Now think of the interaction in a different way. Overall, Hanna is more vulnerable than Westar, but the interaction says that the degree of that greater vulnerability depends on the type of fungus. Look at all pairwise comparisons of the DIFFERENCE between Hanna and Westar. First, verify that the interaction can be expressed this way. Of course it can.

F. Plant by MCG followup, Hanna-Westar subset

All pairwise differences of Westar minus Hanna differences

```
proc reg;
     model meanlng = mu7-mu18 / noint;
     F_inter: test   mu13-mu7=mu14-mu8=mu15-mu9
                   = mu16-mu10=mu17-mu11=mu18-mu12;
     F_1vs2:  test   mu13-mu7=mu14-mu8;
     F_1vs3:  test   mu13-mu7=mu15-mu9;
     F_1vs7:  test   mu13-mu7=mu16-mu10;
     F_1vs8:  test   mu13-mu7=mu17-mu11;
     F_1vs9:  test   mu13-mu7=mu18-mu12;
     F_2vs3:  test   mu14-mu8=mu15-mu9;
     F_2vs7:  test   mu14-mu8=mu16-mu10;
     F_2vs8:  test   mu14-mu8=mu17-mu11;
     F_2vs9:  test   mu14-mu8=mu18-mu12;
     F_3vs7:  test   mu15-mu9=mu16-mu10;
     F_3vs8:  test   mu15-mu9=mu17-mu11;
     F_3vs9:  test   mu15-mu9=mu18-mu12;
     F_7vs8:  test   mu16-mu10=mu17-mu11;
     F_7vs9:  test   mu16-mu10=mu18-mu12;
     F_8vs9:  test   mu17-mu11=mu18-mu12;
```

```
Dependent Variable: MEANLNG
Test: F_INTER  Numerator:   5364.0437  DF:    5   F value:   3.8699
               Denominator:  1386.077  DF:   60   Prob>F:    0.0042


Dependent Variable: MEANLNG
Test: F_1VS2   Numerator:  14956.1036  DF:    1   F value:  10.7902
               Denominator:  1386.077  DF:   60   Prob>F:    0.0017


Dependent Variable: MEANLNG
Test: F_1VS3   Numerator:   2349.9777  DF:    1   F value:   1.6954
               Denominator:  1386.077  DF:   60   Prob>F:    0.1979


Dependent Variable: MEANLNG
Test: F_1VS7   Numerator:  15006.4293  DF:    1   F value:  10.8265
               Denominator:  1386.077  DF:   60   Prob>F:    0.0017
```

```
Dependent Variable: MEANLNG
Test: F_1VS8   Numerator:   1147.2776  DF:    1   F value:   0.8277
               Denominator:  1386.077  DF:   60   Prob>F:    0.3666


Dependent Variable: MEANLNG
Test: F_1VS9   Numerator:    630.3018  DF:    1   F value:   0.4547
               Denominator:  1386.077  DF:   60   Prob>F:    0.5027


Dependent Variable: MEANLNG
Test: F_2VS3   Numerator:   5449.1829  DF:    1   F value:   3.9314
               Denominator:  1386.077  DF:   60   Prob>F:    0.0520


Dependent Variable: MEANLNG
Test: F_2VS7   Numerator:      0.0423  DF:    1   F value:   0.0000
               Denominator:  1386.077  DF:   60   Prob>F:    0.9956


Dependent Variable: MEANLNG
Test: F_2VS8   Numerator:   7818.7443  DF:    1   F value:   5.6409
               Denominator:  1386.077  DF:   60   Prob>F:    0.0208


Dependent Variable: MEANLNG
Test: F_2VS9   Numerator:   9445.7674  DF:    1   F value:   6.8147
               Denominator:  1386.077  DF:   60   Prob>F:    0.0114


Dependent Variable: MEANLNG
Test: F_3VS7   Numerator:   5479.5767  DF:    1   F value:   3.9533
               Denominator:  1386.077  DF:   60   Prob>F:    0.0513


Dependent Variable: MEANLNG
Test: F_3VS8   Numerator:    213.3084  DF:    1   F value:   0.1539
               Denominator:  1386.077  DF:   60   Prob>F:    0.6962


Dependent Variable: MEANLNG
Test: F_3VS9   Numerator:    546.1923  DF:    1   F value:   0.3941
               Denominator:  1386.077  DF:   60   Prob>F:    0.5326


Dependent Variable: MEANLNG
Test: F_7VS8   Numerator:   7855.1432  DF:    1   F value:   5.6672
               Denominator:  1386.077  DF:   60   Prob>F:    0.0205
```

```
Dependent Variable: MEANLNG
Test: F_7VS9   Numerator:   9485.7704  DF:    1   F value:   6.8436
               Denominator:  1386.077  DF:   60   Prob>F:    0.0112


Dependent Variable: MEANLNG
Test: F_8VS9   Numerator:     76.8370  DF:    1   F value:   0.0554
               Denominator:  1386.077  DF:   60   Prob>F:    0.8147
```

These analyses are summarized in the table below. Westar-Hanna differences marked with the same letter are not significantly different.

MCG	Westar-Hanna Difference		
7	120.35	A	
2	120.18	A	
3	59.91	A	B
8	47.98		B
9	40.83		B
1	20.33		B

The last two tests investigate whether there are significant differences in response to type of fungus, separately within Hanna and within Westar. We see that they are statistically significant for Westar, and almost reach significance for Hanna.

```
     G_Hanaeq: test   mu7=mu8=mu9=mu10=mu11=mu12;
     H_Westeq: test   mu13=mu14=mu15=mu16=mu17=mu18;
```

```
Dependent Variable: MEANLNG
Test: G_HANAEQ Numerator:   3223.5872  DF:    5   F value:   2.3257
               Denominator:  1386.077  DF:   60   Prob>F:    0.0536


Dependent Variable: MEANLNG
Test: H_WESTEQ Numerator:  17889.2114  DF:    5   F value:  12.9064
               Denominator:  1386.077  DF:   60   Prob>F:    0.0001
```

It makes sense to follow up with pairwise comparisons of the means with Westar, but first let's review what we've done so far, limiting the discussion to just the Hanna-Westar subset of the data. We've tested

- Overall difference among the 12 means
- Main effect for PLANT
- Main effect for MCG
- PLANT*MCG interaction
- 15 pairwise comparisons of the Hanna-Westar difference, following up the interaction
- One comparison of the 6 means for Hanna
- One comparison of the 6 means for Westar

That's 21 tests in all, and we really should do at least 15 more, testing for pairwise differences among the Westar means. Somehow, we should make this into a set of proper post-hoc tests, and correct for the fact that we've done a lot of them. But how?

Tukey tests are only good for pairwise comparisons, and a Bonferroni correction is very ill-advised, since these tests were not all planned before seeing the data. This pretty much leaves us with Scheffé or nothing. The earlier discussion of Scheffé tests was limited to testing single contrasts. Here, some of our involve testing collections of contrasts, so we need a little more generality.

General Scheffé Tests Assume a multifactor design. Create a combination independent variable whose values are all combinations of factor levels. All the tests we do will be tests for collections consisting of one or more contrasts of the cell means.

Start with an *initial* test, an F-test for s contrasts. A Scheffé follow-up test will be a test for d contrasts, *not* necessarily a subset of the contrasts of the initial test. The follow-up test must obey these rules:

- $d < s$
- If all s contrasts of the initial test are zero in the population, then all d contrasts of the follow-up test must be zero in the population. In other words, the null hypothesis of the follow-up test must be implied by the null hypothesis of the initial test.

Next, compute the ordinary one-at-a-time F statistic for the follow-up test (it will have *d* and *n-p* degrees of freedom). Then, use a calculator to compute

$$F_{sch} = \frac{d}{s} F, \qquad (4.2)$$

and if F_{sch} is bigger than the critical value of F for the initial test, the Scheffé follow-up is significant.

Actually, Formula (4.2) is more general. It applies to testing linear combinations of regression coefficients in a multiple regression setting. The initial test is a test of *s* linear constraints on the regression coefficients, and the follow-up test is a test of *d* linear constraints, where $d < s$ and the linear constraints of the initial test imply the linear constraints of the follow-up test. This is very nice because it allows, for example, Scheffé follow-ups to a significant analysis of covariance.

Before applying Scheffé follow-ups to the greenhouse data, a few comments are in order.

- The term "linear constraints" sounds imposing, but a linear constraint is just a statement that some linear combination equals a constant. Almost always, the constant is zero. So for example, saying that a contrast of cell means is equal to zero is the same as specifying a linear constraint on the betas of a multiple regression model (with cell means coding).
- If you're testing 6 independent variables controlling for some other set of independent variables, the null hypothesis says that 6 regression coefficients are equal to zero. That's six linear constraints on the regression coefficients.
- In the initial one-way ANOVA setting where we were testing single contrasts of *p* cell means, the Scheffe F statistic was defined by $F_{sch} = F/(p-1)$. This was a special case of formula (4.2). The initial test for equality of *p* means involved *p-1* contrasts, so $s = p-1$. The followup tests were all for single contrasts, so $d=1$.
- As in the case of testing single contrasts in a one-way design, it is impossible for a followup to be significant if the initial test is not. And if the initial test is significant, there is always something to find in the family of Scheffé follow-ups.

° Suppose we have a follow-up test for *d* linear constraints, and it's not significant. Then *every* follow-up test whose null hypothesis is implied by those constraints will also be non-significant. To use the metaphor of data fishing, once you've looked for fish in a particular region of the lake and determined that there's nothing there, further detailed exploration in that region is a waste of time.

Formula (4.2) is very simple to apply. There are only two potential complications, and they are related to one another.

° First, you have to know what significance test you are following up. For example, if your initial test is the test for equality of *all* cell means, then the test for a given main effect could be carried out as a Scheffé followup, and a pairwise comparison of marginal means would be another followup to the same initial test. Or, you could start with the test for the main effect. Then, the pairwise comparison of marginal means would be a follow-up to the one-at-a-time test for the main effect. You could do it either way, and the conclusions might differ. Where you start is a matter of data-analytic philosophy. But starting with the standard tests for main effects and interactions is more traditional.

° The second potential complication is that you really have to be sure that the null hypothesis of the initial test implies the null hypothesis of the follow-up test. In terms of `proc reg` syntax, it means that the `test` statement of the initial test implies the `test` statements of all the follow-up tests. Sometimes this is easy to check, and sometimes it is tricky. To a large extent, how easy it is to check depends on what the initial test is.

a. If the initial test is a test for all cell means being equal (a one-way ANOVA on the combination variable), then it's easy, because if all the cell means are equal, then any possible contrast of the cell means equals zero. The proof is one line of High School algebra.

b. Similarly, suppose we are using a regression model with an intercept, and the initial test is for all the regression coefficients except β_0 simultaneously. This means that the null hypothesis of the initial test is $\beta_1 = ... = \beta_{p-1} = 0$, and therefore any linear combination of those quantities is zero. This means that you can test any subset of independent variables controlling for all the others as a proper Scheffé follow-up to the first test SAS prints.

c. If you're following up tests for main effects, then the standard test for any contrast of marginal means is a proper follow-up to the test for the main effect.

Beyond these principles, the logical connection between initial and follow-up tests really needs to be checked on a case-by-case basis. Often, the initial test can be expressed more than one way in the `test` statement of `proc reg`, and one of those statements will make things clear enough so you don't need to do any algebra. This is what I did with the significant Plant by Fungus interaction for the Hanna-Westar subset. When the interaction was written as

```
F_inter: test   mu13-mu7=mu14-mu8=mu15-mu9
              = mu16-mu10=mu17-mu11=mu18-mu12;
```

it was clear that all the pairwise comparisons of Westar-Hanna differences were implied.

```
F_1vs2:  test   mu13-mu7=mu14-mu8;
F_1vs3:  test   mu13-mu7=mu15-mu9;
F_1vs7:  test   mu13-mu7=mu16-mu10;
F_1vs8:  test   mu13-mu7=mu17-mu11;
F_1vs9:  test   mu13-mu7=mu18-mu12;
F_2vs3:  test   mu14-mu8=mu15-mu9;
F_2vs7:  test   mu14-mu8=mu16-mu10;
F_2vs8:  test   mu14-mu8=mu17-mu11;
F_2vs9:  test   mu14-mu8=mu18-mu12;
F_3vs7:  test   mu15-mu9=mu16-mu10;
F_3vs8:  test   mu15-mu9=mu17-mu11;
F_3vs9:  test   mu15-mu9=mu18-mu12;
F_7vs8:  test   mu16-mu10=mu17-mu11;
F_7vs9:  test   mu16-mu10=mu18-mu12;
F_8vs9:  test   mu17-mu11=mu18-mu12;
```

Sometimes it is easy to get this wrong. Just note that SAS will do all pairwise comparisons of marginal means (in the `means` statement of `proc glm)` as Scheffé follow-ups, but don't trust it unless the sample sizes are equal. Do it yourself. This warning applies up to SAS version 6.10. Is it a real error, or was it done deliberately to minimize calls to technical support? It's impossible to tell.

Now let's proceed, limiting the analysis to the Hanna-Westar subset. Just for fun, we'll start in two places. Our initial test will be either the test for equality of all 12 cell means, or the test for the Plant by Fungus interaction. Thus, we need two critical values.

```
proc iml; /* Critical values for Scheffe tests */
     interac = finv(.95,5,60) ; print interac;
     oneway = finv(.95,11,60); print oneway;

                                         INTERAC
                                       2.3682702

                                          ONEWAY
                                       1.9522119
```

Initial Test is for Difference Among 12 Cell Means

Let's start by treating the tests for main effects and the interaction as follow-ups to the significant ANOVA on the combination variable (F = 12.43; df=11,71; p < .0001). The table below is based on numbers displayed earlier.

Effect	One-at-a-time F	$F_{sch} = \frac{d}{s} F$	d	Significant with Scheffé?
PLANT	60.52	5.50	1	Yes
MCG	11.36	5.16	5	Yes
PLANT*MCG	3.87	1.75	5	No
All Hanna Equal?	2.33	1.06	5	No
All Westar Equal?	12.91	5.87	5	Yes

The main effect for Plant is still significant; it means that Westar is more vulnerable than Hanna. The main effect for Fungus (MCG) is significant, but as mentioned eralier, it should not be interpreted.

The interesting Plant by MCG interaction is no longer significant as a Scheffe test. This means that all the pairwise comparisons among Westar-Hanna differences will also be non-significant, as Scheffe follow-ups to the oneway ANOVA on the combination variable. There are no fish in that part of the lake. Just to check, the biggest Westar-Hanna difference was 120.35 for MCG 7, and the smallest was 20.33 for MCG 1. Comparing these two differences yielded a one-at-a-time F of 10.83. But d=1 here and s=11, so that F_{sch}=.98. This falls short of the 1.95 required for significance, and as expected, none of the proper follow-ups to a non-significant follow-up are significant.

Pairwise comparisons of the Westar means are of interest, and the easiest way to get them is to ask `proc glm` for all pairwise comparisons of cell means.

```
proc glm data=hanstar;
     class combo;
     model meanlng = combo;
     means combo / scheffe;

                   Scheffe's test for variable: MEANLNG

          NOTE: This test controls the type I experimentwise error rate but
                generally has a higher type II error rate than REGWF for all
                pairwise comparisons

                      Alpha= 0.05  df= 60  MSE= 1386.077
                         Critical Value of F= 1.95221
                    Minimum Significant Difference= 99.608

          Means with the same letter are not significantly different.

               Scheffe Grouping              Mean      N  COMBO

                             A             187.48      6  14
                             A
                             A             173.97      6  16
                             A
                      B      A             154.10      6  15
                      B      A
                      B      A   C          95.82      6  17
                      B      A   C
                      B      A   C          94.19      6  9
                      B          C
                      B          C          67.30      6  8
                      B          C
                      B          C          66.50      6  18
                      B          C
                      B          C          65.91      6  13
                                 C
                                 C          53.62      6  10
                                 C
                                 C          47.84      6  11
                                 C
                                 C          45.58      6  7
                                 C
                                 C          25.67      6  12
```

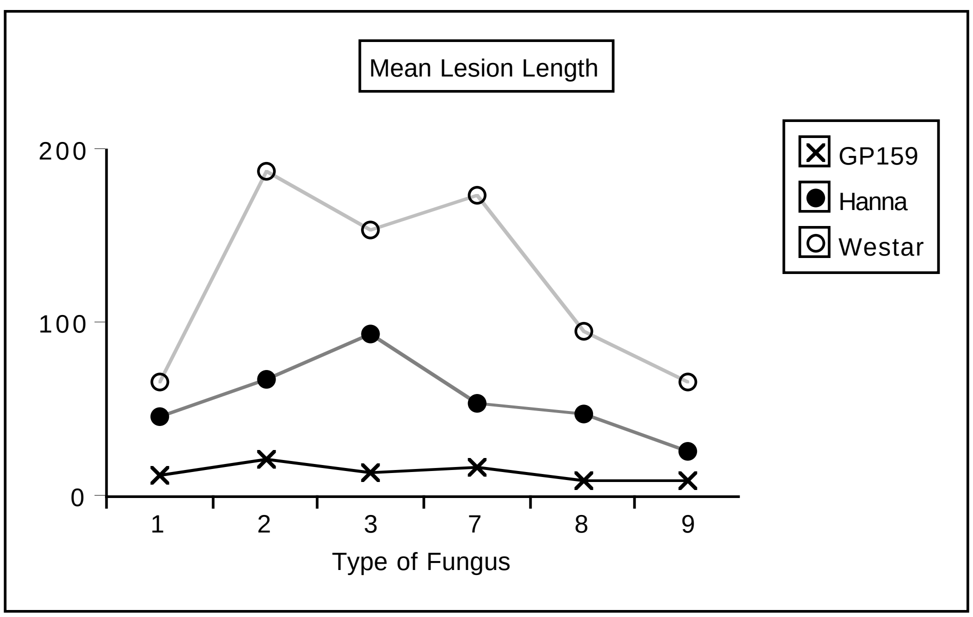
On Westar, Fungus types 2, 3 and 7 grow significantly faster than types 1 and 9, while type 8 is not significantly different from either group. As expected, there are no significant differences among Fungus types for Hanna.

Starting with the Interaction

Logically, a test for interaction can be a follow-up test, but almost no one ever does this in practice. It's much more traditional to start with a one-at a time test for interaction and then, if you're very sophisticated, do Scheffe follow-ups to that initial test. Now s = 5 and the critical value is `2.3682702`.

Again, the biggest Westar-Hanna difference was 120.35 for MCG 7, and the smallest was 20.33 for MCG 1. Comparing these two differences yielded a one-at-a-time F of 10.83. This yields $F_{sch} = \frac{d}{s} F = \frac{1}{5} * 10.83 = 2.16$. But this falls short of the critical value of 2.37, so none of the pairwise comparisons of Westar-Hanna differences reaches significance as a Scheffe follow-up -- even though they look very promising.

As a mathematical certainty, there *is* a single-contrast Scheffe follow-up to the interaction that is significant, but I am still looking for it. The next place I will look is: pairwise comparisons of the differences of line-segment slopes from the interaction plot.



Interactions as Products of Independent Variables

Categorical by Quantitative

An interaction between a quantitative variable and a categorical variable means that differences in E[Y] between categories depend on the value of the quantitative variable, or (equivalently) that the slope of the lines relating x to E[Y] are different, depending on category membership. Such an interaction is represented by **products** of the quantitative variable and the dummy variables for the categorical variable.

For example, consider the metric cars data (`mcars.dat`). It has length, weight, origin and fuel efficiency in kilometeres per litre, for a sample of cars. The three origins are US, Japanese and Other. Presumably these refer tothe location of the head office, not to where the car was manufactured.

Let's use indicator dummy variable coding for origin, with an intercept. In an Analysis of Covariance (ANCOVA), we'd test country of origin controlling, say, for weight. Letting x represent weight and c1 and c2 the dummy variables for country of origin, the model would be

$$E[Y] = \beta_0 + \beta_1 x + \beta_2 c_1 + \beta_3 c_2.$$

This model assumes no interaction between country and weight. The following model includes product terms for the interaction, and would allow you to test it.

$$E[Y] = \beta_0 + \beta_1 x + \beta_2 c_1 + \beta_3 c_2 + \beta_4 c_1 x + \beta_5 c_2 x$$

Country	c1	c2	Expected KPL (let x = weight)
U. S.	1	0	$(\beta_0 + \beta_2) + (\beta_1+\beta_4)x$
Japan	0	0	$\beta_0 + \beta_1 x$
European	0	1	$(\beta_0 + \beta_3) + (\beta_1+\beta_5)x$

It's clear that the slopes are parallel if and only if $\beta_4=\beta_5=0$, and that in this case the relationship of fuel efficiency to country would not depend on weight of the car.

As the program below shows, interaction terms are created by literally multiplying independent variables, and using products as additional independent variables in the regression equation.

```
/********************* mcars.sas *************************/
options linesize=79 pagesize=100 noovp formdlim='-';
title 'Metric Cars Data: Dummy Vars and Interactions';

proc format; /* Used to label values of the categorical variables */
     value carfmt    1 = 'US'
                     2 = 'Japanese'
                     3 = 'European' ;
data auto;
     infile 'mcars.dat';
     input id country kpl weight length;
/* Indicator dummy vars: Ref category is Japanese */
     if country = 1 then c1=1;  else c1=0;
     if country = 3 then c2=1;  else c2=0;
     /* Interaction Terms */
     cw1 = c1*weight; cw2 = c2*weight;
     label country = 'Country of Origin'
           kpl = 'Kilometers per Litre';
     format country carfmt.;

proc means;
     class country;
     var weight kpl;

proc glm;
     title 'One-way ANOVA';
     class country;
     model kpl = country;
     means country / tukey;

proc reg;
     title 'ANCOVA';
     model kpl = weight c1 c2;
     country: test c1 = c2 = 0;

proc reg;
     title 'Test parallel slopes (Interaction)';
     model kpl = weight c1 c2 cw1 cw2;
     interac: test cw1 = cw2 = 0;
     useuro:  test cw1=cw2;
     country: test c1 = c2 = 0;
     eqreg:   test c1=c2=cw1=cw2=0;

proc iml; /* Critical value for Scheffe tests */
     critval = finv(.95,4,94) ; print critval;
```

```
/* Could do most of it with proc glm: ANCOVA, then test interaction */

proc glm;
     class country;
     model kpl = weight country;
     lsmeans country;

proc glm;
     class country;
     model kpl = weight country weight*country;
```

Let's take a look at the output. First, proc means indicates that the US cars get lower gas mileage, and that weight is a potential confounding variable.

```
 COUNTRY  N Obs  Variable  Label                    N         Mean
-----------------------------------------------------------------
US           73  WEIGHT                            73      1540.23
                 KPL       Kilometers per Litre   73    8.1583562

Japanese     13  WEIGHT                            13      1060.27
                 KPL       Kilometers per Litre   13    9.8215385

European     14  WEIGHT                            14      1080.32
                 KPL       Kilometers per Litre   14   11.1600000
-----------------------------------------------------------------

 COUNTRY  N Obs  Variable  Label                       Std Dev       Minimum
------------------------------------------------------------------------------
US           73  WEIGHT                            327.7785402   949.5000000
                 KPL       Kilometers per Litre     1.9760813     5.0400000

Japanese     13  WEIGHT                            104.8370989   891.0000000
                 KPL       Kilometers per Litre     2.3976719     7.5600000

European     14  WEIGHT                            240.9106607   823.5000000
                 KPL       Kilometers per Litre     4.2440764     5.8800000
------------------------------------------------------------------------------

 COUNTRY  N Obs  Variable  Label                         Maximum
--------------------------------------------------------------
US           73  WEIGHT                              2178.00
                 KPL       Kilometers per Litre    12.6000000

Japanese     13  WEIGHT                              1237.50
                 KPL       Kilometers per Litre    14.7000000

European     14  WEIGHT                              1539.00
                 KPL       Kilometers per Litre    17.2200000
--------------------------------------------------------------
```

The one-way ANOVA indicates that fuel efficiency is significantly related to country of origin; country explains 17% of the variation in fuel efficiency.

```
                           General Linear Models Procedure

   Dependent Variable: KPL   Kilometers per Litre
                                           Sum of            Mean
   Source                   DF            Squares          Square  F Value     Pr > F

   Model                     2       121.59232403     60.79616201    10.09     0.0001
   Error                    97       584.29697197      6.02368012
   Corrected Total          99       705.88929600

                     R-Square               C.V.        Root MSE            KPL Mean
                     0.172254           27.90648       2.4543187           8.7948000
```

The Tukey follow-ups are not shown, but they indicate that only the US-European difference is significant. Maybe the US cars are less efficient because they are big and heavy. So let's do the same test, controlling for weight of car. Here's the SAS code. Note this is a standard Analysis of Covariance, and we're *assuming* no interaction.

```
proc reg;
     title 'ANCOVA';
     model kpl = weight c1 c2;
     country: test c1 = c2 = 0;
```

```
   Dependent Variable: KPL        Kilometers per Litre

                                Analysis of Variance

                                     Sum of         Mean
          Source           DF       Squares       Square       F Value       Prob>F

          Model             3     436.21151    145.40384        51.761       0.0001
          Error            96     269.67779      2.80914
          C Total          99     705.88930

              Root MSE       1.67605     R-square       0.6180
              Dep Mean       8.79480     Adj R-sq       0.6060
              C.V.          19.05728
```

```
                         Parameter Estimates

                   Parameter     Standard    T for H0:
Variable  DF        Estimate        Error   Parameter=0    Prob > |T|

INTERCEP   1       16.226336    0.76312281       21.263        0.0001
WEIGHT     1       -0.006041    0.00057080      -10.583        0.0001
C1         1        1.236147    0.57412989        2.153        0.0338
C2         1        1.459591    0.64565633        2.261        0.0260

-------------------------------------------------------------------------------

Dependent Variable: KPL
Test: COUNTRY  Numerator:      8.6168  DF:    2   F value:   3.0674
               Denominator:  2.809144  DF:   96   Prob>F:    0.0511
```

First notice that by including weight, we're now explaining 61% of the variation, while before we explined just 17%. Also, while the effect for country was comfortably significant before we controlled for weight, now it narrowly fails to reach the traditional criterion (p = `0.0511`). But to really appreciate these results, we need to make a table.

Country	c1	c2	$E[Y] = \beta_0 + \beta_1 x + \beta_2 c_1 + \beta_3 c_2$
U. S.	1	0	$(\beta_0 + \beta_2) + \beta_1 x$
Japan	0	0	$\beta_0 \qquad + \beta_1 x$
European	0	1	$(\beta_0 + \beta_3) + \beta_1 x$


```
                         Parameter Estimates

                   Parameter     Standard    T for H0:
Variable  DF        Estimate        Error   Parameter=0    Prob > |T|

INTERCEP   1       16.226336    0.76312281       21.263        0.0001
WEIGHT     1       -0.006041    0.00057080      -10.583        0.0001
C1         1        1.236147    0.57412989        2.153        0.0338
C2         1        1.459591    0.64565633        2.261        0.0260
```


Observe that both b_2 and b_3 are positive -- and significant. Before we controlled for weight, Japanese gas mileage was a little better than US, though not significantly so. Now, because b_2 estimates β_2, and β_2 is the population difference between U.S. and Japanese mileage (for any fixed weight), a positive value of b_2 means that once you control for weight, the U.S. cars are getting better gas mileage than the Japanese -- significantly better, too, if you believe the t-test and not the F-test.

The *direction* of the results has changed because we controlled for weight. This can happen.

Also, may seem strange that the tests for β_2 and β_3 are each significant individually, but the simultaneous test for both of them is not. But this the simultaneous test implicitly includes a comparison between U.S. and European cars, and theyare very close, once you control for weight.

The best way to summarize these results would be to calculate Y-hat for each country of origin, with weight set equal to its mean value in the sample. Instead of doing that, though, let's first test the interaction, which this analysis is *assuming* to be absent.

```
proc reg;
     title 'Test parallel slopes (Interaction)';
     model kpl = weight c1 c2 cw1 cw2;
        interac: test cw1 = cw2 = 0;
        useuro:  test cw1=cw2;
        country: test c1 = c2 = 0;
        eqreg:   test c1=c2=cw1=cw2=0;
```

```
   Dependent Variable: KPL        Kilometers per Litre

                                  Sum of         Mean
          Source          DF      Squares        Square       F Value       Prob>F

          Model            5    489.27223     97.85445       42.463       0.0001
          Error           94    216.61706      2.30444
          C Total         99    705.88930

              Root MSE       1.51804     R-square       0.6931
              Dep Mean       8.79480     Adj R-sq       0.6768
              C.V.          17.26062

                                 Parameter Estimates

                         Parameter      Standard    T for H0:
        Variable  DF      Estimate         Error   Parameter=0    Prob > |T|

        INTERCEP   1     29.194817    4.45188417         6.558        0.0001
        WEIGHT     1     -0.018272    0.00418000        -4.371        0.0001
        C1         1    -12.973668    4.53404398        -2.861        0.0052
        C2         1     -4.891978    4.85268101        -1.008        0.3160
        CW1        1      0.013037    0.00421549         3.093        0.0026
        CW2        1      0.006106    0.00453064         1.348        0.1810
```

```
-------------------------------------------------------------------------------

Dependent Variable: KPL
Test: INTERAC  Numerator:     26.5304  DF:    2   F value:  11.5127
               Denominator:  2.304437  DF:   94   Prob>F:    0.0001

Dependent Variable: KPL
Test: USEURO   Numerator:     33.0228  DF:    1   F value:  14.3301
               Denominator:  2.304437  DF:   94   Prob>F:    0.0003

Dependent Variable: KPL
Test: COUNTRY  Numerator:     24.4819  DF:    2   F value:  10.6238
               Denominator:  2.304437  DF:   94   Prob>F:    0.0001

Dependent Variable: KPL
Test: EQREG    Numerator:     17.5736  DF:    4   F value:   7.6260
               Denominator:  2.304437  DF:   94   Prob>F:    0.0001
```

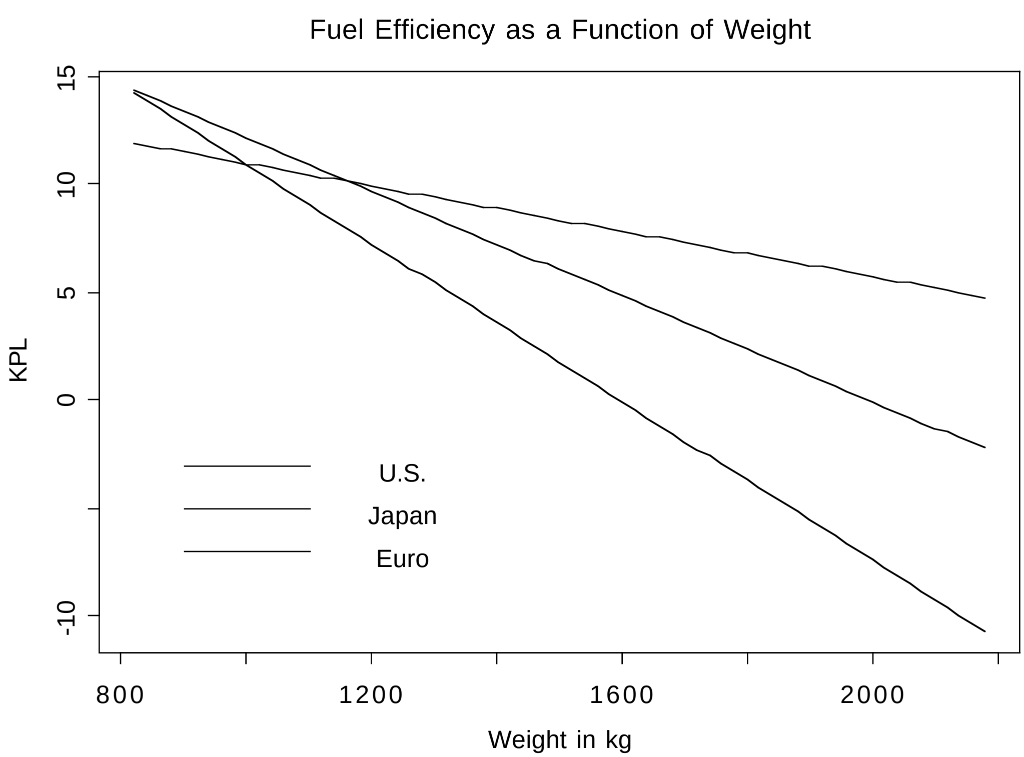
Now the coefficients for the dummy variables are both negative, and the coefficients for the interaction terms are positive. To see what's going on, we need a table *and* a picture -- of $\hat{Y}$.

$\hat{Y} = b_0 + b_1x + b_2c_1 + b_3c_2 + b_4c_1x + b_5c_2x$

$= 29.194817 - 0.018272x - 12.973668c_1 - 4.891978c_2 + 0.013037c_1x + 0.006106c_2x$

Country	c1	c2	Predicted KPL (let x = weight)
U. S.	1	0	$(b_0 + b_2) + (b_1+b_4)x$ = 16.22 - 0.005235 x
Japan	0	0	$b_0 + b_1 x$ = 29.19 - 0.018272 x
European	0	1	$(b_0 + b_3) + (b_1+b_5)x$ = 24.30 - 0.012166 x

From the proc means output, we find that the lightest car was 823.5kg, while the heaviest was 2178kg. So we will let the graph range from 820 to 2180.



When there were no interaction terms, b2 and b3 represented a main effect for country. What do they represent now?

From the picture, it is clear that the most interesting thing is that the slope of the line relating weight to fuel efficiency is least steep for the U.S. Is it significant? 0.05/3 = 0.0167.

Repeating earlier material, ...

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
INTERCEP	1	29.194817	4.45188417	6.558	0.0001
WEIGHT	1	-0.018272	0.00418000	-4.371	0.0001
C1	1	-12.973668	4.53404398	-2.861	0.0052
C2	1	-4.891978	4.85268101	-1.008	0.3160
CW1	1	0.013037	0.00421549	3.093	0.0026
CW2	1	0.006106	0.00453064	1.348	0.1810

```
useuro:  test cw1=cw2;
```

```
Dependent Variable: KPL
Test: USEURO   Numerator:      33.0228  DF:    1   F value:  14.3301
               Denominator:   2.304437  DF:   94   Prob>F:    0.0003
```

The conclusion is that with a Bonferroni correction, the slope is less (less steep) for US than for either Japanese or European, but Japanese and European are not significantly different from each other.

Another interesting follow-up would be to use Scheffe tests to compare the heights of the regression lines at many values of weight; infinitely many comparisons would be protected simultaneously.