We will use real-world data, but a very simple and naive model to keep the example easy to understand. What is interesting about the example is that the outcome of interest, perhaps the probability or alternately the odds of a salary increase (`raise`), has an inherently nonlinear relationship with any covariate. What's more, `age` enters the model in a nonlinear fashion so that interpreting its impact on the outcome or outcomes of interest becomes even more difficult.

First, let's create the reciprocal of age variable and the dependent variable we will be analyzing – `raise`. We will derive it from the wage variable. We are actually losing information in creating the binary raise variable from the continuous wage, but out goal is to create and interesting model to analyze rather than a reasonable one.

```
. use nls , clear
(National Longitudinal Survey.  Young Women 14-26 years of age in 1968)

.
. drop if age >= .
(24 observations deleted)

. gen age1 = 1 / age

.
. sort idcode year

. gen raise = wage > wage[_n-1] & idcode == idcode[_n-1]
```

We can now estimate the probit model.

```
. probit raise age age1 collgrad

Iteration 0:   log likelihood = -19761.451
Iteration 1:   log likelihood = -19292.141
Iteration 2:   log likelihood = -19290.786
Iteration 3:   log likelihood = -19290.786

Probit estimates                                  Number of obs   =      28510
                                                  LR chi2(3)      =     941.33
                                                  Prob > chi2     =     0.0000
Log likelihood = -19290.786                       Pseudo R2       =     0.0238

------------------------------------------------------------------------------
       raise |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |  -.0686738   .0049641   -13.83   0.000    -.0784033   -.0589444
        age1 |  -77.45967   3.952097   -19.60   0.000    -85.20564    -69.7137
    collgrad |  -.1222773   .0203146    -6.02   0.000    -.1620933   -.0824614
       _cons |   4.822641   .2860679    16.86   0.000     4.261958    5.383324
------------------------------------------------------------------------------
```

At this point there are a number of things we might do to analyze our model. We might perform tests of linear and nonlinear combinations of the coefficients using `test` and `testnl`. We might look directly at marginal effects of the regressors on the probability of a salary increase using `mfx`. Or, we might use any of the other post estimation facilities discussed earlier.

Many of these analyses would involve simply typing a command and looking at the output. Instead, let's do some analysis that requires us to combine one of the post-estimation tools `predictnl` with some data manipulation. `predictnl` is extremely flexible and as we will see allows us to mimic some of the results from other post-estimation commands, with increased flexibility.

Looking at the probit results, it is clear that interpreting the effect of age is rather difficult. Let's focus on interpreting that effect in ways that an audience unfamiliar with probit could understand.

We know that the effect of age is nonlinear, so if we are interested in its changing effect over a range of different ages, it is clear that we cannot look at just one age. What we would like to see is the effect of age evaluated over the range of ages observed in our dataset. Recalling that `predictnl` can work on datasets other than the estimation dataset, we can create an artificial dataset with observations at a set of ages – say ages by single year of age from 18 through 47. Further, let's set the value of the college graduate indicator to 1 so that we are looking at the effect of age for college graduates. We could just as easily have set the indicator to 0 and looked an non-college graduates, or even to the proportion of college graduates and looked that the effect of age for the overall sample.

We will just clear our current sample and create our artificial dataset.

```
. drop _all

. set obs 30
obs was 0, now 30

. gen collgrad = 1

. gen age = _n + 17

. gen age1 = 1 / age

. gen raise = .
(30 missing values generated)
```

There are a number of ”outcomes” that we might find interesting from our binary dependent variable regression. Perhaps the easiest to understand is the probability that someone will get an increase in salary from one year to the next. In turns out the probability is something that `predict` after `probit` is willing to produce. What's more `predictnl` is willing to use anything that can be produced by `predict` as part of its nonlinear expression. All we need do is include the term `predict`(*statistic*) in our expression, where *statistic* is any statistic that `predict` will produce after our estimator.

So, we evaluate the effect of age at each observation in our artificial dataset by typing,

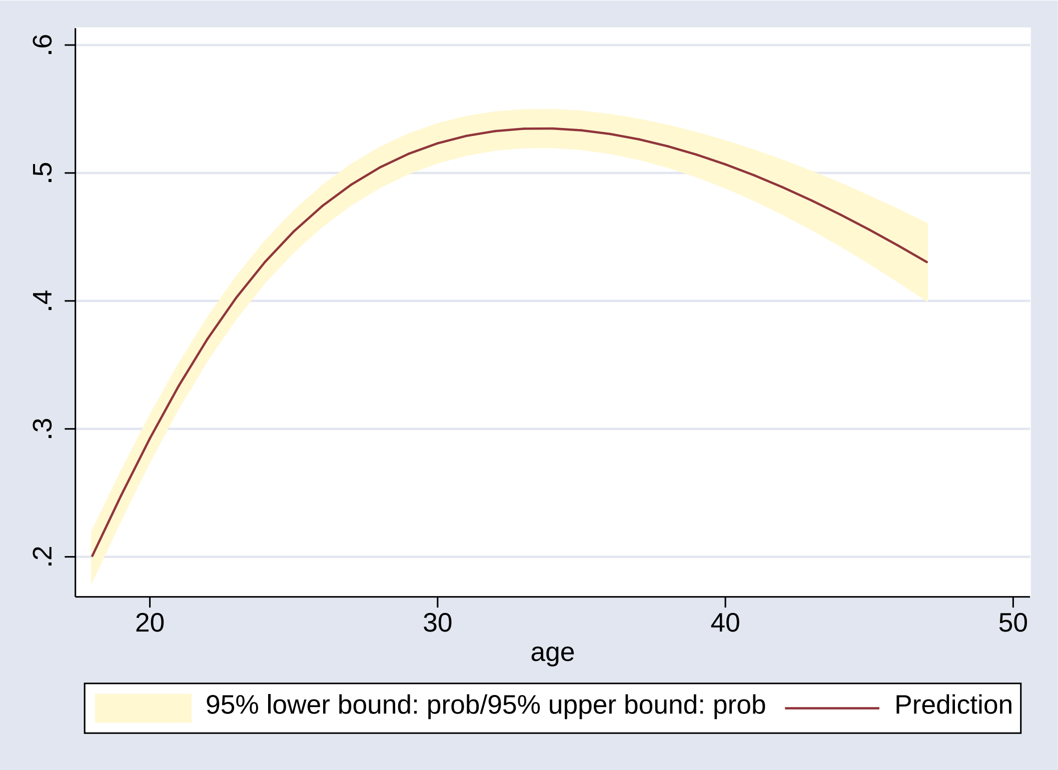
The option `ci(probl probh)` requests that in addition to computing predicted probabilities, `predictnl` should also produce lower and upper confidence bounds (CIs) for these predictions. Since we did not specify a confidence level, it is assumed to be 95method where the required derivatives are taken numerically, but that is really a side note.

We can see the results of our efforts, but just listing the probabilities and their CIs for all of the ages.

What we are seeing is not so much a set of predictions as an alternate representation of our model, and one that we can hopefully interpret more easily. If we had estimated our model in the probability metric rather than the probit metric (where the coefficients represent one standard deviation changes in our underlying latent z variable), these probabilities and their CIs are the age ”coefficients” we would have observed.

The table is interesting and nice if we really need a probability for a specific age, a graph will, however, let us see the overall relationship more clearly.

```
. twoway rarea probl probh age, p(ci2) || line prob age , ytitle("")
```

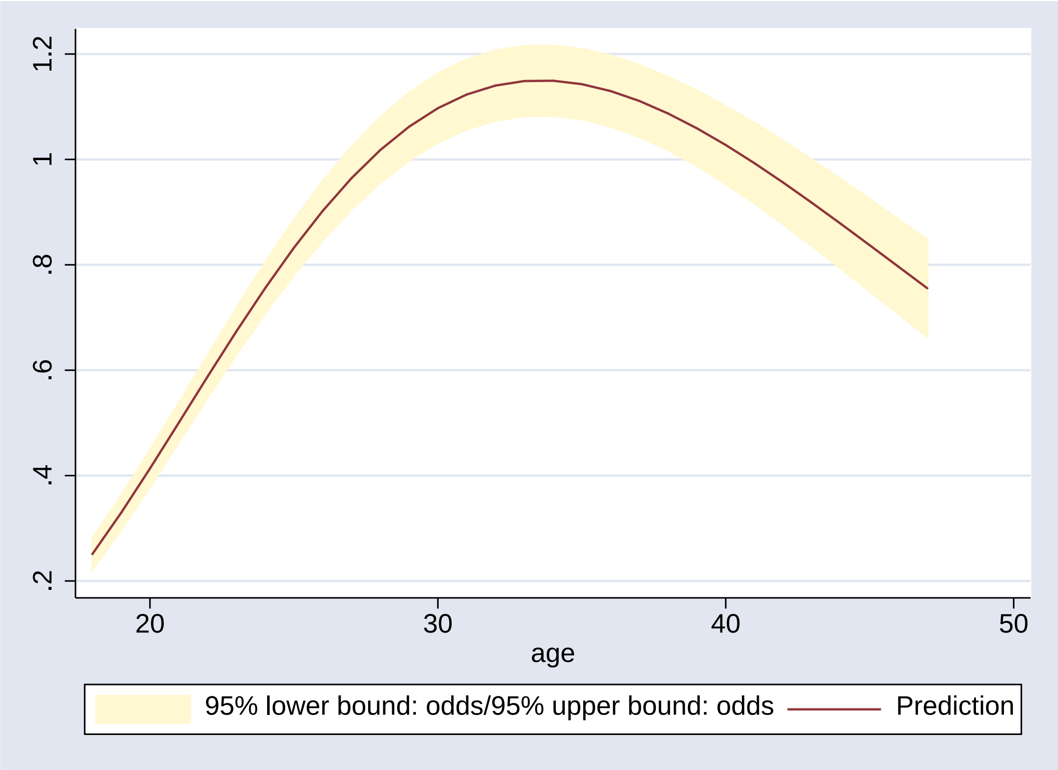


We see that our estimates show the annual probability of a wage increase begins at about .2 for 18 year-olds (though 18 year-old college graduates may be nearly counterfactuals) and increases to about .52 for those age 33, then begins to decrease, but at a rate slower than the rate of increase.

To keep things simple we used a basic graph command to look at our probability ”coefficients”, this graph could obviously benefit from some improved labeling, but we won’t bother here.

Rather than the probability of a salary increase, we might consider the odds of an increase. The odds are just a nonlinear function of the probability $odds = p/(1-p)$ where p is the probability. Similarly to the probability, we use `predictnl` to compute the odds and the CI of the odds, and then graph the result.

```
. predictnl odds = predict(p) / (1 - predict(p)) , ci(oddsl oddsh)
note: Confidence intervals calculated using Z critical values

. twoway rarea oddsl oddsh age, p(ci2) || line odds age , ytitle("")
```

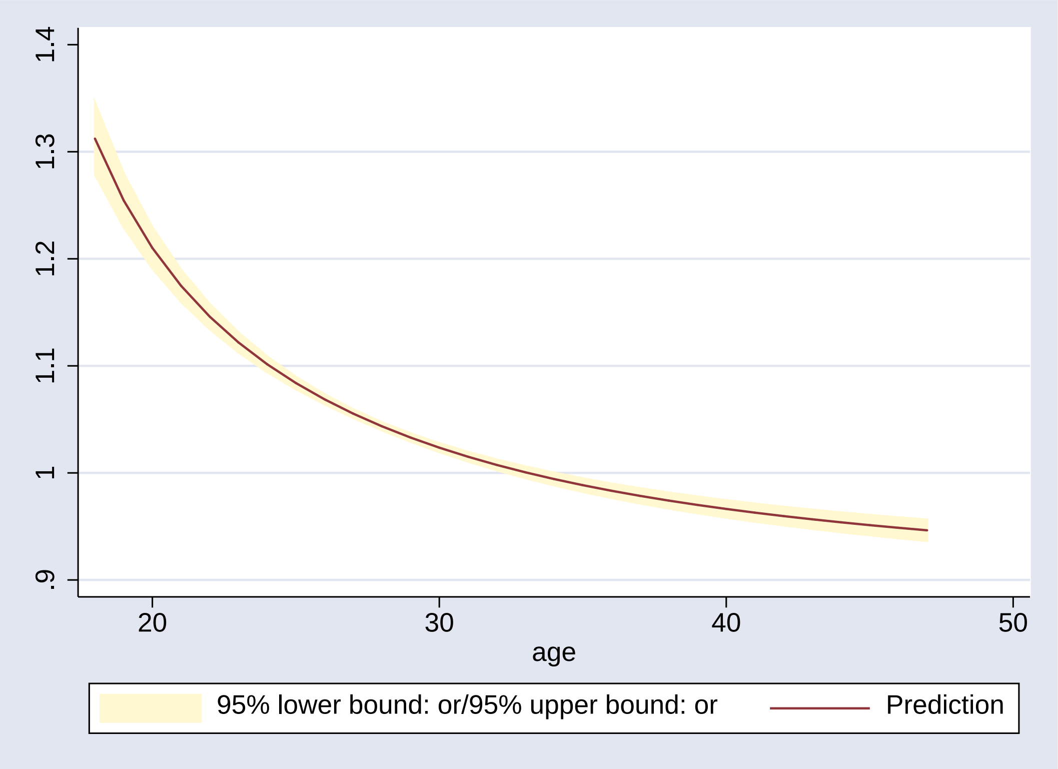


We see that the odds are a similar but somewhat differently shaped function of age. We can clearly see that, according to our model, women college graduates are more likely to get raises than not from about age 27 through age 42. That is to say their odds of a wage increase exceed 1.

Biostatisticians and epidemiologists in particular often like to work with odds ratios – the change in the odds for a unit change in a covariate. We noted earlier that native metric a probit model is that a coefficient represent a one standard deviation change in the latent metric. Similarly the native metric of a logistic (or logit) model is that a coefficient represents a unit change in the odds ratio. For probit the change in the standard deviation is constant over all values of the regressor and all other coefficients while for logistic regression the change is constant for the odds ratio. One result of this is that even if age did not enter the model nonlinearly, if we were interested in odds ratios after estimating a probit model, these odds ratios would have a nonlinear relationship to the variable and they would also depend on the values of all of the other variables. This is also what we observed for probabilities and odds, which are nonlinear relationships of the coefficients for both the logistic and probit models.

We will ”cheat” just a little and compute a discrete form of the odds ratio where we assume a whole unit increase in the covariate, rather than computing the continuous form that requires a bit more manipulation. Here is the computation and resulting graph.

```
. predictnl or = (normprob(xb()+_b[age]-_b[age1]/(age*(age+1))) /         ///
>                 (1 - normprob(xb()+_b[age]-_b[age1]/(age*(age+1))))) /  ///
>                (normprob(xb()) / (1 - normprob(xb())))                  ///
>                 , ci(orl orh)
note: Confidence intervals calculated using Z critical values

. twoway rarea orl orh age, p(ci2) || line or age , ytitle("")
```



Basically this is just one plus the rate of change in the odds w.r.t age. Looking at the graph of the odds ratio, we see that the ratio is monotonically declining – meaning that the effect age on the odds of obtaining a raise becomes becomes more negative with increasing age. We might say that the acceleration of age is negative. Looking at both the odds ratio and odds graphs we see that the odds ratio crosses 1.0 at the same age as the odds switches from increasing to decreasing.

If age were a policy variable that we could change, we might be very interested in the change in probability of an increase for a unit change in age – the marginal effect of age on probability. Clearly age is not a policy variable we can change, but let's ignore that detail.

The expression for the marginal effect of age is fairly easy to derive and we can see it in the `predictnl` expression below.

```
. predictnl mfx = (_b[age] - _b[age1] / age^2) * normden(predict(xb))     ///
>                 , ci(mfxl mfxh) se(se) wald(wald) p(pvalue)
note: significance levels are with respect to the chi-squared(1) distribution
note: Confidence intervals calculated using Z critical values
```

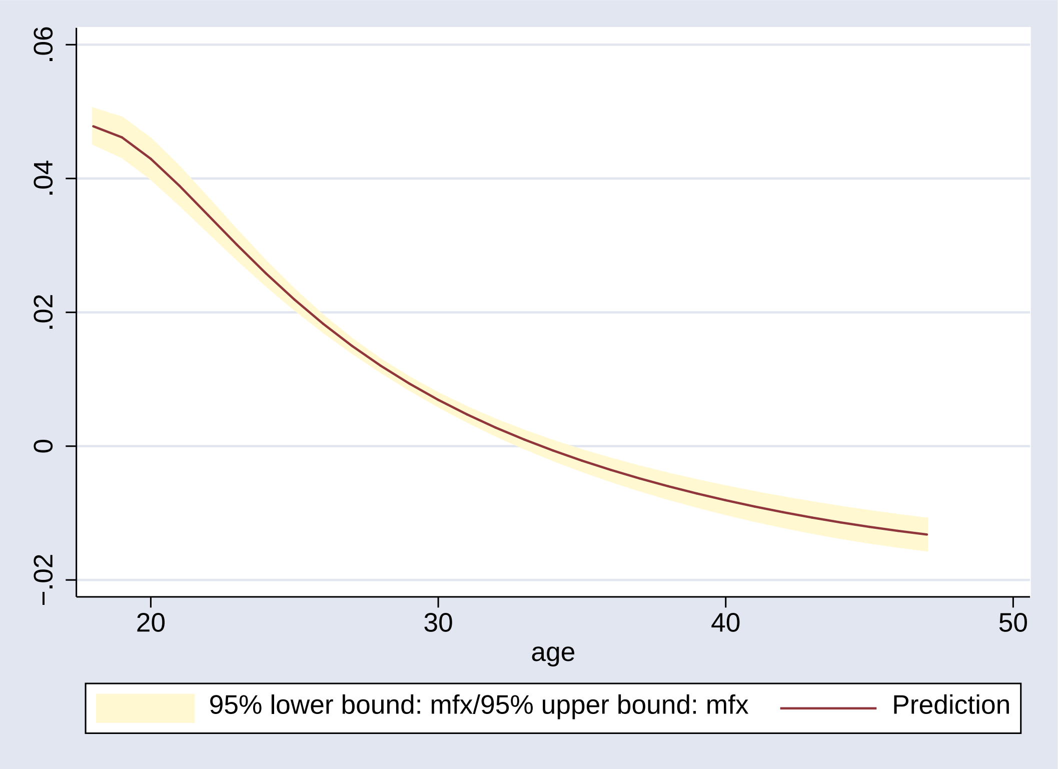
Note that we predictnl to compute a number of additional statistics for us – the standard error of the marginal effect, the Wald statistic against the null hypothesis that the marginal effect is 0, and the p-value of this test. In other words, we have asked for all of the ingredients of an estimation coefficient table. Let's now list those,

```
. format pvalue %6.3f

. list age mfx se wald pvalue mfxl mfxh

     +------------------------------------------------------------------+
     | age         mfx         se        wald   pvalue        mfxl        mfxh |
     |------------------------------------------------------------------|
  1. |  18    .0478073   .0013407    1271.603    0.000    .0451796    .0504349 |
  2. |  19    .0461499    .001493    955.4506    0.000    .0432236    .0490762 |
  3. |  20    .0429513   .0015133    805.6086    0.000    .0399854    .0459173 |
  4. |  21    .0389013   .0014245    745.8185    0.000    .0361095    .0416932 |
  5. |  22     .034503   .0012704    737.6323    0.000    .0320131    .0369929 |
     |------------------------------------------------------------------|
  6. |  23    .0300873   .0010888    763.5726    0.000    .0279533    .0322214 |
  7. |  24     .025852   .0009067    812.9969    0.000     .024075    .0276291 |
  8. |  25    .0219012    .000742    871.3121    0.000     .020447    .0233555 |
  9. |  26    .0182781   .0006074    905.6754    0.000    .0170877    .0194685 |
 10. |  27    .0149887   .0005127     854.595    0.000    .0139838    .0159937 |
     |------------------------------------------------------------------|
 11. |  28    .0120181   .0004643    669.9606    0.000    .0111081    .0129281 |
 12. |  29    .0093407   .0004602    411.9077    0.000    .0084387    .0102428 |
 13. |  30    .0069269   .0004894    200.2936    0.000    .0059676    .0078862 |
 14. |  31    .0047465   .0005384    77.71658    0.000    .0036913    .0058018 |
 15. |  32    .0027714    .000597    21.54703    0.000    .0016012    .0039416 |
     |------------------------------------------------------------------|
 16. |  33    .0009759   .0006594    2.190479    0.139   -.0003164    .0022682 |
 17. |  34   -.0006626   .0007221    .8418617    0.359    -.002078    .0007528 |
 18. |  35   -.0021632   .0007837    7.618898    0.006   -.0036993   -.0006272 |
 19. |  36   -.0035424   .0008431    17.65299    0.000   -.0051949   -.0018899 |
 20. |  37   -.0048138   .0008998    28.62359    0.000   -.0065773   -.0030503 |
     |------------------------------------------------------------------|
 21. |  38   -.0059884   .0009531    39.47705    0.000   -.0078565   -.0041204 |
 22. |  39   -.0070755   .0010027    49.79154    0.000   -.0090407   -.0051102 |
 23. |  40    -.008082   .0010481    59.46309    0.000   -.0101362   -.0060278 |
 24. |  41   -.0090137   .0010887    68.54877    0.000   -.0111475   -.0068799 |
 25. |  42   -.0098749    .001124    77.18836    0.000   -.0120778   -.0076719 |
     |------------------------------------------------------------------|
 26. |  43   -.0106685   .0011533    85.56606    0.000    -.012929   -.0084081 |
 27. |  44    -.011397   .0011762    93.89397    0.000   -.0137023   -.0090918 |
 28. |  45   -.0120618   .0011919     102.408    0.000   -.0143979   -.0097257 |
 29. |  46   -.0126636      .0012    111.3723    0.000   -.0150155   -.0103118 |
 30. |  47   -.0132032   .0011998    121.0904    0.000   -.0155548   -.0108515 |
     +------------------------------------------------------------------+
```

What we see is the marginal effect at each age along with all of the statistics commonly reported in an estimation coefficient table. This is a particularly nice representation of the marginal effect, because they are indeed derived ”coefficients” from our estimation. We can see that at age 18 the marginal effect of increasing age by 1 year is to increase the probability of obtaining a salary increase by about 4.78%. We also note that this effect is very significantly different from 0 and has a 955.00%. Likewise we can read off the marginal effect for any age.

The marginal effect looks to be decreasing each year and eventually moves from positive to negative. Again, we can get a feel for the whole relationship by graphing the marginal effect against age.

```
. twoway rarea mfxl mfxh age, p(ci2) || line mfx age , ytitle("")
```



Like the odds ratio, the marginal effect is monotonically decreasing and intersects zero at the point where the probability of a raise begins to decrease. The marginal effect and the odds ratio are closely related, one being the rate of change in the probability w.r.t. age and the other one plus the percentage change in the odds w.r.t. age.

What if we were interested in examining the along both the age and college graduate dimensions? As seen in the steps below we could again start with an empty dataset and create two stacked sets of observations by age where one set has `collgrad` set to 0 and the other set to 1. The only real trick in this code is the use of fillin to ”fill in” all possible combinations of `age` and `collgrad`.

```
. drop _all
. set obs 30
obs was 0, now 30
. gen     collgrad = 1
. replace collgrad = 0  in 1
(1 real change made)
. gen age = _n + 17
. fillin collgrad age
. gen age1 = 1 / age
. gen raise = .
(60 missing values generated)
```

We could now use exactly the same `predictnl` statements to create our probabilities, odds, odds ratios, and marginal effects, only now they will be created for college graduates and non-graduates.

```
. predictnl prob = predict() , ci(plow phigh)
note: Confidence intervals calculated using Z critical values
```

```
. predictnl odds = predict() / (1 - predict()) , ci(oddsl oddsh)
note: Confidence intervals calculated using Z critical values
. predictnl or = (normprob(xb()+_b[age]-_b[age1]/(age*(age+1))) /         ///
>                 (1 - normprob(xb()+_b[age]-_b[age1]/(age*(age+1))))) /  ///
>                (normprob(xb()) / (1 - normprob(xb())))                  ///
>                 , ci(orl orh)
note: Confidence intervals calculated using Z critical values
. predictnl mfx = (_b[age] - _b[age1] / age^2) * normden(predict(xb))     ///
>                 , ci(mfxl mfxh)
note: Confidence intervals calculated using Z critical values
```

We could again list the data, but instead we will use the `tabdisp` command to produce a twoway table of annual probability of a salary increase by age and college graduation.

```
. label variable plow ""
. label variable phigh ""
. tabdisp age collgrad , cell(prob)

----------------------------------
          |        collgrad
      age |         0          1
----------+-----------------------
       18 |  .2367478   .2007121
       19 |  .2880854   .2478523
       20 |  .3358488   .2924998
       21 |  .3790163   .3334732
       22 |  .4171939   .3701882
       23 |  .4503928   .4024749
       24 |  .4788591   .4304241
       25 |  .5029556   .4542746
       26 |  .5230871   .4743362
       27 |  .5396575   .4909422
       28 |  .5530444     .50442
       29 |  .5635898   .5150763
       30 |  .5715953   .5231893
       31 |  .5773245   .5290079
       32 |  .5810045   .5327508
       33 |  .5828313   .5346105
       34 |  .5829731   .5347549
       35 |  .5815747   .5333312
       36 |  .5787613   .5304688
       37 |  .5746419   .5262822
       38 |  .5693122   .5208735
       39 |  .5628569   .5143345
       40 |  .5553524   .5067493
       41 |  .5468687   .4981955
       42 |  .5374702   .4887454
       43 |  .5272182   .4784682
       44 |  .5161709     .46743
       45 |  .5043854   .4556954
       46 |  .4919177   .4433275
       47 |  .4788234   .4303889
----------------------------------
```

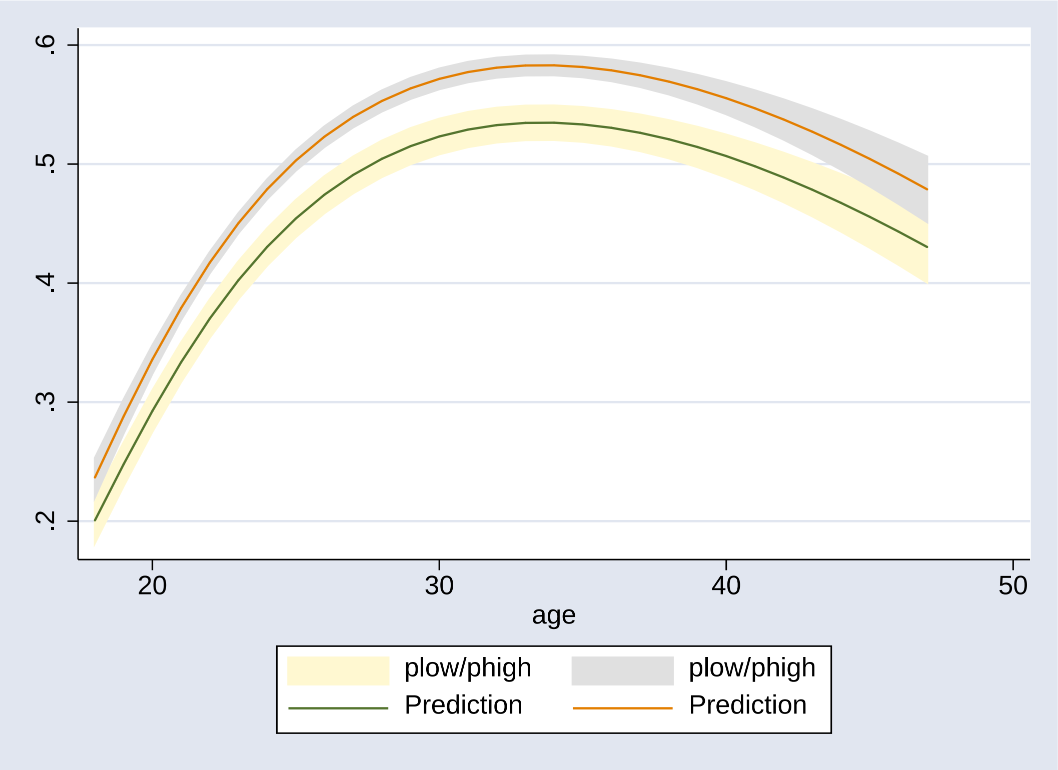
Or, we could include the CIs in the table.

```
. tabdisp age collgrad , cell(prob plow phigh)

------------------------------
          |      collgrad
      age |        0         1
----------+-------------------
       18 | .2367478  .2007121
          | .2205016  .1819141
          |  .252994  .2195101
          |
       19 | .2880854  .2478523
          |  .274085  .2297721
          | .3020858  .2659324
          |
       20 | .3358488  .2924998
          | .3241054  .2752971
          | .3475923  .3097025
          |
       21 | .3790163  .3334732
          | .3691117  .3170219
          |  .388921  .3499245
          |
       22 | .4171939  .3701882
          | .4084745  .3542495
          | .4259132  .3861269
          |
       23 | .4503928  .4024749
          |    .4422  .3868231
          | .4585857  .4181268
          |
       24 | .4788591  .4304241
          | .4707236  .4149063
          | .4869946  .4459419
          |
       25 | .5029556  .4542746
          | .4946519  .4388186
          | .5112591  .4697305
          |
       26 | .5230871  .4743362
          | .5145723   .458932
          |  .531602  .4897403
          |
       27 | .5396575  .4909422
          | .5309871  .4756165
          | .5483279  .5062679
          |
       28 | .5530444    .50442
          | .5443122  .4892144
          | .5617766  .5196257
          |
       29 | .5635898  .5150763
          | .5548908  .5000303
          | .5722888  .5301222
          |
       30 | .5715953  .5231893
          |  .563004  .5083292
          | .5801865  .5380494
          |
       31 | .5773245  .5290079
          | .5688801  .5143371
          | .5857687  .5436786
          |
32 | .5810045  .5327508
   | .5726995  .5182443
   | .5893094  .5472573
   |
33 | .5828313  .5346105
   | .5746014  .5202096
   | .5910611  .5490112
   |
34 | .5829731  .5347549
   |  .574692  .5203651
   | .5912542  .5491446
   |
35 | .5815747  .5333312
   | .5730563   .518822
   | .5900932  .5478404
   |
36 | .5787613  .5304688
   | .5697733  .5156769
   | .5877494  .5452608
   |
37 | .5746419  .5262822
   | .5649276  .5110185
   | .5843562  .5415459
   |
38 | .5693122  .5208735
   | .5586137  .5049331
   | .5800107  .5368139
   |
39 | .5628569  .5143345
   |  .550932  .4975072
   | .5747817  .5311619
   |
40 | .5553524  .5067493
   | .5419835  .4888298
   | .5687214  .5246689
   |
41 | .5468687  .4981955
   | .5318651  .4789915
   | .5618724  .5173994
   |
42 | .5374702  .4887454
   | .5206671  .4680835
   | .5542733  .5094073
   |
43 | .5272182  .4784682
   | .5084753  .4561969
   | .5459611  .5007395
   |
44 | .5161709    .46743
   | .4953703  .4434217
   | .5369716  .4914384
   |
45 | .5043854  .4556954
   | .4814302  .4298465
   | .5273407  .4815443
   |
46 | .4919177  .4433275
   | .4667315  .4155592
   | .5171039  .4710958
   |
47 | .4788234  .4303889
```

```
          |  .4513499   .4006465
          |  .5062969   .4601313
------------------------------
```

Finally, we can compare the two probability profiles on a single graph.

```
. twoway  rarea plow phigh age    if collgrad==1, p(ci2)  ||              ///
>         rarea plow phigh age    if collgrad==0, p(ci)   ||              ///
>         line prob age           if collgrad==1          ||              ///
>         line prob age           if collgrad==0 , ytitle("")
```



We have ignored the odds, odds ratios, and marginal effects in our twoway analysis, though we could have tabulated and graphed them also.

We have only scratched the surface of what we might do with `predictnl` and the other post-estimation commands. We have focused on various transforms of the dependent variable, but could just as easily focused on transforms of one or more of the independent variables. Still, hopefully we have gotten a sense of how interesting questions can be approached.