Group Comparisons:
Differences in Composition Versus Differences in Models and Effects

Richard Williams, University of Notre Dame, https://www3.nd.edu/~rwilliam/
Last revised February 15, 2015

Overview. This is the first of a series of handouts that will deal with techniques for comparing groups. This initial handout notes that, when comparing groups, it is important to realize that groups can differ in two ways:

- There can be compositional differences between groups. Specifically, means on the Independent Variables (IVs) may differ between groups.
- The effects of IVs can differ between groups. A variable might have a stronger effect on one group than it does on the other. Indeed, the direction of an effect may even differ between groups. The model that describes one group may be very different from the model that describes another.

For example, blacks may have lower levels of education and less job experience than do whites. As a result, they may tend to have lower levels of income, even if the effects of education and job experience are the same for both groups. Simple T-tests or ANOVA tests can determine whether there are significant compositional differences between groups.

Or, blacks may have similar levels of education and job experience, but the effects of these variables may be less for them, e.g. a year of education is worth less to a black than it is to a white. As a result, blacks may tend to have lower incomes than comparable whites.

Compositional, or mean, differences between groups on the IVs may suggest that differences on the DVs are “justified”, e.g. blacks earn less than whites because they are less educated; women earn less than men because they are concentrated in lower-paying occupations, or have less continuous service with the same company. Of course, one must then ask what produced the compositional differences. It may be, for example, that race is a cause of education and job experience; that is, race may be an *indirect* cause of income, because race affects education and job experience which in turn affect income.

Differences in effects raise questions about why those differences exist. If blacks benefit less from education than whites, is this perhaps because of discrimination? Or do other factors need to be considered in the model?

It is important to keep compositional differences and differences in effects separate. Researchers will sometimes confuse the two, muddling the discussion of why group differences exist. In particular, researchers sometimes focus a lot on their models, and overlook how important compositional factors can be in explaining group differences.

Differences in Composition. Returning again to our hypothetical data from 400 whites and 100 blacks: the following t-tests and descriptive statistics reveal that blacks have lower levels of education and job experience than do whites. These differences are all highly significant. These lower levels of education and job experience probably are part of the reason that black income is also lower than white income.

```
. use https://www3.nd.edu/~rwilliam/statafiles/blwh.dta, clear
. ttest   educ, by(black)

Two-sample t test with equal variances

------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
   white |     400        13.9     .175505      3.5101    13.55497    14.24503
   black |     100        10.2    .4376244    4.376244    9.331658    11.06834
---------+--------------------------------------------------------------------
combined |     500       13.16     .178023    3.980715    12.81023    13.50977
---------+--------------------------------------------------------------------
    diff |                 3.7     .413502                2.887576    4.512424
------------------------------------------------------------------------------
Degrees of freedom: 498

                  Ho: mean(white) - mean(black) = diff = 0

     Ha: diff < 0               Ha: diff != 0              Ha: diff > 0
       t =   8.9480               t =   8.9480              t =   8.9480
   P < t =   1.0000         P > |t| =   0.0000          P > t =   0.0000

. ttest   jobexp, by(black)

Two-sample t test with equal variances

------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
   white |     400        14.1    .2395171    4.790341    13.62913    14.57087
   black |     100        11.2    .5464301    5.464301    10.11576    12.28424
---------+--------------------------------------------------------------------
combined |     500       13.52    .2263663    5.061703    13.07525    13.96475
---------+--------------------------------------------------------------------
    diff |                 2.9    .5513765                1.816689    3.983311
------------------------------------------------------------------------------
Degrees of freedom: 498

                  Ho: mean(white) - mean(black) = diff = 0

     Ha: diff < 0               Ha: diff != 0              Ha: diff > 0
       t =   5.2596               t =   5.2596              t =   5.2596
   P < t =   1.0000         P > |t| =   0.0000          P > t =   0.0000

. ttest   income, by(black)

Two-sample t test with equal variances

------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
   white |     400       30.04    .3897187    7.794375    29.27384    30.80616
   black |     100       18.79    .7664749    7.664749    17.26915    20.31085
---------+--------------------------------------------------------------------
combined |     500       27.79    .4013067    8.973491    27.00154    28.57846
---------+--------------------------------------------------------------------
    diff |               11.25    .8685758                9.543475    12.95652
------------------------------------------------------------------------------
Degrees of freedom: 498

                  Ho: mean(white) - mean(black) = diff = 0

     Ha: diff < 0               Ha: diff != 0              Ha: diff > 0
       t =  12.9522               t =  12.9522              t =  12.9522
   P < t =   1.0000         P > |t| =   0.0000          P > t =   0.0000
```

Differences in Effects. The effects of variables may also differ across groups. For example, education may have a greater effect on whites than it does blacks. If whites get greater benefits from their education than blacks do, this will further contribute to racial differences in outcome variables.

There are various ways to estimate differences in effects across groups. Later handouts will show how models with interaction effects provide a powerful and flexible means to detect and test differences across groups. However, it is also possible to simply estimate separate models for each group. Note that I store the model results for each group because I am going to use them in later calculations.

```
. use https://www3.nd.edu/~rwilliam/statafiles/blwh.dta, clear
. * estimate model for blacks
. regress income educ jobexp if black == 1

      Source |       SS       df       MS              Number of obs =     100
-------------+------------------------------           F(  2,    97) =  267.80
       Model |  4924.27286     2  2462.13643           Prob > F      =  0.0000
    Residual |   891.81705    97  9.19399021           R-squared     =  0.8467
-------------+------------------------------           Adj R-squared =  0.8435
       Total |  5816.08991    99   58.748383           Root MSE      =  3.0322

------------------------------------------------------------------------------
      income |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        educ |   1.677949   .0725479    23.13   0.000     1.533962    1.821936
      jobexp |    .421975   .0581021     7.26   0.000     .3066585    .5372915
       _cons |    -3.0512   1.154604    -2.64   0.010    -5.342771   -.7596302
------------------------------------------------------------------------------
. est store black
. * estimate model for whites
. regress income educ jobexp if black == 0

      Source |       SS       df       MS              Number of obs =     400
-------------+------------------------------           F(  2,   397) =  620.07
       Model |  18361.9894     2  9180.99472           Prob > F      =  0.0000
    Residual |  5878.16991   397  14.8064733           R-squared     =  0.7575
-------------+------------------------------           Adj R-squared =  0.7563
       Total |  24240.1594   399  60.7522791           Root MSE      =  3.8479

------------------------------------------------------------------------------
      income |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        educ |   1.893338   .0562591    33.65   0.000     1.782735    2.003941
      jobexp |    .722255   .0412236    17.52   0.000     .6412111    .8032988
       _cons |  -6.461189   1.089219    -5.93   0.000    -8.602547    -4.31983
------------------------------------------------------------------------------
. est store white
```

Just looking at the coefficients, the estimated effect of education is greater for whites than it is for blacks (1.893338 vs 1.677949). Similarly, it is estimated that whites gain more from each year of job experience than blacks do (.722255 vs .421975). Ergo, not only do whites average more years of education and job experience than do blacks, each one of those years benefits them more than it does blacks.

Testing Differences in Effects. Of course, simply eyeballing the results can be deceptive. Apparent differences in coefficients across groups may be due to sampling variability. We will therefore want to do more formal tests. For example, we might want to test whether the effects of education and job experience on income are the same for both whites and blacks, i.e. do blacks get as much benefit from their education and job experience as do whites? That is, we want to test the hypothesis

H_0: $\beta^{(0)} = \beta^{(1)}$ for all corresponding betas across groups
H_A: $\beta^{(0)} <> \beta^{(1)}$ for at least one of the corresponding betas across groups

where the superscripts stand for the group (in this case 0 = white, 1 = black) and the Betas stand for all the coefficients (including the intercepts) estimated for each group.

There are various ways to do this. We can, as before, do incremental F tests. However, when we are estimating separate models for each group, as we are doing here, the procedure can be a bit tedious and prone to error since hand calculations are required. Further, probably few people would do it this way, since it is much easier to do the exact same calculations by specifying models with interaction terms. The procedure is therefore relegated to Appendix A.

A more popular approach uses Wald Tests and the `suest` command (seemingly unrelated estimation). Here you estimate separate models for each group. I personally don't use this approach very often but others often do so it is illustrated in Appendix B.

For now, we will illustrate the use of a *Likelihood Ratio test*. Likelihood ratio tests are widely used with logistic regression and other maximum likelihood techniques, but they can also be used for our current problem. With an LR test, you estimate a constrained and unconstrained model. A chi-square statistic can then be used to test whether the differences between the models are statistically significant.

To do the LR test, we need to estimate one additional model: The model in which both blacks and whites are analyzed simultaneously.

```
. * estimate pooled model
. regress income educ jobexp

      Source |       SS       df       MS              Number of obs =     500
-------------+------------------------------           F(  2,   497) = 1103.96
       Model |  32798.4018     2  16399.2009           Prob > F      =  0.0000
    Residual |  7382.84742   497  14.8548238           R-squared     =  0.8163
-------------+------------------------------           Adj R-squared =  0.8155
       Total |  40181.2493   499  80.5235456           Root MSE      =  3.8542

------------------------------------------------------------------------------
      income |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        educ |    1.94512   .0436998    44.51   0.000     1.859261     2.03098
      jobexp |   .7082212   .0343672    20.61   0.000     .6406983     .775744
       _cons |  -7.382935   .8027781    -9.20   0.000    -8.960192   -5.805678
------------------------------------------------------------------------------
. est store both
```

This last model can be thought of as the *constrained model* (or, if you prefer, the model that corresponds to the null hypothesis specified above). It is constrained in that it does not allow for racial differences in the effects of any variables, or even in the intercepts. That is, for both blacks and whites, the effect of education is 1.945, the effect of job experience is .708, and the intercept is -7.383. (For that matter, the same is true for brown-eyed people and blue-eyed people, and for men and for women; if the model is true then the variables have the same effect on you regardless of what group you happen to be in).

What, then, is the unconstrained model? When we estimated separate models for blacks and whites, we allowed *all* coefficients, including the intercepts, to differ by race; that is, we did not constrain any effects to be equal across the races, whereas in the constrained model we constrained all corresponding effects to be the same across the races.

The *unconstrained model*, then, in this case, is *the combination of the two sets of coefficients from the regressions for each race separately*. Because two separate regressions had to be run to come up with the unconstrained model, the calculation of an incremental F test is awkward but doable (see Appendix A). Luckily, the `lrtest` command is easy to use. You just have to group the components of the unconstrained model in parentheses and then specify the constrained model, using the estimates we previously stored.

```
. lrtest (black white) both, stats

Likelihood-ratio test                                 LR chi2(3)  =     52.33
                                                      Prob > chi2 =    0.0000

Assumption: (both) nested in (black, white)

Akaike's information criterion and Bayesian information criterion

-----------------------------------------------------------------------------
       Model |    Obs    ll(null)   ll(model)     df          AIC         BIC
-------------+---------------------------------------------------------------
        both |    500   -1806.106   -1382.546      3     2771.092    2783.736
       black |    100   -345.0545   -251.2984      3     508.5968    516.4123
       white |    400   -1388.436   -1105.083      3     2216.165     2228.14
-----------------------------------------------------------------------------
               Note:  N=Obs used in calculating BIC; see [R] BIC note
```

How do we interpret the results? For the unconstrained model, we had to estimate 6 parameters: three for blacks and three for whites. For the constrained model, we only had to estimate three parameters. If all of the coefficients were the same across races (other than differences caused by sampling variability) the chi-square statistic would not be significant. That is, the apparent differences we saw in the coefficients for each race would be small enough to attribute to sampling error. However, the LR chi-square value is highly significant, indicating that at least one coefficient differs by race (i.e. either the effect of educ, or the effect of jobexp, or the intercept, or come combination of these three).

As a sidelight, if you divide the LR Chi square statistic by its degrees of freedom, you get a value that is very close to the value of the corresponding F statistic (at least if the sample size is large). As Appendix A shows the corresponding incremental F test yields a value of 14.91; in the LR Chi-square test above, 52.33 / 3 = 17.44. If for some reason you care why then see http://www.stata.com/support/faqs/statistics/chi-squared-and-f-distributions/ .

Note that the above (as well as similar approaches) is sometimes called a *Chow Test*. A Chow Test is a test of whether any parameters differ across populations. It is also common (indeed, probably more common) to have the term Chow Test used when the intercept is allowed to differ across populations but all other parameters are constrained to be the same.

There are two main concerns with the approach of estimating separate models for each group:

- The tests presented in this handout do *not* tell you which coefficients differ across populations. Indeed, they do not even tell you whether it is one of the IV effects that differs across populations. It could just be that the intercept term differs in the two groups.
- You may think that some IVs have different effects in different groups, while other IVs have the same effects in each group. The approaches shown in this handout do not allow you to easily test whether a subset of the variables has different effects across groups. For good theoretical reasons, you may believe that some effects will differ across groups, while others will not.

An alternative approach, which we will describe shortly, makes it possible to overcome these limitations. This approach uses interaction effects rather than estimating separate models for each group.

Conclusion. Both compositional differences and differences in variable effects appear to contribute to income differences between blacks and white. The descriptive statistics reveal that blacks have lower levels of education and job experience. These lower levels of education and experience, combined with the apparently smaller effects of the education and experience they do have, make black income lower than white income.

Remember our earlier caution, however. We do not know which effects significantly differ across populations. Indeed, it may not even be an effect of an IV that is different; it could be that the intercepts significantly differ. Subsequent handouts will discuss alternative and more flexible ways for making comparisons across groups.

More generally, the researcher should be aware that both differences in composition and differences in independent variable effects could be important when trying to explain why differences in outcomes exist across groups. *Further, differences in composition may well reflect the indirect effects of group membership on the outcome variable.* As we said earlier, *a failure to find direct and/or interactive effects of group membership does not mean that group membership is irrelevant for the outcome; it may just be that the effects of group membership are indirect rather than direct.*

In the next handout we will further consider how both differences in composition and differences in effects can combine to produce differences in outcomes.

Appendix A: Incremental F Test

Another procedure that will let us test whether there are <u>any</u> differences in effects across groups is as follows [NOTE: In practice, you would rarely do it this way because, as we will see, simpler and equivalent approaches are available.]:

- Estimate separate regressions for blacks and whites.
  - Add up the error sums of squares from both groups; this is SSE_u.
  - Also note the sample size for each group, i.e. N_1 and N_2.
  - We refer to this as the unconstrained model, because coefficients are free to differ between populations.
  - For group 1, error d.f. = N_1 - K – 1
  - for group 2 error d.f. = N_2 - K – 1
  - hence the error d.f. for both together is $N_1 + N_2$ - 2K - 2.
  - Put another way, a total of 2K + 2 coefficients are estimated: 2 sets of betas, and 2 intercepts, hence the error d.f. in the unconstrained model = $N_1 + N_2$ - 2K - 2.
  - Put another way – we have Group 0 and Group 1. The unconstrained model is obtained by estimating the regressions

$$Y = \alpha^{(0)} + \beta_1^{(0)} X_1 + \beta_2^{(0)} X_2 + \varepsilon \text{ for group 0}$$

$$Y = \alpha^{(1)} + \beta_1^{(1)} X_1 + \beta_2^{(1)} X_2 + \varepsilon \text{ for group 1}$$

where the superscripts stand for the group number.

- Estimate a regression for both groups together. This will give you SSE_c. We refer to this as the constrained model, because parameters (including the intercept) are constrained to be equal in both populations.
- Note that J (the number of restrictions) = K + 1. This is because, not only are all the X's constrained to have equal effects across groups, the intercepts are also constrained to be equal. Also, Total N = $N_1 + N_2$.
- You then compute the incremental F:

$$F_{K+1, N_1+N_2-2K-2} = \frac{(SSE_c - SSE_u) * (N_1 + N_2 - 2K - 2)}{SSE_u * (K+1)}$$

- If the F value is significant, you reject the null, and conclude that coefficients are not the same across groups.
- This strategy can easily be modified for more than 2 groups. Just run separate regressions for each group, add up the SSE's to get the unconstrained SSE. Remember that J = (Number of groups - 1) * (K + 1), unconstrained error d.f. = total sample size - [number of groups*(K + 1)].

EXAMPLE. In our modified Income/Job experience/Education example, there are 100 blacks and 400 whites. First, we estimate separate regressions for blacks and whites. This is easily done in either Stata or SPSS. Using Stata,

```
. use https://www3.nd.edu/~rwilliam/statafiles/blwh.dta, clear
. bysort black: regress income educ jobexp

---------------------------------------------------------------------------------------
-> black = white

      Source |       SS       df       MS              Number of obs =     400
-------------+------------------------------           F(  2,   397) =  620.07
       Model |  18361.9894     2  9180.99472           Prob > F      =  0.0000
    Residual |  5878.16991   397  14.8064733           R-squared     =  0.7575
-------------+------------------------------           Adj R-squared =  0.7563
       Total |  24240.1594   399  60.7522791           Root MSE      =  3.8479

------------------------------------------------------------------------------
      income |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        educ |   1.893338   .0562591    33.65   0.000     1.782735    2.003941
      jobexp |    .722255   .0412236    17.52   0.000     .6412111    .8032988
       _cons |  -6.461189   1.089219    -5.93   0.000    -8.602546   -4.319831
------------------------------------------------------------------------------

---------------------------------------------------------------------------------------
-> black = black

      Source |       SS       df       MS              Number of obs =     100
-------------+------------------------------           F(  2,    97) =  267.80
       Model |  4924.27286     2  2462.13643           Prob > F      =  0.0000
    Residual |   891.81705    97  9.19399021           R-squared     =  0.8467
-------------+------------------------------           Adj R-squared =  0.8435
       Total |  5816.08991    99   58.748383           Root MSE      =  3.0322

------------------------------------------------------------------------------
      income |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        educ |   1.677949   .0725479    23.13   0.000     1.533962    1.821936
      jobexp |    .421975   .0581021     7.26   0.000     .3066585    .5372915
       _cons |    -3.0512   1.154604    -2.64   0.010    -5.342771   -.7596303
------------------------------------------------------------------------------
```

Or alternatively, you could do

```
. regress income educ jobexp if black
. regress income educ jobexp if !black
```

Hence, for whites, $N_w = 400$, $SSE_w = 5878.17$, $DFE_w = 397$. For blacks, $N_b = 100$, $SSE_b = 891.82$, $DFE_b = 97$. Combining the black and white numbers for the unconstrained model,

$$N_u = 500, SSE_u = 6770, DFE_u = 494$$

For the constrained model, Income is regressed on Educ and Jobexp for both groups together:

```
. reg  income educ jobexp

      Source |       SS       df       MS              Number of obs =     500
-------------+------------------------------           F(  2,   497) = 1103.96
       Model |  32798.4018     2  16399.2009           Prob > F      =  0.0000
    Residual |  7382.84742   497  14.8548238           R-squared     =  0.8163
-------------+------------------------------           Adj R-squared =  0.8155
       Total |  40181.2493   499  80.5235456           Root MSE      =  3.8542

------------------------------------------------------------------------------
      income |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        educ |    1.94512   .0436998    44.51   0.000     1.859261     2.03098
      jobexp |   .7082212   .0343672    20.61   0.000     .6406983     .775744
       _cons |  -7.382935   .8027781    -9.20   0.000    -8.960192   -5.805678
------------------------------------------------------------------------------
```

Hence, N_c = 500, SSE_c = 7382.85, DFE_c = 497.

We now compute the incremental F:

$$F_{K+1,N_1+N_2-2K-2} = \frac{(SSE_c - SSE_u)*(N_1 + N_2 - 2K - 2)}{SSE_u*(K+1)} = \frac{(7383-6770)*494}{6770*3} = 14.91$$

Stata commands make it easy to tell what the critical value is for an F with d.f. = 3, 494, and how significant an F value of 14.91 is.

```
. di invF(3,494,.95)
2.6229522
. di Ftail(3,494,14.91)
2.632e-09
```

Using the .05 level of significance, the critical value for an F with d.f. = 3, 494 is only about 2.62, and an F value of 14.91 is highly significant. Therefore, we reject the null hypothesis: coefficients are not the same for both blacks and whites. Just from "eyeballing" the coefficients, it appears that both education and years of job experience have smaller effects on blacks than on whites. We need further tests to identify exactly where the statistically significant differences are.

Appendix B: Using a Wald chi-square test and the suest command

It is also possible to examine whether coefficients differ across groups with a Wald chi-square test. For this you use the Stata `suest` (seemingly unrelated estimation) command. Stata does all the calculations for you. To do this, you estimate separate models for each group, store the results, use `suest` to combine the results into a single model, and then test whether coefficients differ across groups. You also have the option to either include or not include the constants in the equality test. The following example leads to the same conclusion that the other tests did: one or more of the coefficients significantly differ across groups. In addition, the last test command indicates that the differences in coefficients are not just limited to differences in the intercepts.

```
. use https://www3.nd.edu/~rwilliam/statafiles/blwh.dta, clear
. reg income educ jobexp if !black

      Source |       SS       df       MS              Number of obs =     400
-------------+------------------------------           F(  2,   397) =  620.07
       Model |  18361.9894     2  9180.99472           Prob > F      =  0.0000
    Residual |  5878.16991   397  14.8064733           R-squared     =  0.7575
-------------+------------------------------           Adj R-squared =  0.7563
       Total |  24240.1594   399  60.7522791           Root MSE      =  3.8479

------------------------------------------------------------------------------
      income |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        educ |   1.893338   .0562591    33.65   0.000     1.782735    2.003941
      jobexp |    .722255   .0412236    17.52   0.000     .6412111    .8032988
       _cons |  -6.461189   1.089219    -5.93   0.000    -8.602547    -4.31983
------------------------------------------------------------------------------

. est store white
. reg income educ jobexp if black

      Source |       SS       df       MS              Number of obs =     100
-------------+------------------------------           F(  2,    97) =  267.80
       Model |  4924.27286     2  2462.13643           Prob > F      =  0.0000
    Residual |   891.81705    97  9.19399021           R-squared     =  0.8467
-------------+------------------------------           Adj R-squared =  0.8435
       Total |  5816.08991    99   58.748383           Root MSE      =  3.0322

------------------------------------------------------------------------------
      income |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        educ |   1.677949   .0725479    23.13   0.000     1.533962    1.821936
      jobexp |    .421975   .0581021     7.26   0.000     .3066585    .5372915
       _cons |    -3.0512   1.154604    -2.64   0.010    -5.342771   -.7596302
------------------------------------------------------------------------------

. est store black
```

```
. suest black white

Simultaneous results for black, white

                                                  Number of obs   =        500

------------------------------------------------------------------------------
             |               Robust
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
black_mean   |
        educ |   1.677949   .0434112    38.65   0.000     1.592865    1.763034
      jobexp |    .421975   .0328505    12.85   0.000     .3575893    .4863607
       _cons |    -3.0512   .6827896    -4.47   0.000    -4.389443   -1.712957
-------------+----------------------------------------------------------------
black_lnvar  |
       _cons |    2.21855   .1432128    15.49   0.000     1.937858    2.499242
-------------+----------------------------------------------------------------
white_mean   |
        educ |   1.893338   .0454701    41.64   0.000     1.804218    1.982458
      jobexp |    .722255   .0312211    23.13   0.000     .6610628    .7834471
       _cons |  -6.461189   .6423953   -10.06   0.000     -7.72026   -5.202117
-------------+----------------------------------------------------------------
white_lnvar  |
       _cons |   2.695064   .0652246    41.32   0.000     2.567227    2.822902
------------------------------------------------------------------------------

. * Include the constant in the equality tests
. test [black_mean = white_mean], constant coef

 ( 1)  [black_mean]educ - [white_mean]educ = 0
 ( 2)  [black_mean]jobexp - [white_mean]jobexp = 0
 ( 3)  [black_mean]_cons - [white_mean]_cons = 0

           chi2(  3) =  121.57
         Prob > chi2 =    0.0000


Constrained coefficients

------------------------------------------------------------------------------
             |               Robust
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
black_mean   |
        educ |   1.829709   .0309947    59.03   0.000     1.768961    1.890458
      jobexp |   .5614209   .0223398    25.13   0.000     .5176356    .6052062
       _cons |  -4.352208   .4557838    -9.55   0.000    -5.245527   -3.458888
-------------+----------------------------------------------------------------
black_lnvar  |
       _cons |   2.193622   .1274118    17.22   0.000     1.943899    2.443344
-------------+----------------------------------------------------------------
white_mean   |
        educ |   1.829709   .0309947    59.03   0.000     1.768961    1.890458
      jobexp |   .5614209   .0223398    25.13   0.000     .5176356    .6052062
       _cons |  -4.352208   .4557838    -9.55   0.000    -5.245527   -3.458888
-------------+----------------------------------------------------------------
white_lnvar  |
       _cons |   3.001294   .0574212    52.27   0.000      2.88875    3.113837
------------------------------------------------------------------------------
```

```
. * Allow the constants to differ while other coefficients are the same

. test [black_mean = white_mean], coef

 ( 1)  [black_mean]educ - [white_mean]educ = 0
 ( 2)  [black_mean]jobexp - [white_mean]jobexp = 0

           chi2(  2) =   62.05
         Prob > chi2 =    0.0000


Constrained coefficients

------------------------------------------------------------------------------
             |               Robust
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
black_mean   |
        educ |   1.793616   .0313458    57.22   0.000     1.732179    1.855053
      jobexp |   .5865008   .0225751    25.98   0.000     .5422544    .6307473
       _cons |  -6.162922   .5126647   -12.02   0.000    -7.167727   -5.158118
-------------+----------------------------------------------------------------
black_lnvar  |
       _cons |    2.53209   .1347536    18.79   0.000     2.267978    2.796202
-------------+----------------------------------------------------------------
white_mean   |
        educ |   1.793616   .0313458    57.22   0.000     1.732179    1.855053
      jobexp |   .5865008   .0225751    25.98   0.000     .5422544    .6307473
       _cons |   -3.93801    .458935    -8.58   0.000    -4.837506   -3.038514
-------------+----------------------------------------------------------------
white_lnvar  |
       _cons |   2.907211   .0587019    49.52   0.000     2.792157    3.022264
------------------------------------------------------------------------------
```