Adjusted Predictions & Marginal Effects for Multiple Outcome Models & Commands (including ologit, mlogit, oglm, & gologit2)

Richard Williams, University of Notre Dame, https://www3.nd.edu/~rwilliam/
Last revised January 25, 2021

As was the case with logit models, the parameters for an ordered logit model and other multiple outcome models can be hard to interpret. Adjusted predictions and marginal effects can again make results more understandable. Stata 14 made the `margins` command much easier to use after multiple outcome commands like `ologit`, `oprobit`, `mlogit`, `oglm` and `gologit2`. While the examples here use `ologit`, the same procedures can be used with other commands. If you are condemned to using Stata 12 or 13, the appendix describes the much more tedious process that is required. However, regardless of which version of Stata you are using, you may prefer to use spost13 commands (`findit spost13_ado`) like `mtable` as they can produce easier to read output.

Consider the following. We will once again use the Nhanes2f data, but this time the dependent variable will be self-reported health, whose values range between 1 (poor) and 5 (excellent):

```
. version 14.1
. webuse nhanes2f, clear
. keep if !missing(diabetes, black, female, age)
(2 observations deleted)
. label define black 0 "nonBlack" 1 "black"
. label define female 0 "male" 1 "female"
. label values black black
. label values female female
. fre health

health -- 1=poor,..., 5=excellent
-----------------------------------------------------------------
                    |      Freq.    Percent      Valid       Cum.
--------------------+--------------------------------------------
Valid   1 poor      |        729       7.05       7.05       7.05
        2 fair      |       1670      16.16      16.16      23.21
        3 average   |       2938      28.43      28.43      51.64
        4 good      |       2591      25.07      25.07      76.71
        5 excellent |       2407      23.29      23.29     100.00
        Total       |      10335     100.00     100.00           
-----------------------------------------------------------------

. ologit health i.female i.black c.age, nolog

Ordered logistic regression                     Number of obs     =      10335
                                                LR chi2(3)        =    1682.10
                                                Prob > chi2       =     0.0000
Log likelihood = -14923.345                     Pseudo R2         =     0.0534

------------------------------------------------------------------------------
      health |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      female |
     female  |  -.1170992   .0355732    -3.29   0.001    -.1868215   -.0473769
             |
       black |
      black  |  -.8845093   .0583106   -15.17   0.000     -.998796   -.7702227
         age |  -.0410673   .0010907   -37.65   0.000     -.043205   -.0389295
-------------+----------------------------------------------------------------
       /cut1 |  -4.910859   .0743281                     -5.056539   -4.765179
       /cut2 |  -3.428162   .0648869                     -3.555338   -3.300986
       /cut3 |  -2.004318   .0586634                     -2.119296    -1.88934
       /cut4 |  -.7512595   .0561222                     -.8612569   -.6412621
```

The results tell us that, on an all other things being equal basis, females, blacks, and older people tend to have lower levels of self-reported health. However, other than sign and significance, it is difficult to get a tangible feel for how large and important these differences are.

Using the Margins and Spost13 Commands. Now let's see what happens when we use the `margins` command to get the AAPs (Average Adjusted Predictions) for whites and blacks:

```
. * AAPs using margins
. margins black

Predictive margins                              Number of obs     =     10,335
Model VCE    : OIM

1._predict   : Pr(health==1), predict(pr outcome(1))
2._predict   : Pr(health==2), predict(pr outcome(2))
3._predict   : Pr(health==3), predict(pr outcome(3))
4._predict   : Pr(health==4), predict(pr outcome(4))
5._predict   : Pr(health==5), predict(pr outcome(5))

------------------------------------------------------------------------------
               |            Delta-method
               |     Margin   Std. Err.      z    P>|z|     [95% Conf. Interval]
---------------+--------------------------------------------------------------
_predict#black |
   1#nonBlack  |   .0635826   .0023288    27.30   0.000     .0590182     .068147
      1#black  |   .1374343   .0070777    19.42   0.000     .1235622    .1513064
   2#nonBlack  |   .1548809   .0034868    44.42   0.000     .1480469     .161715
      2#black  |   .2475271   .0072026    34.37   0.000     .2334102     .261644
   3#nonBlack  |   .2846398   .0043742    65.07   0.000     .2760666     .293213
      3#black  |   .3054912   .0048495    62.99   0.000     .2959863    .3149961
   4#nonBlack  |   .2540548   .0043035    59.03   0.000     .2456201    .2624894
      4#black  |   .1863406   .0057371    32.48   0.000     .1750961     .197585
   5#nonBlack  |   .2428419   .0041509    58.50   0.000     .2347062    .2509776
      5#black  |   .1232068   .0060262    20.45   0.000     .1113957     .135018
------------------------------------------------------------------------------
```

You get the AAPs for blacks and whites for each category of the ordinal dependent variable. Looking at the results, you can see that, on an all other things equal basis, blacks are more than twice as likely as whites to say they are in poor health (13.7% as opposed to 6.4%), about as likely to report they are in average health (30.5% to 28.5%) and about half as likely to say they are in excellent health (12.3% versus 24.3%).

Long & Freese's `mtable` command (`findit spost13_ado`) produces output that is much easier to read and that does not require Stata 14:

```
. *spost13
. mtable, at(black = (0 1))

Expression: Pr(health), predict(outcome())

          |    black      poor      fair   average      good  excellent
 ----------+-----------------------------------------------------------
        1 |        0     0.064     0.155     0.285     0.254      0.243
        2 |        1     0.137     0.248     0.305     0.186      0.123
```

You could simplify this a bit and say things like, according to the model, on an all other things being equal basis, almost 50% of whites say their health is good or excellent, compared to less than 31% of blacks. Or, if you prefer, you can say that more than 38% of blacks say their health is fair or poor, compared to less than 22% of non-blacks. (Note that the AAPs for each group sum to 1, i.e. everybody has to fall into one of the five health categories.)

Given that there are a lot of numbers here, it might be easier to look at the AMEs:

```
. * AMEs using margins
. margins, dydx(black)

Average marginal effects                          Number of obs     =     10,335
Model VCE    : OIM

dy/dx w.r.t. : 1.black
1._predict   : Pr(health==1), predict(pr outcome(1))
2._predict   : Pr(health==2), predict(pr outcome(2))
3._predict   : Pr(health==3), predict(pr outcome(3))
4._predict   : Pr(health==4), predict(pr outcome(4))
5._predict   : Pr(health==5), predict(pr outcome(5))

------------------------------------------------------------------------------
             |            Delta-method
             |      dy/dx   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
1.black      |
    _predict |
          1  |   .0738517    .006432    11.48   0.000     .0612452    .0864582
          2  |   .0926462   .0061925    14.96   0.000     .0805091    .1047833
          3  |   .0208514   .0018031    11.56   0.000     .0173174    .0243855
          4  |  -.0677142   .0053068   -12.76   0.000    -.0781153   -.0573131
          5  |  -.1196351   .0064726   -18.48   0.000    -.1323211    -.106949
------------------------------------------------------------------------------
Note: dy/dx for factor levels is the discrete change from the base level.

. mtable, dydx(black)

Expression: Marginal effect of Pr(health), predict(outcome())

    poor      fair   average      good  excellent
-------------------------------------------------
   0.074     0.093     0.021    -0.068     -0.120
```

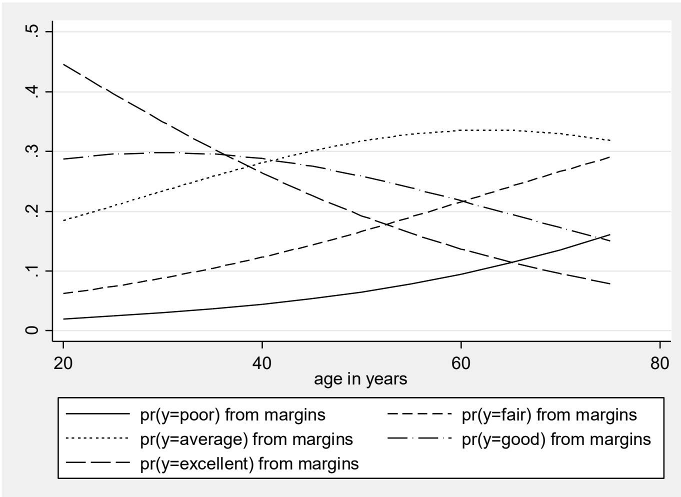
Consistent with the earlier results, the marginal effects show you that, on average, blacks are 7.4 percentage points more likely than whites to say their health is poor, and about 12 percentage points less likely to say their health is excellent. Personally, I find numbers like this much more tangible and meaningful than the raw coefficients.

Here is an example of how you can use Long & Freese's `mtable` command to compute adjusted predictions for prototypical cases. The first command does everything with one command but the subsequent commands produce an easier to read table.

```
. * mtable
. mtable, at (black = (0 1) age = 20 ) at (black = (0 1) age = 47 ) at (black = (0 1) age = 74 ) dec(4)

Expression: Pr(health), predict(outcome())

          |    black       age      poor      fair   average      good  excellent
----------+------------------------------------------------------------------------
        1 |        0        20    0.0175    0.0553    0.1731    0.2869     0.4672
        2 |        1        20    0.0414    0.1184    0.2813    0.2930     0.2659
        3 |        0        47    0.0513    0.1410    0.3046    0.2786     0.2245
        4 |        1        47    0.1157    0.2498    0.3395    0.1882     0.1068
        5 |        0        74    0.1407    0.2781    0.3305    0.1634     0.0872
        6 |        1        74    0.2839    0.3517    0.2430    0.0834     0.0380

. quietly mtable, at (black = 0 age = 20 ) rown(20 year old white) dec(4)
. quietly mtable, at (black = 1 age = 20 ) rown(20 year old black) dec(4) below
. quietly mtable, at (black = 0 age = 47 ) rown(47 year old white) dec(4) below
. quietly mtable, at (black = 1 age = 47 ) rown(47 year old black) dec(4) below
. quietly mtable, at (black = 0 age = 74 ) rown(74 year old white) dec(4) below
. mtable, at (black = 1 age = 74 ) rown(74 year old black) dec(4) below

Expression: Pr(health), predict(outcome())

                   |     poor      fair   average      good  excellent
-------------------+---------------------------------------------------
  20 year old white |   0.0175    0.0553    0.1731    0.2869     0.4672
  20 year old black |   0.0414    0.1184    0.2813    0.2930     0.2659
  47 year old white |   0.0513    0.1410    0.3046    0.2786     0.2245
  47 year old black |   0.1157    0.2498    0.3395    0.1882     0.1068
  74 year old white |   0.1407    0.2781    0.3305    0.1634     0.0872
  74 year old black |   0.2839    0.3517    0.2430    0.0834     0.0380

Specified values of covariates

          |    black       age
----------+-------------------
    Set 1 |        0        20
    Set 2 |        1        20
    Set 3 |        0        47
    Set 4 |        1        47
    Set 5 |        0        74
  Current |        1        74
```

Producing readable graphs can be a little tricky since so many different lines can be produced at the same time because of the multiple outcomes for the dependent variable. The spost13 `mgen` command might be useful for this purpose. I'll just show the commands and then the final graph.

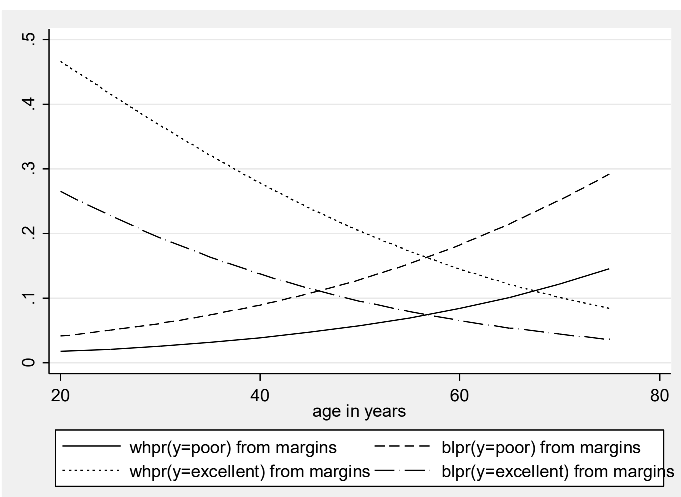
```
* Graphics using mgen
* mgen for all groups pooled together
mgen, at(age = (20(5)75)) stub(all)
list  allpr1 allpr2 allpr3 allpr4 allpr5 allage in 1/15
line allpr1 allpr2 allpr3 allpr4 allpr5 allage, scheme(sj) name(pooled)
```



The line that starts at the top left shows that the probability of being in excellent health is around 45% for twenty year olds and less than 10% by age 70. Conversely, the line that starts at the bottom left shows that hardly anyone is in poor health at age 20 but by age 70 close to 20% are.

If you are doing multiple groups (e.g. blacks and whites) you may want to graph only some of the outcomes. In the following graph we include only the predicted probabilities of poor health and excellent health, separately for blacks and whites.

```
* mgen for groups
drop allpr1 - allCpr5
mgen, at(age = (20(5)75) black  = 0) stub(wh) predn(whpr)
mgen, at(age = (20(5)75) black  = 1) stub(bl) predn(blpr)
line whwhpr1 blblpr1 whwhpr5 blblpr5 whage, scheme(sj) name(byrace)
```



If you look at the two lines that start at the bottom left, you see that the likelihood of being in poor health increases with age and that at every age blacks are more likely to be in poor health than are whites (and the gap increases with age). Conversely the top two lines show you that the probability of being in excellent health decreases with age and that at every age blacks are less likely to have excellent health.

You may also want to look at the user-written `combomarginsplot`, available from SSC. It shows how you can combine categories (e.g. poor and fair, good and excellent) when doing a graph.

Finally the spost13 `mchange` command can be good for producing lots of potentially useful statistics. Read the help for the command if you aren't already familiar with it.

```
. * mchange
. mchange black female age, stats(change start end) dec(5) delta(10)

ologit: Changes in Pr(y) | Number of obs = 10335

Expression: Pr(health), predict(outcome())

                   |      poor       fair    average       good  excellent
-------------------+-------------------------------------------------------
black              |
 black vs nonBlack |   0.07385    0.09265    0.02085   -0.06771   -0.11964
              From |   0.06358    0.15488    0.28464    0.25405    0.24284
                To |   0.13743    0.24753    0.30549    0.18634    0.12321
female             |
    female vs male |   0.00743    0.01179    0.00652   -0.00681   -0.01893
              From |   0.06707    0.15829    0.28411    0.25087    0.23966
                To |   0.07450    0.17008    0.29062    0.24406    0.22074
age                |
                +1 |   0.00266    0.00415    0.00220   -0.00244   -0.00657
              From |   0.07100    0.16450    0.28747    0.24726    0.22977
                To |   0.07365    0.16865    0.28967    0.24482    0.22321
            +delta |   0.03077    0.04251    0.01537   -0.02859   -0.06006
              From |   0.07100    0.16450    0.28747    0.24726    0.22977
                To |   0.10177    0.20701    0.30284    0.21867    0.16972
          Marginal |   0.00261    0.00413    0.00227   -0.00239   -0.00663
              From |        .z         .z         .z         .z         .z
                To |        .z         .z         .z         .z         .z

Average predictions

             |      poor       fair    average       good  excellent
-------------+-------------------------------------------------------
  Pr(y|base) |   0.07100    0.16450    0.28747    0.24726    0.22977
```

Appendix:
Adjusted Predictions & Marginal Effects for Multiple Outcome Models & Commands
(Stata 13 and Earlier)

NOTE: Stata 14 made it much easier to estimate adjusted predictions and marginal effects for multiple outcome commands like `mlogit` and `ologit` and `oglm` and `gologit2`. Therefore use Stata 14, or the spost13 commands, if at all possible. However if you are condemned to using Stata 12 or 13 and do not want to use spost13, this appendix shows the (much more tedious) commands that are needed. Otherwise you can skip this appendix. `ologit` is used in these examples but similar and probably identical procedures can be used with other multiple outcome commands.

As was the case with logit models, the parameters for an ordered logit model can be hard to interpret. Adjusted predictions and marginal effects can again make results more understandable. The `margins` command, however, is harder to use, because we need to issue a separate command for *each* of the outcomes of the ordinal variable. Consider the following. We will once again use the Nhanes2f data, but this time the dependent variable will be self-reported health, whose values range between 1 (poor) and 5 (excellent):

```
. version 13.1
. webuse nhanes2f, clear
. keep if !missing(diabetes, black, female, age)
(2 observations deleted)
. label define black 0 "nonBlack" 1 "black"
. label define female 0 "male" 1 "female"
. label values black black
. label values female female
. fre health

health -- 1=poor,..., 5=excellent
-----------------------------------------------------------------
                    |      Freq.    Percent      Valid       Cum.
--------------------+--------------------------------------------
Valid   1 poor      |        729       7.05       7.05       7.05
        2 fair      |       1670      16.16      16.16      23.21
        3 average   |       2938      28.43      28.43      51.64
        4 good      |       2591      25.07      25.07      76.71
        5 excellent |       2407      23.29      23.29     100.00
        Total       |      10335     100.00     100.00
-----------------------------------------------------------------

. ologit health i.female i.black c.age, nolog

Ordered logistic regression                       Number of obs   =      10335
                                                  LR chi2(3)      =    1682.10
                                                  Prob > chi2     =     0.0000
Log likelihood = -14923.345                       Pseudo R2       =     0.0534

------------------------------------------------------------------------------
      health |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      female |
     female  |  -.1170992   .0355732    -3.29   0.001    -.1868215   -.0473769
             |
       black |
      black  |  -.8845093   .0583106   -15.17   0.000     -.998796   -.7702227
         age |  -.0410673   .0010907   -37.65   0.000     -.043205   -.0389295
-------------+----------------------------------------------------------------
       /cut1 |  -4.910859   .0743281                     -5.056539   -4.765179
       /cut2 |  -3.428162   .0648869                     -3.555338   -3.300986
       /cut3 |  -2.004318   .0586634                     -2.119296    -1.88934
       /cut4 |  -.7512595   .0561222                     -.8612569   -.6412621
------------------------------------------------------------------------------
```

The results tell us that, on an all other things being equal basis, females, blacks, and older people tend to have lower levels of self-reported health. However, other than sign and significance, it is difficult to get a tangible feel for how large and important these differences are.

Using the Margins Command. Now let's see what happens when we use the `margins` command to get the AAPs (Average Adjusted Predictions) for whites and blacks:

```
. margins black

Predictive margins                                Number of obs   =      10335
Model VCE    : OIM

Expression   : Pr(health==1), predict()

------------------------------------------------------------------------------
             |            Delta-method
             |     Margin   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       black |
   nonBlack  |   .0635826   .0023288    27.30   0.000     .0590182     .068147
      black  |   .1374343   .0070777    19.42   0.000     .1235622    .1513064
------------------------------------------------------------------------------
```

We see that blacks score more than 7 percentage points higher than do whites – which may seem odd given that the coefficient for black is negative. However, you have to realize that this is ONLY the predicted probability for health = 1, i.e. the probability that somebody will say their health is poor. So, it is not surprising that blacks are more than twice as likely to report they are in poor health as whites are. This is only one of the five possible outcomes though; if we want to get all of them we need to run five separate margins commands.

```
. * AAPs using margins
. margins black, predict(outcome(#1))

Predictive margins                                Number of obs   =      10335
Model VCE    : OIM

Expression   : Pr(health==1), predict(outcome(#1))

------------------------------------------------------------------------------
             |            Delta-method
             |     Margin   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       black |
   nonBlack  |   .0635826   .0023288    27.30   0.000     .0590182     .068147
      black  |   .1374343   .0070777    19.42   0.000     .1235622    .1513064
------------------------------------------------------------------------------

. margins black, predict(outcome(#2))

Predictive margins                                Number of obs   =      10335
Model VCE    : OIM

Expression   : Pr(health==2), predict(outcome(#2))

------------------------------------------------------------------------------
             |            Delta-method
             |     Margin   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       black |
   nonBlack  |   .1548809   .0034868    44.42   0.000     .1480469     .161715
      black  |   .2475271   .0072026    34.37   0.000     .2334102     .261644
------------------------------------------------------------------------------
```

```
. margins black, predict(outcome(#3))

Predictive margins                                Number of obs   =      10335
Model VCE    : OIM

Expression   : Pr(health==3), predict(outcome(#3))

------------------------------------------------------------------------------
             |            Delta-method
             |     Margin   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       black |
   nonBlack  |   .2846398   .0043742    65.07   0.000     .2760666     .293213
      black  |   .3054912   .0048495    62.99   0.000     .2959863    .3149961
------------------------------------------------------------------------------

. margins black, predict(outcome(#4))

Predictive margins                                Number of obs   =      10335
Model VCE    : OIM

Expression   : Pr(health==4), predict(outcome(#4))

------------------------------------------------------------------------------
             |            Delta-method
             |     Margin   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       black |
   nonBlack  |   .2540548   .0043035    59.03   0.000     .2456201    .2624894
      black  |   .1863406   .0057371    32.48   0.000     .1750961     .197585
------------------------------------------------------------------------------

. margins black, predict(outcome(#5))

Predictive margins                                Number of obs   =      10335
Model VCE    : OIM

Expression   : Pr(health==5), predict(outcome(#5))

------------------------------------------------------------------------------
             |            Delta-method
             |     Margin   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       black |
   nonBlack  |   .2428419   .0041509    58.50   0.000     .2347062    .2509776
      black  |   .1232068   .0060262    20.45   0.000     .1113957     .135018
------------------------------------------------------------------------------
```

Looking at the results, you can see that, on an all other things equal basis, blacks are more than twice as likely as whites to say they are in poor health (13.7% as opposed to 6.4%), about as likely to report they are in average health (30.5% to 28.5%) and about half as likely to say they are in excellent health (12.3% versus 24.3%). You could simplify this a bit and say things like, according to the model, on an all other things being equal basis, almost 50% of whites say their health is good or excellent, compared to less than 31% of blacks. Or, if you prefer, you can say that more than 38% of blacks say their health is fair or poor, compared to less than 22% of non-blacks. (Note that the AAPs for each group sum to 1, i.e. everybody has to fall into one of the five health categories.)

You also get the marginal effects one outcome at a time. For variety, rather than use the dydx option, I will use the r. operator, which gives you some additional useful output. For convenience, I will just show the two most extreme outcomes, poor health (outcome 1) and excellent health (outcome 5).

```
. * AMEs using margins
. margins r.black, predict(outcome(#1))

Contrasts of predictive margins
Model VCE    : OIM

Expression   : Pr(health==1), predict(outcome(#1))

------------------------------------------------
             |         df        chi2     P>chi2
-------------+----------------------------------
       black |          1      131.83     0.0000
------------------------------------------------

----------------------------------------------------------------------
                     |            Delta-method
                     |   Contrast   Std. Err.     [95% Conf. Interval]
---------------------+------------------------------------------------
               black |
(black vs nonBlack)  |   .0738517    .006432      .0612452    .0864582
----------------------------------------------------------------------

. margins r.black, predict(outcome(#5))

Contrasts of predictive margins
Model VCE    : OIM

Expression   : Pr(health==5), predict(outcome(#5))

------------------------------------------------
             |         df        chi2     P>chi2
-------------+----------------------------------
       black |          1      341.63     0.0000
------------------------------------------------

----------------------------------------------------------------------
                     |            Delta-method
                     |   Contrast   Std. Err.     [95% Conf. Interval]
---------------------+------------------------------------------------
               black |
(black vs nonBlack)  |  -.1196351   .0064726     -.1323211    -.106949
----------------------------------------------------------------------
```

Consistent with the earlier results, the marginal effects show you that, on average, blacks are 7.4 percentage points more like than whites to say their health is poor, and about 12 percentage points less likely to say their health is very good.

Personally, I find numbers like this much more tangible and meaningful than the raw coefficients; but the process of generating them is tedious and prone to error. It is probably even worse when you want to make statements about the effects of continuous variables. Fortunately, the m* commands in Spost13 and other user-written routines can make these tasks easier.