On the Power of the *F*-test for Linear Restrictions in a Regression Model

Willim E. Griffiths

University of Melbourne

R. Carter Hill

Louisiana State University

8 May 2020

Abstract

Consider testing a joint null hypothesis on regression coefficients where some of the restrictions in the null hypothesis are true and some are false. We show that an F-test that includes the valid restrictions frequently has greater power than the test that excludes the valid restrictions, a result that may be counterintuitive.

Keywords: noncentrality parameter; joint hypotheses; surplus restrictions

JEL codes: C12, C20

Corresponding author
William Griffiths
Department of Economics
University of Melbourne
Vic 3010
Australia
wegrif@unimelb.edu.au

1. Introduction

Consider the linear regression model $E(y \mid X) = X\beta$ with covariance matrix $\sigma^2 I$. We assume *X* is of dimension $(N \times K)$ with rank K, that $N > K$, and that *y* is normally distributed, conditional on *X*. We are concerned with the power of the *F*-test for a set of linear restrictions $H_0 : R\beta = r$ which can be written as

$$\begin{pmatrix} R_1 \\ R_2 \end{pmatrix} \beta = \begin{pmatrix} r_1 \\ r_2 \end{pmatrix}$$

where $R_1\beta_1 \neq r_1$ and $R_2\beta_2 = r_2$. That is, the first set of restrictions is false and the second set is true. We assume that R_1 is of dimension $(J_1 \times K)$, R_2 is $(J_2 \times K)$, and $J_1 + J_2 < K$.

We ask the question: which test has the greater power, the *F*-test for $H_0 : R\beta = r$, or the *F*-test for $H_0 : R_1\beta = r_1$? Since the latter test omits the superfluous valid restrictions $R_2\beta_2 = r_2$, intuition suggests it should have the greater power. In fact, Ruud (2000, p.235) "proves" that the noncentrality parameters for the two tests are equal, and argues that the power for $H_0 : R_1\beta = r_1$ will be greater because of the reduction in the numerator degrees of freedom. There is an error in Ruud's proof, however. We show that, except under special circumstances, the noncentrality parameter for testing $H_0 : R_1\beta = r_1$ *is less* than the non-centrality parameter for testing $H_0 : R\beta = r$, and hence, whether testing $H_0 : R_1\beta = r_1$ is more powerful will depend on the competing effects of changes in the noncentrality parameter, the degrees of freedom in the numerator, and the critical value of the test.

We then go on to ask: what happens if the model is re-estimated subject to the restrictions $R_2\beta_2 = r_2$ which are known to be true, and the *F*-test for $H_0 : R_1\beta = r_1$ is applied to this restricted model? We show that the noncentrality parameter for this *F*-test is identical to that for testing $H_0 : R\beta = r$ in the unrestricted model. This result brings the earlier result more in line with our intuition. While it might seem counterintuitive for the noncentrality parameter to be larger when valid restrictions are included in the null hypothesis $H_0 : R\beta = r$, our intuition does suggest that it should be larger when testing $H_0 : R_1\beta = r_1$ in a correctly restricted model. We conclude by exploring the conditions under which the noncentrality parameter for testing $H_0: R_1\beta = r_1$ from the restricted model is equal to the noncentrality parameter for testing $H_0: R_1\beta = r_1$ from the unrestricted model.

To summarise, we are interested in comparing the noncentrality parameters for the *F*-test under the following three scenarios.

1. Testing $H_0: R\beta = r$ when $R_1\beta_1 \neq r_1$ and $R_2\beta_2 = r_2$.
2. Testing $H_0: R_1\beta = r_1$ when $R_1\beta_1 \neq r_1$ and $R_2\beta_2 = r_2$.
3. Testing $H_0: R_1\beta = r_1$ when $R_1\beta_1 \neq r_1$ and the valid restrictions $R_2\beta_2 = r_2$ have been imposed on the model.

Let the three noncentrality parameters for these cases be λ_1, λ_2 and λ_3, respectively. We show that $\lambda_1 = \lambda_3 \geq \lambda_2$, and we specify the conditions under which $\lambda_1 = \lambda_3 = \lambda_2$.

The *F*-statistic for the complete set of restrictions is $H_0: R\beta = r$ is

$$F = \frac{\left(R\hat{\beta} - r\right)'\left[\hat{\sigma}^2 R\left(X'X\right)^{-1} R'\right]^{-1}\left(R\hat{\beta} - r\right)}{J_1 + J_2}$$

where $\hat{\beta} = \left(X'X\right)^{-1} X'y$ is the least squares estimator and $\hat{\sigma}^2 = \left(y - X\hat{\beta}\right)'\left(y - X\hat{\beta}\right) \big/ (N - K)$ is the variance estimator. The noncentrality parameter is

$$\lambda = \left(R\beta - r\right)'\left[\sigma^2 R\left(X'X\right)^{-1} R'\right]^{-1}\left(R\beta - r\right)$$

In Section 2 we illustrate our results with some simple examples. Section 3 is devoted to proofs of the general results. Section 4 contains some numerical examples of the power differences. A concluding remark is provided in Section 5.

2. Simple Examples

Consider a model with two explanatory variables x_1 and x_2 which are expressed in terms of deviations from their means such that $\sum_{i=1}^{N} x_{1i} = 0$, $\sum_{i=1}^{N} x_{2i} = 0$ and $E\left(y_i \mid X\right) = x_{1i}\beta_1 + x_{2i}\beta_2$.

Example 1

For the first example we compare noncentrality parameters for null hypotheses $H_0: \beta_1 = \beta_2 = 0$ and $H_0: \beta_1 = 0$ when $\beta_1 \neq 0$ and $\beta_2 = 0$. For testing $H_0: \beta_1 = \beta_2 = 0$, we have

$$R = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad r = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \qquad \left[\sigma^2 R\left(X'X\right)^{-1} R'\right]^{-1} = \frac{X'X}{\sigma^2}$$

and

$$\lambda_1 = \frac{1}{\sigma^2}\left(\beta_1 \quad 0\right) X'X \begin{pmatrix} \beta_1 \\ 0 \end{pmatrix} = \frac{\beta_1^2 \sum_{i=1}^N x_{1i}^2}{\sigma^2}$$

For testing $H_0 : \beta_1 = 0,$ we have $R_1 = \left(1 \quad 0\right), r_1 = 0$, and

$$\lambda_2 = \frac{1}{\sigma^2}\beta_1^2\left[\left(1 \quad 0\right)\left(X'X\right)^{-1}\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right]^{-1}$$

$$= \frac{1}{\sigma^2}\beta_1^2\left(\frac{\sum_{i=1}^N x_{2i}^2}{\sum_{i=1}^N x_{1i}^2 \sum_{i=1}^N x_{2i}^2 - \left(\sum_{i=1}^N x_{1i}x_{2i}\right)^2}\right)^{-1}$$

$$= \frac{\beta_1^2 \sum_{i=1}^N x_{1i}^2}{\sigma^2}\left(1 - r^2\right)$$

where $r^2 = \left(\sum_{i=1}^N x_{i1}x_{21}\right)^2 \Big/ \left(\sum_{i=1}^N x_{1i}^2 \sum_{i=1}^N x_{2i}^2\right)$ is the squared sample correlation between x_1 and x_2. Thus, we have $\lambda_1 \geq \lambda_2$, with equality holding when x_1 and x_2 are uncorrelated, a result that contradicts Ruud's general result that $\lambda_1 = \lambda_2$.

For the third scenario, we impose the restriction $\beta_2 = 0$, leading to the model $E\left(y_i \mid X\right) = x_{1i}\beta_1$. Then, for testing $H_0 : \beta_1 = 0$, we see immediately that

$$\lambda_3 = \frac{\beta_1^2 \sum_{i=1}^N x_{1i}^2}{\sigma^2} = \lambda_1$$

While it may seem strange that $\lambda_1 \geq \lambda_2$, it is clear that we would expect $\lambda_3 \geq \lambda_2$; otherwise, we could increase the noncentrality parameter for testing $H_0 : \beta_1 = 0$ by simply including irrelevant explanatory variables in the model.

Example 2

For a second example, consider the hypothesis $H_0 : \beta_1 = 0, \beta_1 = \beta_2$ where $\beta_1 \neq 0$ and $\beta_1 = \beta_2$. In this case

$$R = \begin{pmatrix} R_1 \\ R_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix} \qquad r = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

The noncentrality parameter for testing the joint hypothesis $H_0:\beta_1=0, \beta_1=\beta_2$ is

$$\lambda_1=\frac{1}{\sigma^2}\begin{pmatrix}\beta_1 & 0\end{pmatrix}\left[\begin{pmatrix}1 & 0\\ 1 & -1\end{pmatrix}(X'X)^{-1}\begin{pmatrix}1 & 1\\ 0 & -1\end{pmatrix}\right]^{-1}\begin{pmatrix}\beta_1\\ 0\end{pmatrix}$$

$$=\frac{\sum_{i=1}^N x_{1i}^2\sum_{i=1}^N x_{2i}^2-\left(\sum_{i=1}^N x_{1i}x_{2i}\right)^2}{\sigma^2}\begin{pmatrix}\beta_1 & 0\end{pmatrix}\left[\begin{pmatrix}1 & 0\\ 1 & -1\end{pmatrix}\begin{pmatrix}\sum_{i=1}^N x_{2i}^2 & -\sum_{i=1}^N x_{1i}x_{2i}\\ -\sum_{i=1}^N x_{1i}x_{2i} & \sum_{i=1}^N x_{1i}^2\end{pmatrix}\begin{pmatrix}1 & 1\\ 0 & -1\end{pmatrix}\right]^{-1}\begin{pmatrix}\beta_1\\ 0\end{pmatrix}$$

$$=\frac{\sum_{i=1}^N x_{1i}^2\sum_{i=1}^N x_{2i}^2-\left(\sum_{i=1}^N x_{1i}x_{2i}\right)^2}{\sigma^2}\begin{pmatrix}\beta_1 & 0\end{pmatrix}\begin{pmatrix}\sum_{i=1}^N x_{2i}^2 & \sum_{i=1}^N x_{2i}^2+\sum_{i=1}^N x_{1i}x_{2i}\\ \sum_{i=1}^N x_{2i}^2+\sum_{i=1}^N x_{1i}x_{2i} & \sum_{i=1}^N x_{1i}^2\sum_{i=1}^N x_{2i}^2+2\sum_{i=1}^N x_{1i}x_{2i}\end{pmatrix}^{-1}\begin{pmatrix}\beta_1\\ 0\end{pmatrix}$$

$$=\frac{\sum_{i=1}^N x_{1i}^2\sum_{i=1}^N x_{2i}^2-\left(\sum_{i=1}^N x_{1i}x_{2i}\right)^2}{\sigma^2}\begin{pmatrix}\beta_1 & 0\end{pmatrix}\begin{pmatrix}\sum_{i=1}^N x_{2i}^2 & \sum_{i=1}^N z_i x_{2i}\\ \sum_{i=1}^N z_i x_{2i} & \sum_{i=1}^N z_i^2\end{pmatrix}^{-1}\begin{pmatrix}\beta_1\\ 0\end{pmatrix}$$

where $z_i=x_{1i}+x_{2i}$. It can be shown that $\sum_{i=1}^N x_{1i}^2\sum_{i=1}^N x_{2i}^2-\left(\sum_{i=1}^N x_{1i}x_{2i}\right)^2=\sum_{i=1}^N z_i^2\sum_{i=1}^N x_{2i}^2-\left(\sum_{i=1}^N z_i x_{2i}\right)^2$, and hence that

$$\lambda_1=\frac{\beta_1^2\sum_{i=1}^N z_i^2}{\sigma^2}$$

Omitting $\beta_1=\beta_2$ from the null hypothesis and testing only $H_0:\beta_1=0$ yields the same noncentrality parameter as in Example 1, namely

$$\lambda_2=\frac{\beta_1^2\sum_{i=1}^N x_{1i}^2}{\sigma^2}\left(1-r^2\right)$$

To test $H_0:\beta_1=0$ on the model with the restriction $\beta_1=\beta_2$ imposed, we note that the restricted model is

$$E\left(y_i\mid X\right)=\left(x_{1i}+x_{2i}\right)\beta_1=z_i\beta_1$$

and the noncentrality parameter is

$$\lambda_3=\frac{\beta_1^2\sum_{i=1}^N z_i^2}{\sigma^2}=\lambda_1$$

After some algebra, we can show that

$$\lambda_1-\lambda_2=\frac{\beta_1^2\sum_{i=1}^N z_i^2}{\sigma^2}-\frac{\beta_1^2\sum_{i=1}^N x_{1i}^2}{\sigma^2}\left(1-r^2\right)$$
$$=\frac{\beta_1^2}{\sigma^2\sum_{i=1}^N x_{2i}^2}\left(\sum_{i=1}^N x_{2i}^2+\sum_{i=1}^N x_{1i}x_{2i}\right)^2>0$$

Thus, $\lambda_1 = \lambda_3 > \lambda_2$. Moreover, when x_1 and x_2 are uncorrelated, $\lambda_1 - \lambda_2 = \beta_1^2 \sum_{i=1}^{N} x_{2i}^2 / \sigma^2 > 0$; in contrast to Example 1, the strict inequality always holds.

3. General Case

To explore the general case of testing $H_0 : R\beta = r$ when $R\beta - r$ can be partitioned as

$$\begin{pmatrix} R_1\beta \\ R_2\beta \end{pmatrix} - \begin{pmatrix} r_1 \\ r_2 \end{pmatrix} = \begin{pmatrix} \delta \\ 0 \end{pmatrix}$$

with $\delta \neq 0$, we adopt the notation used by Ruud (2000, p.233-235) and set $AA' = \sigma^2 R\left(X'X\right)^{-1} R'$, with *A* being lower triangular, partitioned conformably with *R*,

$$A = \begin{pmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{pmatrix}$$

and with the inverse of *A* given by

$$B = \begin{pmatrix} B_{11} & 0 \\ B_{21} & B_{22} \end{pmatrix}$$

It follows that

$$\begin{aligned} \sigma^2 R\left(X'X\right)^{-1} R' &= \sigma^2 \begin{pmatrix} R_1\left(X'X\right)^{-1} R_1' & R_1\left(X'X\right)^{-1} R_2' \\ R_2\left(X'X\right)^{-1} R_1' & R_2\left(X'X\right)^{-1} R_2' \end{pmatrix} \\ &= \begin{pmatrix} A_{11}A_{11}' & A_{11}A_{21}' \\ A_{21}A_{11}' & A_{21}A_{21}' + A_{21}A_{21}' \end{pmatrix} \end{aligned} \qquad (1)$$

and

$$\begin{aligned} \left[\sigma^2 R\left(X'X\right)^{-1} R'\right]^{-1} &= \left(AA'\right)^{-1} = A'^{-1}A^{-1} = B'B \\ &= \begin{pmatrix} B_{11}'B_{11} + B_{21}'B_{21} & B_{21}'B_{22} \\ B_{22}'B_{21} & B_{22}'B_{22} \end{pmatrix} \end{aligned} \qquad (2)$$

The noncentrality parameter can be written as

$$\begin{aligned} \lambda_1 &= \left(\delta' \quad 0'\right)\left[\sigma^2 R\left(X'X\right)^{-1} R'\right]^{-1} \begin{pmatrix} \delta \\ 0 \end{pmatrix} \\ &= \left(\delta' \quad 0'\right)\begin{bmatrix} B_{11}'B_{11} + B_{21}'B_{21} & B_{21}'B_{22} \\ B_{22}'B_{21} & B_{22}'B_{22} \end{bmatrix}\begin{pmatrix} \delta \\ 0 \end{pmatrix} \\ &= \delta'\left(B_{11}'B_{11} + B_{21}'B_{21}\right)\delta \end{aligned}$$

Now, from (1), $B_{11}'B_{11} = \left(A_{11}A_{11}'\right)^{-1} = \left[\sigma^2 R_1 \left(X'X\right)^{-1} R_1'\right]^{-1}$, and thus,

$$\lambda_1 = \delta'\left[\sigma^2 R_1 \left(X'X\right)^{-1} R_1'\right]^{-1}\delta + \delta' B_{21}' B_{21}\delta$$

Discarding the surplus restrictions $R_2\beta = r_2$ from the null hypothesis and testing $H_0 : R_1\beta = r_1$, we get the noncentrality parameter

$$\lambda_2 = \delta'\left[\sigma^2 R_1 \left(X'X\right)^{-1} R_1'\right]^{-1}\delta$$

Since $B_{21}'B_{21}$ is nonnegative, we have $\lambda_1 \geq \lambda_2$; discarding the surplus restrictions decreases the non-centrality parameter and hence may decrease the power of the test if it is not compensated by a reduction in the degrees of freedom. We provide some insights into the conditions under which $B_{21}'B_{21} = 0$ and hence $\lambda_1 = \lambda_2$ after showing that $\lambda_3 = \lambda_1$.

Estimating β subject to the valid restrictions $R_2\beta = r_2$ yields the restricted estimator

$$\hat{\beta}_R = \hat{\beta} - \left(X'X\right)^{-1} R_2'\left[R_2\left(X'X\right)^{-1} R_2'\right]^{-1}\left(R_2\hat{\beta} - r_2\right)$$

with covariance matrix (see, for example, Judge et al. 1988, p.237-239)

$$V\left(\hat{\beta}_R \mid X\right) = \sigma^2\left(X'X\right)^{-1} - \sigma^2\left(X'X\right)^{-1} R_2'\left[R_2\left(X'X\right)^{-1} R_2'\right]^{-1} R_2\left(X'X\right)^{-1}$$

The non-centrality parameter for testing $H_0 : R_1\beta = r_1$ using $\hat{\beta}_R$ is

$$\begin{aligned}\lambda_3 &= \delta'\left[\sigma^2 R_1\left(X'X\right)^{-1} R_1' - \sigma^2 R_1\left(X'X\right)^{-1} R_2'\left[R_2\left(X'X\right)^{-1} R_2'\right]^{-1} R_2\left(X'X\right)^{-1} R_1'\right]^{-1}\delta \\ &= \delta'\left[A_{11}A_{11}' - A_{11}A_{21}'\left(A_{21}A_{21}' + A_{22}A_{22}'\right)^{-1} A_{21}A_{11}'\right]^{-1}\delta\end{aligned}$$

The term $\left[A_{11}A_{11}' - A_{11}A_{21}'\left(A_{21}A_{21}' + A_{22}A_{22}'\right)^{-1} A_{21}A_{11}'\right]^{-1}$ can be recognised as the top-left block in the partitioned inverse of AA'. The inverse of AA' is $B'B$ whose top-left block is $B_{11}'B_{11} + B_{21}'B_{21}$. Hence,

$$\lambda_3 = \delta'\left(B_{11}'B_{11} + B_{21}'B_{21}\right)\delta = \lambda_1$$

To examine the conditions under which $\lambda_1 = \lambda_2 = \lambda_3$, suppose that the model can be rewritten as

$$E(y \mid X) = X_1\beta_1 + X_2\beta_2 \tag{3}$$

where X_1 is $(N \times K_1)$ with $J_1 \leq K_1$, X_2 is $(N \times K_2)$ with $J_2 \leq K_2$, and $J_1 + J_2 < K_1 + K_2 = K$. Also suppose that the restrictions $R\beta = r$ are of the form

$$R\beta = \begin{pmatrix} R_1 \\ R_2 \end{pmatrix} \beta = \begin{pmatrix} R_{11} & 0 \\ 0 & R_{22} \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = \begin{pmatrix} r_1 \\ r_2 \end{pmatrix} \tag{4}$$

with dimensions conformable to the partition $X = \begin{pmatrix} X_1 & X_2 \end{pmatrix}$. There is no overlap between the coefficients in the invalid restrictions $R_{11}\beta_1 = r_1$ and those in the valid restrictions $R_{22}\beta_2 = r_2$. We encountered this situation in Example 1 where we found that $\lambda_1 = \lambda_2 = \lambda_3$ if x_1 and x_2 were uncorrelated. Thus, we suspect the requirement for $\lambda_1 = \lambda_2 = \lambda_3$ in the more general set up in (3) and (4) would be that X_1 and X_2 are orthogonal.

Noting that

$$B_{11}'B_{11} + B_{21}'B_{21} = \left[\sigma^2 R_1 \left(X'X \right)^{-1} R_1' - \sigma^2 R_1 \left(X'X \right)^{-1} R_2' \left[R_2 \left(X'X \right)^{-1} R_2' \right]^{-1} R_2 \left(X'X \right)^{-1} R_1' \right]^{-1}$$

it follows that $B_{11}'B_{11} + B_{21}'B_{21} = B_{11}'B_{11} = \left[\sigma^2 R_1 \left(X'X \right)^{-1} R_1' \right]^{-1}$ if

$$\sigma^2 R_1 \left(X'X \right)^{-1} R_2' \left[R_2 \left(X'X \right)^{-1} R_2' \right]^{-1} R_2 \left(X'X \right)^{-1} R_1' = 0$$

Partitioning $\left(X'X \right)^{-1}$ in line with (3),

$$\left(X'X \right)^{-1} = \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix}$$

we find that $R_2 \left(X'X \right)^{-1} R_2' = R_{22} Q_{22} R_{22}'$, $R_1 \left(X'X \right)^{-1} R_2' = R_{11} Q_{12} R_{22}'$, and

$$\sigma^2 R_1 \left(X'X \right)^{-1} R_2' \left[R_2 \left(X'X \right)^{-1} R_2' \right]^{-1} R_2 \left(X'X \right)^{-1} R_1' = \sigma^2 R_{11} Q_{12} R_{22}' \left(R_{22} Q_{22} R_{22}' \right)^{-1} R_{22} Q_{21} R_{11}'$$

This quantity will be zero if $Q_{12} = 0$ which in turn will be true if X_1 and X_2 are orthogonal. Thus, the three noncentrality parameters are equal if the valid coefficient restrictions are separate from the invalid ones, and the explanatory variables associated with the valid and invalid restrictions are orthogonal. As we saw in Example 2, it is not sufficient for X_1 and X_2 to be orthogonal; we also require the two sets of restrictions to be separable.

4. Numerical Examples

Example 1

To illustrate how the power of a test that includes a valid hypothesis can be greater than the test that excludes it, consider the regression model $y_i = \beta_1 + \beta_2 x_i + e_i$, $i = 1,\ldots,40$. Let $x_i = 10,\ i = 1,\ldots,20$ and $x_i = 20,\ i = 21,\ldots,40$. We have $N = 40$, $\sum_{i=1}^{40} x_i = 600$ and $\sum_{i=1}^{40} x_i^2 = 10{,}000$. Let the true parameter values be $\beta_1 = 100$, $\beta_2 = 10$ and $\sigma^2 = 2500$. First consider testing the joint hypothesis $H_0 : \beta_1 = 100, \beta_2 = 9$ against the alternative $H_1 : \beta_1 \neq 100 \text{ and/or } \beta_2 \neq 9$. At the 5% level of significance we reject the joint null hypothesis if the F-test statistic is greater than the critical value $F_{(0.95,2,38)} = 3.24482$. The noncentrality parameter is

$$\begin{aligned}\lambda_1 &= \begin{pmatrix}\beta_1 - 100 & \beta_2 - 9\end{pmatrix}\frac{X'X}{\sigma^2}\begin{pmatrix}\beta_1 - 100 \\ \beta_2 - 9\end{pmatrix} \\ &= \frac{1}{2500}\begin{pmatrix}0 & 1\end{pmatrix}\begin{pmatrix}40 & 600 \\ 600 & 10{,}000\end{pmatrix}\begin{pmatrix}0 \\ 1\end{pmatrix} \\ &= 4\end{aligned}$$

The probability of rejecting the joint null hypothesis is the probability that a value from a non-central F-distribution with noncentrality parameter $\lambda = 4$ will exceed $F_{(0.95,2,38)} = 3.24482$. This value for the test power is $\Pr\left(F_{(2,38,\lambda=4)} > 3.24482\right) = 0.38738$.

Now we omit the true hypothesis $\beta_1 = 100$, and test $H_0 : \beta_2 = 9$ against $H_1 : \beta_2 \neq 9$ using an F-test. The test critical value is the 95th percentile of the F-distribution, $F_{(0.95,1,38)} = 4.09817$. The noncentrality parameter of the F-distribution for this single hypothesis is

$$\begin{aligned}\lambda_2 &= (\beta_2 - 9)^2\left(\begin{pmatrix}0 & 1\end{pmatrix}\sigma^2\left(X'X\right)^{-1}\begin{pmatrix}0 \\ 1\end{pmatrix}\right)^{-1} \\ &= \frac{1}{2500}\left(\begin{pmatrix}0 & 1\end{pmatrix}\begin{pmatrix}40 & 600 \\ 600 & 10{,}000\end{pmatrix}^{-1}\begin{pmatrix}0 \\ 1\end{pmatrix}\right)^{-1} \\ &= 0.4\end{aligned}$$

The probability of rejecting the null hypothesis $H_0 : \beta_2 = 9$ versus $H_1 : \beta_2 \neq 9$ when $\beta_2 = 10$ is $\Pr\left(F_{(1,38,\lambda=0.4)} > 4.09817\right) = 0.09457$. Removing the true hypothesis has reduced the power of the test.

Example 2

All the above results have been conditional on X. To investigate the effect of relaxing this assumption, we again use the model in Example 1, but generate 10,000 samples of x from a $N\left(15, 1.6^2\right)$ distribution, and, correspondingly, 10,000 samples of $(y \mid x)$ from $N\left(100 + 10x, 50^2\right)$. This choice of x gives values from 10.2 to 19.8 with probability over 0.99, mimicking the range of x-values in Example 1. Using the F-test, the percentage of rejections for the single hypothesis was 0.0569, and for the joint hypothesis, 0.3545. Using the chi-square test, with chi-square value equal to the F-value multiplied by the numerator degrees of freedom, the single hypothesis rejection percentage was 0.0640, and the joint test rejection percentage was 0.3911. Removing the true hypothesis reduced the simulated power of the test.

5. Concluding Remark

We have shown that, except under special circumstances, including valid superfluous restrictions in an F-test for linear restrictions in a regression model increases the non-centrality parameter of the F-statistic, and hence can also improve its power. This apparently counterintuitive result makes sense when we observe that the noncentrality parameter from testing both sets of restrictions jointly is identical to that obtained when the F-test for the invalid restrictions is applied to a model that has been restricted in line with the valid restrictions.

References

Judge, G.G., R.C. Hill, W.E. Griffiths, H. Lütkepohl and T.-C. Lee (1988), *Introduction to the Theory and Practice of Econometrics*, second edition, New York: John Wiley and Sons.

Ruud, P. A. (2000), *An Introduction to Classical Econometric Theory*, Oxford: Oxford University Press.