highly recommended book can be found at

  http://www.stata.com/bookstore/regression-models-categorical-dependent-variables/

- Patrick Royston's mcp command (available from SSC) provides an excellent means for using margins with continuous variables and graphing the results. From within Stata type findit mcp. For more details see

  http://www.stata-journal.com/article.html?article=gr0056

References

- Williams, Richard. 2012. “Using the margins command to estimate and interpret adjusted predictions and marginal effects.” The Stata Journal 12(2):308-331.
- Available for free at

  http://www.stata-journal.com/article.html?article=st0260

- This handout is adapted from the article. The article includes more information than in this presentation. However, this presentation also includes some additional points that were not in the article.
- Please cite the above article if you use this material in your own research.