1972

Another Look at Tests of Equality between Sets of Coefficients in Two Linear Regressions

Ronald L. Oaxaca

Research Report 7204

ANOTHER LOOK AT TESTS OF EQUALITY

BETWEEN SETS OF COEFFICIENTS IN

TWO LINEAR REGRESSIONS

by

Ronald L. Oaxaca

June, 1972

Another Look at Tests of Equality

Between Sets of Coefficients in

Two Linear Regressions*

by

RONALD L. OAXACA

1. Introduction

Undoubtedly, anyone with at least an introductory course in econome-
trics has been exposed to the "Chow" test. As popularized by Johnston
[4, pp. 136-138], the "Chow" test is a statistical means for deciding whether
or not a linear regression run on one group of observations is significantly
different from the same linear regression run on another group.

For some purposes it suffices to test only for the overall structural
difference; however, there are many instances in which the researcher would
like to know which regressors give rise to the structural difference. In
certain other cases a researcher may have reason to include a subset of
regressors for one group of observations that are not applicable to the
other group. By virtue of the difference in specification, the overall
structure is different between the two groups of observations. Yet we may
be interested in structural differences with respect to the subset of regres-
sors that the two linear regressions have in common.

The answers to these additional questions can be found in Chow's
article on the subject [1], but their significance from an operational
point of view is not immediately clear from the exposition. Recognizing the
difficulty that some may have in following Chow's development of the testing
procedures, F. M. Fisher has sought to simplify the development and illustrate
the unity of Chow's tests with the F test for the significance of subsets of

---

coefficients in a single linear regression [2].

While all of this is familiar terrain to econometricians, it is not at all obvious to students nor to noneconometricians how one actually goes about implementing these tests. In so far as I know, there is no published source that clearly illustrates the operational manner in which one tests for differences in coefficients in the variety of situations previously discussed. It is therefore the purpose of this paper to present a straightforward, step by step, procedure for carrying out these tests. The emphasis is on ease of application and not on the actual derivation of the F tests themselves. We shall be examining only those situations in which there are sufficient degrees of freedom to allow estimation of separate regressions for the two groups of observations.

2. Tests for Structural Differences When the Two Regressions are Identically Specified

We begin by postulating the following linear relationships:

$$(2.1) \qquad y_{1i} = \sum_{j=1}^{k} \beta_{1j} X_{1ij} + e_{1i} , \qquad i = 1,\ldots,n_1$$

$$y_{2i} = \sum_{j=1}^{k} \beta_{2j} X_{2ij} + e_{2i} , \qquad i = 1,\ldots,n_2$$

We assume that e_1 and e_2 have the same normal distribution. With the exception of the coefficients, we can drop the subscripts that denote which sample the observation is drawn from. The two linear relationships of (2.1) can then be combined into a single linear model:

$$(2.2) \qquad y_i = \sum_{j=1}^{k} \left[\beta_{1j} X_{ij} D_i + \beta_{2j} X_{ij} (1 - D_i) \right] + e_i , \qquad i = 1,\ldots,n_1 + n_2$$

where D_i = 1 if the observation is from group one, and

$\qquad$ = 0 otherwise.

Thus D_i and $(1 - D_i)$ are merely dummy variables that are interacted with all k regressors. After collecting terms on $X_{ij} D_i$ in (2.2), we obtain

$$y_i = \sum_{j=1}^{k} \left[(\beta_{1j} - \beta_{2j}) X_{ij} D_i + \beta_{2j} X_{ij} \right] + e_i$$

or

$$(2.3) \qquad y_i = \sum_{j=1}^{k} \left[\Delta \beta_j X_{ij} D_i + \beta_{2j} X_{ij} \right] + e_i$$

where $\Delta \beta_j \equiv \beta_{1j} - \beta_{2j}$.

With 2k regressors and $n_1 + n_2$ observations, equation (2.3) is estimated directly by ordinary least squares. The standard error and 't' value corresponding to $\Delta \hat{\beta}_j$ tell us how significant the difference is between the coefficients if they had been estimated separately by ordinary least squares, i.e., $\hat{\beta}_{1j} - \hat{\beta}_{2j}$ for $j = 1, \ldots, k$.

Often one is interested in a joint test of significance for a subset of the $\Delta \hat{\beta}_j$'s. For example, there may be some interest in testing for a structural difference with respect to a variable that is specified as a set of regressors. Suppose that the variable of interest consists of $k - h$ regressors, then the null hypothesis is that $\Delta \beta_{h+1} = \ldots = \Delta \beta_k = 0$. In other words the null hypothesis states that y_i in (2.3) does not depend on $X_{ij} D_i$, $j = h+1, \ldots, k$. Following the standard procedure, we re-estimate equation (2.3) omitting those regressors whose coefficients are to be jointly tested for significance. The equation to be estimated is given by

$$(2.4) \qquad y_i = \sum_{j=1}^{h} \Delta \beta_j' X_{ij} D_i + \sum_{j=1}^{k} \beta_{2j}' X_{ij} + e_i', \qquad 0 \leq h < k.$$

Basically, the joint test of significance involves a comparison between the explained sums of squares for equations (2.3) and (2.4).

Let $\quad$ SST $\quad = \quad \sum\limits_{i=1}^{n_1+n_2} (y_i - \bar{y})^2$ $\qquad$ (Total sum of squares)

$\quad$ SSR $\quad = \quad \sum\limits_{i=1}^{n_1+n_2} (\hat{y}_i - \bar{\hat{y}})^2$ $\qquad$ (Explained sum of squares from the full regression (2.3)) and,

$\quad$ SSE $\quad = \quad \sum\limits_{i=1}^{n_1+n_2} \hat{e}_i^2$ $\qquad$ (Sum of squared residuals from (2.3)) .

SSR' $\quad$ and $\quad$ SSE' $\quad$ are the corresponding statistics from (2.4), $\quad$ and

Δ SSR $\quad = \quad$ SSR - SSR'

On the hypothesis that $\Delta\beta_{h+1} = \ldots = \Delta\beta_k = 0$, the ratio

(2.5) $\qquad \dfrac{\Delta\,SSR/(k-h)}{SSE/(n_1+n_2-2k)}$ $\quad$ is distributed as $\quad F^{2k-h}_{n_1+n_2-2k}$

High values of (2.5) lead to rejection of the null hypothesis. For a detailed development of this testing procedure see Goldberger [3, pp. 173-177].

Some regression programs do not provide the explained sum of squares statistic in the computer print out; however, the F statistic can be computed from the R^2 and the sum of squared residuals. Given that

$\qquad$ SST $\quad = \quad$ SSR + SSE $\qquad$ and

$\qquad R^2 \quad = \quad \dfrac{SSR}{SST} \quad = \quad 1 - \dfrac{SSE}{SST}$ $\qquad$ it follows that

$\qquad \dfrac{R^2}{1-R^2}$ $\quad$ (SSE) $\quad = \quad$ SSR . $\qquad$ Thus

$$\frac{\Delta\,SSR/(k-h)}{SSE/(n_1+n_2-2k)} = \frac{\left[\dfrac{R^2}{1-R^2}\,(SSE) - \dfrac{R^{2'}}{1-R^{2'}}\,(SSE') \right] \Big/ (k-h)}{SSE/(n_1+n_2-2k)}.$$

In terms of the approach taken here, the "Chow" test described in Johnston [4, pp. 136-138; 5, pp. 192-207] is equivalent to and nothing more than a joint test of all the $\Delta\beta_j$'s , i.e., $h = 0$. Under our procedures equation (2.3) is re-estimated omitting $X_{ij} D_i$, $j = 1,\ldots,k$ and therefore becomes

(2.6)
$$y_i = \sum_{j=1}^{k} \beta_j X_{ij} + \mu_i , \qquad i = 1,\ldots,n_1 + n_2 .$$

On the hypothesis that $\Delta\beta_i = \ldots = \Delta\beta_k = 0$, the ratio

(2.7)
$$\frac{(SSR - SSR'')/k}{SSE/(n_1 + n_2 - 2k)}$$
is distributed as $F^k_{n_1+n_2-2k}$

> (where SSR" is the explained sum of squares corresponding to equation (2.6)).

In the appendix to this paper, a short proof is given that (2.7) is identical to the ratio that would be computed by following Johnston's steps for testing for an overall structural difference between two linear relations.

3. Tests for Structural Differences When the Two Regressions have Different Specifications

Testing for structural differences can be generalized to those cases in which not all of the regressors in the linear relations are the same between the two groups of observations. The general functional relationships may differ between the two groups as follows:

$$y_1 = f_1(X,W) + e_1$$

(3.1)

$$y_2 = f_2(X,Z) + e_2$$

where X is a vector of regressors common to both groups, and W and Z are vectors of regressors specific to groups one and two respectively.

The linear regression form of (3.1) can be expressed as

$$y_{1i} = \sum_{j=1}^{k} \beta_{1j} X_{1ij} + \sum_{j=1}^{k_1} \alpha_{1j} W_{1ij} + e_{1i} , \qquad i = 1,\ldots,n_1$$

(3.2)

$$y_{2i} = \sum_{j=1}^{k} \beta_{2j} X_{2ij} + \sum_{j=1}^{k_2} \gamma_{2j} Z_{2ij} + e_{2i} , \qquad i = 1,\ldots,n_2$$

Next, the two separate linear relations of (3.2) are combined into a single expression:

(3.3) $$y_i = \sum_{j=1}^{k} \left[\beta_{1j} X_{ij} D_i + \beta_{2j} X_{ij}(1-D_i) \right] + \sum_{j=1}^{k_1} \alpha_{1j} W_{ij} D_i$$

$$+ \sum_{j=1}^{k_2} \gamma_{2j} Z_{ij}(1-D_i) + e_i , \qquad i = 1,\ldots,n_1 + n_2$$

By rearranging the terms inside the first summation of (3.3), we obtain

(3.4) $$y_i = \sum_{j=1}^{k} \left[\Delta\beta_j X_{ij} D_i + \beta_{2j} X_{ij} \right] + \sum_{j=1}^{k_1} \alpha_{1j} W_{ij} D_i + \sum_{j=1}^{k_2} \gamma_{2j} Z_{ij}(1-D_i) + e_i$$

With $2k + k_1 + k_2$ regressors and $n_1 + n_2$ observations, equation (3.4) is estimated directly by ordinary least squares. The standard error and 't' value corresponding to $\Delta\hat{\beta}_j$ tell us how significant the difference is between $\hat{\beta}_{1j}$ and $\hat{\beta}_{2j}$ had we estimated the two equations of (3.2) separately by ordinary least squares.

The joint test of significance for a subset of the $\Delta\hat{\beta}_j$'s in (3.4) parallels the joint test described in section (2): Equation (3.4) is re-estimated omitting those regressors whose coefficients are to be jointly tested for significance. Of course we are only interested in a joint test from among the coefficients of the regressors that the two groups have in common. Suppose the $k-h$ regressors to be omitted from (3.4) are $X_{ij} D_i$, $j = h+1,\ldots,k$, then the null hypothesis is that $\Delta\beta_{h+1} = \ldots = \Delta\beta_k = 0$.

Accordingly the F ratio for the joint test is

$$(3.5) \qquad \frac{\Delta \, SSR/(k-h)}{SSE/(n_1 + n_2 - 2k - k_1 - k_2)}$$

which is distributed as $F^{k-h}_{n_1 + n_2 - 2k - k_1 - k_2}$.

Although the linear relations for the two groups are specified differently, we find that we can still test for equality between the coefficients of the regressors common to the two linear relationships. Depending on the interests of the researcher, these tests can be made individually or jointly.

4. Example

An application of the testing procedures presented in sections (2) and (3) is illustrated by an example taken from a cross-section study of male-female wage differentials [6]. Separate wage regressions were estimated for the white males and white females in a household sample drawn from the 1967 Survey of Economic Opportunity. The natural logarithm of the hourly wage was the dependent variable. The independent variables are very briefly described here:

Experience, age less years of schooling completed less six years;

Education, years of schooling completed;

Health Problems, dummy variable = 1 if the individual reports health problems that affect the kind or amount of work he can perform, and '0' otherwise;

Part-Time, dummy variable = 1 if the individual works less than thirty-five hours a week, and '0' otherwise;

Migration, dummy variable = 1 if the individual has maintained a residence more than fifty miles from his current address since the age of seventeen, and '0' otherwise;

YRSM, number of years since the individual last migrated;

Spouse Present, Spouse Absent, Widowed, Divorced, dummy variables for

marital status with Never Married as the reference group;

SMSA 250, SMSA 500, SMSA 750, SMSA 750+, dummy variables for size of urban

area where the Standard Metropolitan Statistical Area is less than 250,000,

or greater than or equal to 250,000 but less than 500,000, or greater than

or equal to 500,000 but less than 750,000, or greater than or equal to 750,000,

with Urban, Non SMSA as the reference group;

North East, North Central, West, dummy variables for the census region with

South as the reference group; and

Children, number of children the female worker has given birth to. The

Experience, Education, and YRSM variables also include quadratic terms.

The separate wage equations are given by

$$\ln(w_{fi}) = \sum_{j=1}^{21} \beta_{fj} X_{fij} + \alpha_f W_{fi} + e_{fi} , \qquad i = 1,\ldots,4{,}962$$

(4.1)

$$\ln(w_{mi}) = \sum_{j=1}^{21} \beta_{mj} X_{mij} + e_{mi} , \qquad i = 1,\ldots,8{,}123$$

where w_{fi} represents the children variable and the subscripts

'm' and 'f' denote males and females respectively.

Except for the children variable occurring only in the female wage equation,

the two equations are identically specified. The estimated differences

between the coefficients of the regressors common to both equations are

directly obtained from the combined regression equation

$$(4.2) \qquad \ln(w_i) = \sum_{j=1}^{21} \left[\Delta \beta_j X_{ij} D_i + \beta_{mj} X_{ij} \right] + \alpha_f W_i D_i + e_i ,$$

$$i = 1,\ldots,13{,}085$$

where $\Delta\beta_j \equiv \beta_{fj} - \beta_{mj}$ and $D_i = 1$ if the worker is female, '0' otherwise.

The results for both the separate and combined regressions are presented in Table (4.1). Since the estimated parameters of the separate regression equations of (4.1) are all contained in the combined regression (4.2), it may seem redundant to repeat them; however, they are repeated in Table (4.1) for illustrative purposes only. Although the variances of the true residuals corresponding to the two groups may be the same, in practice the standard error of estimate is seldom identical for separate regressions run on two groups of observations. Consequently, the 't' values and implied standard errors from the combined regression in Table (4.1) differ somewhat from the values obtained through the separate regressions. This of course presents no problem when one considers that the standard errors from the combined regression are based on more information than in the separate regressions.

Aside from the children variable, are the wage equations for males and females significantly different? The answer to this question requires a joint test of significance for differences in the coefficients of the twenty-one regressors common to both groups. Consequently, equation (4.2) is re-estimated omitting $X_{ij} D_i$, $j = 1,\ldots,21$. The resultant equation is

$$(4.3) \qquad \ln(w_i) = \sum_{j=1}^{21} \beta_j X_{ij} + \alpha'_f W_i D_i + e'_i , \qquad i = 1,\ldots,13,085$$

The pertinent information for the F test is provided in Table (4.2).

TABLE 4.1

SEPARATE AND COMBINED WAGE REGRESSIONS[a]

Dependent Variable: log (hourly wage)	Females		Males	Combined		
Independent Variables:	$\hat{\beta}_f$	$\hat{\gamma}_f$	$\hat{\beta}_m$	$\Delta\hat{\beta}$	$\hat{\gamma}_f$	$\hat{\beta}_m$
Constant	.0894 (.94)		-.0681 (-2.03)	.1575 (1.87)		-.0681 (-1.46)
Experience						
Experience	.0182 (10.20)		.0222 (16.51)	-.0039 (-1.81)		.0222 (15.78)
Experience**2	-.000349 (-9.69)		-.000354 (-14.19)	-.000005 (.13)		-.000354 (-13.52)
Education						
Education	-.0394 (-3.30)		.0342 (5.24)	-.0736 (-5.63)		.0342 (4.99)
Education**2	.00450 (8.63)		.00097 (3.51)	.0353 (6.24)		.00097 (3.34)
Health Problems	-.1097 (-3.89)		-.1325 (-7.57)	.0228 (.71)		-.1325 (7.21)
Part-Time	-.1560 (-9.09)		-.3154 (-14.84)	.1594 (5.81)		-.3154 (-14.14)
Migration						
Migration	-.1262 (-5.54)		-.0316 (-2.08)	-.0946 (-3.56)		-.0316 (-1.98)
YRSM	.0107 (3.81)		.0073 (3.98)	.0034 (1.05)		.0073 (3.80)
YRSM**2	-.000169 (-2.29)		-.000140 (-2.91)	-.000029 (-.34)		-.000140 (-2.77)
Marital Status						
Spouse Present	.1246 (5.96)		.2514 (15.44)	-.1268 (-4.90)		.2514 (14.72)

TABLE 4.1 (continued)

	Females		Males	Combined		
	$\hat{\beta}_f$	$\hat{\gamma}_f$	$\hat{\beta}_m$	$\Delta\hat{\beta}$	$\hat{\gamma}_f$	$\hat{\beta}_m$
Marital Status						
Spouse Absent	.0706 (1.07)		.1189 (1.71)	-.0483 (-.51)		.1189 (1.63)
Widowed	.0804 (2.41)		.1389 (3.00)	-.0585 (-1.01)		.1389 (2.86)
Divorced	.1064 (3.61)		.1027 (3.34)	.0037 (.09)		.1027 (3.19)
Never Married	---		---	---		---
Size of Urban Area						
Urban, Non SMSA	---		---	---		---
SMSA 250	.1080 (4.23)		.0412 (2.30)	.0668 (2.20)		.0412 (2.20)
SMSA 500	.1154 (4.11)		.0845 (4.26)	.0309 (.92)		.0845 (4.06)
SMSA 750	.1721 (5.77)		.1739 (8.47)	-.0018 (-.05)		.1739 (8.07)
SMSA 750+	.2543 (12.27)		.1972 (13.45)	.0571 (2.31)		.1972 (12.81)
Region						
North East	.1129 (5.69)		.0655 (4.73)	.0474 (2.01)		.0655 (4.51)
North Central	.0685 (3.52)		.0790 (5.91)	-.0105 (-.46)		.0790 (5.63)
South	---		---	---		---
West	.1174 (5.43)		.1111 (7.49)	.0063 (.25)		.1111 (7.14)
Children		-0.0295 (-6.31)			-0.295 (-6.77)	
R^2	.22		.34	.39		
S.E.E.	.4851		.4308	.4521		
NOBS	4962		8123	13,085		

TABLE 4.2

Regression	SSR	SSE
(4.2)	1696.7331	2665.9758
(4.3)	1503.3711	

Using formula (3.5) with $h = k_2 = 0$, the F ratio is

$$\frac{\Delta SSR/21}{SSE/13,042} = \frac{193.3620/21}{2665.9758/13,042}$$

$$= \frac{9.2077}{.2044}$$

$$= 45.05$$

In order for the differences in the coefficients to be jointly significant at the 1% level, the F ratio must exceed 1.88; therefore, we can easily reject the null hypothesis that the wage equations are the same for males and females with respect to the regressors common to both.

5. Summary

Beginning with identical linear specifications for two groups of observations, we briefly list the steps for testing for equality between sets of coefficients in the two linear relations:

1. Pool the two groups of observations $(n_1 + n_2)$ and generate an additional set of regressors by interacting the original set (X_{ij}) with a dummy variable that distinguishes between the two groups (D_i).

2. Estimate the combined regression with 2k regressors by ordinary least squares.

3. The estimated coefficients $(\Delta \hat{\beta}_j)$ of the interaction terms $(X_{ij} D_i)$

are the estimated differences in the coefficients that could have been obtained by separately estimating the regressions for the two groups and then subtracting the estimated coefficients of one group from the estimated coefficients of the other, i.e., $\Delta\hat{\beta}_j \equiv \hat{\beta}_{1j} - \hat{\beta}_{2j}$. The corresponding 't' values indicate the statistical significance of these differences.

4. The estimated coefficients $(\hat{\beta}_{2j})$ of the noninteracted terms (X_{ij}) are identical to the estimated coefficients that would be obtained from a separate regression run for one of the two groups.

5. To perform a joint test for differences between a subset of coefficients $\Delta\hat{\beta}_j$, $j = h+1,\ldots,k$, re-estimate the combined regression omitting the $k - h$ regressors $X_{ih+1} D_i, \ldots, X_{ik} D_i$. The change in the explained sum of squares (ΔSSR) is the basis for computing the F statistic

$$\frac{\Delta SSR/(k - h)}{SSE/(n_1 + n_2 - 2k)} \sim F_{n_1+n_2-2k}^{k - h} .$$

The standard "Chow" test corresponds to the case when $h = 0$ (or $k - h = k$).

Only slight changes to the above are required when the linear relations for the two groups are specified such that they have a subset of regressors in common and differ with respect to other regressors:

1. The combined regression will now include the regressors specific to each group. These are interacted with the dummy variable that represents the relevant group, i.e., D_i or $(1 - D_i)$.

2. The F statistic for joint tests becomes

$$\frac{\Delta SSR/(k - h)}{SSE/(n_1 + n_2 - 2k - k_1 - k_2)}$$

Hopefully, this paper has provided a simplified procedure for testing for equality between sets of coefficients in two linear regressions. Although the example in section (4) was taken from a cross section study, our testing procedures obviously apply to time series: One may be interested in testing whether the same linear relation applies to two different time periods. For another example of a time series application see [7].

Appendix

In this section we will prove that ordinary least squares estimation of the combined regression for two groups of observations is equivalent to ordinary least squares estimation of the separate regressions. We will also show that our test statistic for the overall structural difference between the coefficients in two linear relations is identical to the test statistic developed in Johnston [4, p. 137; 5, p. 199].

The development of the first proof is greatly facilitated by casting the model in matrix form:

$$y_1 = X_1 \beta_1 + e_1$$

(A.1)

$$y_2 = X_2 \beta_2 + e_2$$

where y_1 and y_2 are $n_1 \times 1$ and $n_2 \times 1$ vectors respectively;

$\qquad X_1$ and X_2 are $n_1 \times k$ and $n_2 \times k$ matrices respectively;

$\qquad \beta_1$ and β_2 are $k \times 1$ vectors; and

$\qquad e_1$ and e_2 are $n_1 \times 1$ and $n_2 \times 1$ vectors respectively.

It is assumed that e_1 and e_2 have the same normal distribution. Through the appropriate use of dummy variables that identify which group the observation is drawn from, the two linear relationships of (A.1) can be combined into a single linear model:

(A.2)
$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}$$

or equivalently,

$$(A.3) \quad \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} X_1 \\ 0 \end{bmatrix} \beta_1 + \begin{bmatrix} 0 \\ X_2 \end{bmatrix} \beta_2 + \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}$$

Define $\Delta\beta \equiv \beta_1 - \beta_2$, then $\beta_1 \equiv \Delta\beta + \beta_2$.

Substituting for β_1 in (A.3) we obtain

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} X_1 \\ 0 \end{bmatrix} \Delta\beta + \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \beta_2 + \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}$$

or

$$(A.4) \quad \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} X_1 & X_1 \\ 0 & X_2 \end{bmatrix} \begin{bmatrix} \Delta\beta \\ \beta_2 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}$$

$$n_1 + n_2 \times 1 \qquad n_1 + n_2 \times 2k \quad 2k \times 1 \qquad n_1 + n_2 \times 1$$

The linear model (A.4) is of course the matrix form of equation (2.3).

We wish to show that the ordinary least squares estimator of $\Delta\beta$ in (A.4) is identical to the difference between the ordinary least squares estimators of β_1 and β_2 in the separate linear relationships of (A.1). We also wish to prove that the ordinary least squares estimators of β_2 in the combined regression of (A.4) and in the separate regression of (A.1) are identical. From ordinary least squares estimation of (A.4) we have

$$\begin{bmatrix} \hat{\Delta\beta} \\ \hat{\beta}_2 \end{bmatrix} = \left[\begin{pmatrix} X_1' & 0' \\ X_1' & X_2' \end{pmatrix} \begin{pmatrix} X_1 & X_1 \\ 0 & X_2 \end{pmatrix} \right]^{-1} \begin{bmatrix} X_1' & 0' \\ X_1' & X_2' \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

$$= \begin{bmatrix} X_1' X_1 & X_1' X_1 \\ X_1' X_1 & X_1' X_1 + X_2' X_2 \end{bmatrix}^{-1} \begin{bmatrix} X_1' y_1 \\ X_1' y_1 + X_2' y_2 \end{bmatrix}$$

Through the application of partitioned matrix inversion we obtain

$$\begin{bmatrix} \Delta\hat{\beta} \\ \hat{\beta}_2 \end{bmatrix} = \begin{bmatrix} (X_1' X_1)^{-1} + (X_2' X_2)^{-1} & -(X_2' X_2)^{-1} \\ -(X_2' X_2)^{-1} & (X_2' X_2)^{-1} \end{bmatrix} \begin{bmatrix} X_1' y_1 \\ X_1' y_1 + X_2' y_2 \end{bmatrix}$$

$$= \begin{bmatrix} (X_1' X_1)^{-1} X_1' y_1 - (X_2' X_2)^{-1} X_2' y_2 \\ (X_2' X_2)^{-1} X_2' y_2 \end{bmatrix}$$

and thus

$$\Delta\hat{\beta} = (X_1' X_1)^{-1} X_1 y_1 - (X_2' X_2)^{-1} X_2' y_2 = \hat{\beta}_1 - \hat{\beta}_2$$

$$\hat{\beta}_2 = (X_2' X_2)^{-1} X_2' y_2 .$$

Using our notation, the F statistic presented in Johnston [4, p. 137] for testing for overall differences between two linear regressions is

$$(A.5) \quad \frac{[SSE'' - (SSE_1 + SSE_2)]/k}{(SSE_1 + SSE_2)/(n_1 + n_2 - 2k)}$$

where SSE'' is the sum of squared residuals from the pooled regression run without the interaction terms, i.e.,

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \beta + \mu \quad \text{as equation (2.6)} ,$$

and SSE_1, SSE_2 are the sums of squared residuals from the separate regressions for groups one and two respectively.

In his second edition, Johnston develops the testing procedures in terms of the analysis of covariance [5, pp. 192-207]. As such, the analysis handles more than two groups of observations, but the sample size is assumed to be the same for all groups. The analysis is easily adapted to the present case of two groups with their sample sizes not necessarily the same. Also in the

second edition, $SSE_1 + SSE_2$ in the denominator of (A.5) was replaced by the identical sum of squared residuals from a combined regression of the form given in (A.2) [5, p. 199]. These are of course identical to the sum of squared residuals (SSE) from our combined regression (A.4). Replacing $SSE_1 + SSE_2$ in (A.5) by SSE, we have

$$\frac{(SSE'' - SSE)/k}{SSE/(n_1 + n_2 - 2k)} \cdot$$

Now $SSE'' = SST - SSR''$ and $SSE = SST - SSR$ so that $SSE'' - SSE = SSR - SSR''$. Therefore (A.5) is identical to

$$\frac{(SSR - SSR'')/k}{SSE/(n_1 + n_2 - 2k)}$$ which is of course formula (2.7).

When the linear relationships for the two groups have different specifications but with a subset of regressors in common, the model in matrix form can be expressed as

$$y_1 = X_1 \beta_1 + W_1 \alpha_1 + e_1$$

(A.6)

$$y_2 = X_2 \beta_2 + Z_2 \gamma_2 + e_2$$

where y_1, y_2, X_1, X_2, β_1, β_2, e_1, and e_2 have the same dimensions as in (A.1);

W_1 and Z_2 are $n_1 \times k_1$ and $n_2 \times k_2$ matrices respectively; and α_1 and γ_2 are $k_1 \times 1$ and $k_2 \times 1$ vectors respectively.

Combining the two linear relations of (A.6) into a single linear model, we have

$$\text{(A.7)} \qquad \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} X_1 & W_1 & 0 & 0 \\ 0 & 0 & X_2 & Z_2 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \alpha_1 \\ \beta_2 \\ \gamma_2 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}$$

or

$$(A.8) \qquad \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} X_1 \\ 0 \end{bmatrix} \beta_1 + \begin{bmatrix} W_1 \\ 0 \end{bmatrix} \alpha_1 + \begin{bmatrix} 0 \\ X_2 \end{bmatrix} \beta_2 + \begin{bmatrix} 0 \\ Z_2 \end{bmatrix} \gamma_2 + \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}.$$

Substituting $\Delta\beta + \beta_2$ for β_1 in (A.8), we obtain

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} X_1 \\ 0 \end{bmatrix} \Delta\beta + \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \beta_2 + \begin{bmatrix} W_1 \\ 0 \end{bmatrix} \alpha_1 + \begin{bmatrix} 0 \\ Z_2 \end{bmatrix} \gamma_2 + \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}$$

or

$$(A.9) \qquad \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} X_1 & X_1 & W_1 & 0 \\ 0 & X_2 & 0 & Z_2 \end{bmatrix} \begin{bmatrix} \Delta\beta \\ \beta_2 \\ \alpha_1 \\ \gamma_2 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}$$

$$n_1 + n_2 \times 1 \qquad n_1 + n_2 \times 2k + k_1 + k_2 \qquad 2k + k_1 + k_2 \times 1 \qquad n_1 + n_2 \times 1$$

It is a straightforward exercise to show that ordinary least squares estimation of $\Delta\beta_1$ β_2, α_1 and γ_2 in the combined regression of (A.9) is identical to ordinary least squares estimation of the separate linear relationships in (A.6); therefore, the proof is omitted.

References

[1] Chow, Gregory C., "Tests of Equality Between Subsets of Coefficients in
      Two Linear Regressions," <u>Econometrica</u>, Vol. 28, No. 3 (July, 1960),
      pp. 591-605.

[2] Fisher, Franklin M., "Tests of Equality Between Sets of Coefficients in
      Two Linear Regressions: an Expository Note," <u>Econometrica</u>, Vol.
      38, No. 2 (March, 1970), pp. 361-366.

[3] Goldberger, Arthur S., <u>Econometric Theory</u>, New York, John Wiley and
      Sons, 1964.

[4] Johnston, J., <u>Econometric Methods</u>, New York, McGraw-Hill Book Company,
      Inc., 1963.

[5] Johnston, J., <u>Econometric Methods</u>, New York, McGraw-Hill Book Company,
      Inc., second edition, 1972.

[6] Oaxaca, Ronald L., "Male-Female Wage Differentials in Urban Labor
      Markets," unpublished Ph.D. dissertation, Princeton University,
      1971.

[7] Valentine, Thomas J., "The Chow Test and Autocorrelation: A Comment,"
      mimeographed paper, Department of Economics, University of
      Western Ontario, London, Canada, 1971.