## Example of Different Ways to Handle Qualitative (Class) Variables

**Method 1**: Create C – 1 indicator variables, each with 1 for class membership, 0 if not in class. Reference class has all 0s for those variables. Overall intercept ($b_0$) is the intercept for the reference class. Other intercepts are adjusted by relevant coefficient for each class.

```
. regress height momheight male

      Source |       SS       df       MS              Number of obs =     165
-------------+------------------------------           F(  2,   162) =   88.85
       Model |    1330.009     2    665.0045           Prob > F      =  0.0000
    Residual |  1212.48494   162  7.48447494           R-squared     =  0.5231
-------------+------------------------------           Adj R-squared =  0.5172
       Total |  2542.49394   164  15.5030118           Root MSE      =  2.7358

------------------------------------------------------------------------------
      height |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   momheight |   .4768522   .0769811     6.19   0.000     .3248363     .628868
        male |   5.014979   .4277661    11.72   0.000     4.170263    5.859696
       _cons |   34.38884   4.876736     7.05   0.000     24.75868    44.01901
------------------------------------------------------------------------------
```

**Method 2**: Create C – 1 indicator variables, each with 1 for class membership, 0 if not in class, *except*, use –1 for one of the classes on *all* indicator variables. With this method $b_0$ is like an average of all the class intercepts. In this example, "malefemale" = 1 for males, –1 for females.

```
. regress height momheight malefemale

      Source |       SS       df       MS              Number of obs =     165
-------------+------------------------------           F(  2,   162) =   88.85
       Model |    1330.009     2    665.0045           Prob > F      =  0.0000
    Residual |  1212.48494   162  7.48447494           R-squared     =  0.5231
-------------+------------------------------           Adj R-squared =  0.5172
       Total |  2542.49394   164  15.5030118           Root MSE      =  2.7358

------------------------------------------------------------------------------
      height |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   momheight |   .4768522   .0769811     6.19   0.000     .3248363     .628868
  malefemale |    2.50749    .213883    11.72   0.000     2.085132    2.929848
       _cons |   36.89633   4.875632     7.57   0.000     27.26835    46.52432
------------------------------------------------------------------------------
```

**How to put indicator variables into Stata regress command** (using Method 1) when the class variable is entered using text instead of numbers. In this case, sex is 'Male' and 'Female' in the data file.

```
. xi: regress height momheight i.sex
i.sex             _Isex_1-2           (_Isex_1 for sex=='Female' omitted)

      Source |       SS       df       MS              Number of obs =     165
-------------+------------------------------           F(  2,   162) =   88.85
       Model |    1330.009     2    665.0045           Prob > F      =  0.0000
    Residual |  1212.48494   162  7.48447494           R-squared     =  0.5231
-------------+------------------------------           Adj R-squared =  0.5172
       Total |  2542.49394   164  15.5030118           Root MSE      =  2.7358

------------------------------------------------------------------------------
      height |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   momheight |   .4768522   .0769811     6.19   0.000     .3248363     .628868
     _Isex_2 |   5.014979   .4277661    11.72   0.000     4.170263    5.859696
       _cons |   34.38884   4.876736     7.05   0.000     24.75868    44.01901
------------------------------------------------------------------------------
```