PhD Econometrics 1: Study Questions Week 1
Imperial College London
Hormoz Ramian

Solutions

Question 1

(1.1) The test setup is:

$$\begin{aligned} H_0 &: \boldsymbol{R\beta} - \boldsymbol{r} = \boldsymbol{0} \\ H_A &: \boldsymbol{R\beta} - \boldsymbol{r} \neq \boldsymbol{0} \end{aligned}$$

at a given confidence level and using theoretical F-distribution critical values.

(1.2) The Lagrangian to this constrained problem is:

$$\mathcal{L} = \frac{1}{2}(\boldsymbol{y} - \boldsymbol{X\beta})'(\boldsymbol{y} - \boldsymbol{X\beta}) + \boldsymbol{\lambda}'(\boldsymbol{R\beta} - \boldsymbol{r})$$

where $\boldsymbol{\beta}$ is the parameter vector of restricted least squares or constrained least squares (CLS).

(1.3) The first order condition(s) with respect to $\boldsymbol{\beta}$ is:

$$\boldsymbol{X}'\boldsymbol{y} - (\boldsymbol{X}'\boldsymbol{X})\widehat{\boldsymbol{\beta}} = \boldsymbol{R}'\boldsymbol{\lambda} \tag{1}$$

or just $\boldsymbol{X}'(\boldsymbol{y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}}) = \boldsymbol{R}'\boldsymbol{\lambda}$. The first order condition(s) with respect to $\boldsymbol{\lambda}$ is:

(1.4) The solution including the Lagrange multiplier (vector) $\boldsymbol{\lambda}$ is:

$$\boldsymbol{0} = \boldsymbol{R}\widehat{\boldsymbol{\beta}} - \boldsymbol{r} \tag{2}$$

Now we'll need to solve the two first order conditions for the values of $\widehat{\boldsymbol{\beta}}$ and $\boldsymbol{\lambda}$ that satisfy the Lagrangian problem. First pre-multiply by $\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}$:

$$\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}\left[\boldsymbol{X}'\boldsymbol{y} - (\boldsymbol{X}'\boldsymbol{X})\widehat{\boldsymbol{\beta}}\right] = \boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}\left[\boldsymbol{R}'\boldsymbol{\lambda}\right]$$

and simplify to get:

$$\begin{aligned} \boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y} - \boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}(\boldsymbol{X}'\boldsymbol{X})\widehat{\boldsymbol{\beta}} &= \boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}'\boldsymbol{\lambda} \\ \boldsymbol{R}\tilde{\boldsymbol{\beta}} - \boldsymbol{R}\widehat{\boldsymbol{\beta}} &= \boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}'\boldsymbol{\lambda} \end{aligned}$$

where $\tilde{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}$ and $(\boldsymbol{X}'\boldsymbol{X})^{-1}(\boldsymbol{X}'\boldsymbol{X}) = \boldsymbol{I}$. Note that under the null hypothesis $\boldsymbol{R}\widehat{\boldsymbol{\beta}} = \boldsymbol{r}$:

$$\boldsymbol{R}\tilde{\boldsymbol{\beta}} - \boldsymbol{r} = \boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}'\boldsymbol{\lambda}$$

Solving for $\boldsymbol{\lambda}$:

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

The expression above does not include $\widehat{\boldsymbol{\beta}}$ and hence is the solution for $\boldsymbol{\lambda}$. Substituting this into equation (1) gives:

$$
\begin{aligned}
\boldsymbol{X}'\boldsymbol{y}-(\boldsymbol{X}'\boldsymbol{X})\widehat{\boldsymbol{\beta}} &= \boldsymbol{R}'\boldsymbol{\lambda}\\
\boldsymbol{X}'\boldsymbol{y}-(\boldsymbol{X}'\boldsymbol{X})\widehat{\boldsymbol{\beta}} &= \boldsymbol{R}'\left[\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}'\right]^{-1}\left(\boldsymbol{R}\tilde{\boldsymbol{\beta}}-\boldsymbol{r}\right)\\
&\Leftrightarrow\\
(\boldsymbol{X}'\boldsymbol{X})\widehat{\boldsymbol{\beta}} &= \boldsymbol{X}'\boldsymbol{y}-\boldsymbol{R}'\left[\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}'\right]^{-1}\left(\boldsymbol{R}\tilde{\boldsymbol{\beta}}-\boldsymbol{r}\right)\\
&\Leftrightarrow\\
\widehat{\boldsymbol{\beta}} &= (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}-(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}'\left[\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}'\right]^{-1}\left(\boldsymbol{R}\tilde{\boldsymbol{\beta}}-\boldsymbol{r}\right)\\
\widehat{\boldsymbol{\beta}} &= (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}-(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}'\left[\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}'\right]^{-1}\left(\boldsymbol{R}\tilde{\boldsymbol{\beta}}-\boldsymbol{r}\right)\\
&= \tilde{\boldsymbol{\beta}}-(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}'\left[\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}'\right]^{-1}\left(\boldsymbol{R}\tilde{\boldsymbol{\beta}}-\boldsymbol{r}\right)
\end{aligned}
$$

The solution to $\widehat{\boldsymbol{\beta}}$ and $\boldsymbol{\lambda}$ is:

$$
\begin{aligned}
\widehat{\boldsymbol{\beta}} &= \tilde{\boldsymbol{\beta}}-(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}'[\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}']^{-1}(\boldsymbol{R}\tilde{\boldsymbol{\beta}}-\boldsymbol{r})\\
\boldsymbol{\lambda} &= [\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}]^{-1}(\boldsymbol{R}\tilde{\boldsymbol{\beta}}-\boldsymbol{r})
\end{aligned}
$$

The Lagrange multiplier is a vector with non-negative elements which measures whether constraint is binding or slack. The interpretation of the Lagrange multiplier depends on economic meaning of each variable. In general, however, this links the constraint to the sum of squared residuals.

(1.5) Let $\widehat{\boldsymbol{\epsilon}} \equiv \boldsymbol{y}-\boldsymbol{X}\widehat{\boldsymbol{\beta}}$ be residuals from the restricted regression:

$$
\begin{aligned}
\widehat{\boldsymbol{\epsilon}} &\equiv \boldsymbol{y}-\boldsymbol{X}\widehat{\boldsymbol{\beta}}\\
&= \boldsymbol{y}-\boldsymbol{X}\widehat{\boldsymbol{\beta}}+\boldsymbol{X}\tilde{\boldsymbol{\beta}}-\boldsymbol{X}\tilde{\boldsymbol{\beta}}\\
&= (\boldsymbol{y}-\boldsymbol{X}\tilde{\boldsymbol{\beta}})+\boldsymbol{X}(\tilde{\boldsymbol{\beta}}-\widehat{\boldsymbol{\beta}})
\end{aligned}
$$

Hence $SSR_R \equiv \widehat{\boldsymbol{\epsilon}}'\widehat{\boldsymbol{\epsilon}}$ is:

$$
\begin{aligned}
\widehat{\boldsymbol{\epsilon}}'\widehat{\boldsymbol{\epsilon}} &= \left[(\boldsymbol{y}-\boldsymbol{X}\tilde{\boldsymbol{\beta}})+\boldsymbol{X}(\tilde{\boldsymbol{\beta}}-\widehat{\boldsymbol{\beta}})\right]'\left[(\boldsymbol{y}-\boldsymbol{X}\tilde{\boldsymbol{\beta}})+\boldsymbol{X}(\tilde{\boldsymbol{\beta}}-\widehat{\boldsymbol{\beta}})\right]\\
&= (\boldsymbol{y}-\boldsymbol{X}\tilde{\boldsymbol{\beta}})'(\boldsymbol{y}-\boldsymbol{X}\tilde{\boldsymbol{\beta}})+(\tilde{\boldsymbol{\beta}}-\widehat{\boldsymbol{\beta}})'\boldsymbol{X}'\boldsymbol{X}(\tilde{\boldsymbol{\beta}}-\widehat{\boldsymbol{\beta}})
\end{aligned}
$$

Noting, firstly that $\boldsymbol{X}'(\boldsymbol{y}-\boldsymbol{X}\boldsymbol{\beta})=0$, and secondly, $SSR_U=(\boldsymbol{y}-\boldsymbol{X}\tilde{\boldsymbol{\beta}})'(\boldsymbol{y}-\boldsymbol{X}\tilde{\boldsymbol{\beta}})$, then:

$$
SSR_R-SSR_U = (\tilde{\boldsymbol{\beta}}-\widehat{\boldsymbol{\beta}})'\boldsymbol{X}'\boldsymbol{X}(\tilde{\boldsymbol{\beta}}-\widehat{\boldsymbol{\beta}})
$$

Using the results $\widehat{\boldsymbol{\beta}}=\tilde{\boldsymbol{\beta}}-(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}'[\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}']^{-1}(\boldsymbol{R}\tilde{\boldsymbol{\beta}}-\boldsymbol{r})$, to simplify the RHS:

$$
\begin{aligned}
. &= (\boldsymbol{R}\tilde{\boldsymbol{\beta}}-\boldsymbol{r})'[\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}']^{-1}\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}'[\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}']^{-1}(\boldsymbol{R}\tilde{\boldsymbol{\beta}}-\boldsymbol{r})\\
&= (\boldsymbol{R}\tilde{\boldsymbol{\beta}}-\boldsymbol{r})'[\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}']^{-1}\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}'[\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}']^{-1}(\boldsymbol{R}\tilde{\boldsymbol{\beta}}-\boldsymbol{r})\\
&= (\boldsymbol{R}\tilde{\boldsymbol{\beta}}-\boldsymbol{r})'[\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}']^{-1}(\boldsymbol{R}\tilde{\boldsymbol{\beta}}-\boldsymbol{r})
\end{aligned}
$$

Using $\boldsymbol{\lambda}=[\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}]^{-1}(\boldsymbol{R}\tilde{\boldsymbol{\beta}}-\boldsymbol{r})$, then:

$$
\begin{aligned}
SSR_R-SSR_U &= (\boldsymbol{R}\tilde{\boldsymbol{\beta}}-\boldsymbol{r})'[\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}']^{-1}(\boldsymbol{R}\tilde{\boldsymbol{\beta}}-\boldsymbol{r})\\
&= (\boldsymbol{R}\tilde{\boldsymbol{\beta}}-\boldsymbol{r})'\boldsymbol{\lambda}\\
&= (\boldsymbol{R}\tilde{\boldsymbol{\beta}}-\boldsymbol{r})'\left\{[\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}']^{-1}[\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}']\right\}\boldsymbol{\lambda}\\
&= \boldsymbol{\lambda}'[\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}]\boldsymbol{\lambda}
\end{aligned}
$$

Use first order conditions, $\boldsymbol{X}'\boldsymbol{y}-(\boldsymbol{X}'\boldsymbol{X})\widehat{\boldsymbol{\beta}}=\boldsymbol{R}'\boldsymbol{\lambda}$, then $\boldsymbol{X}'(\boldsymbol{y}-\boldsymbol{X}\widehat{\boldsymbol{\beta}})=\boldsymbol{R}'\boldsymbol{\lambda}$, and then $\boldsymbol{X}'\widehat{\epsilon}=\boldsymbol{R}'\boldsymbol{\lambda}$. The interpretation is that, given the constraint, then the inner product $\boldsymbol{X}'\widehat{\epsilon}$ is not zero and indeed equal to a function of $\boldsymbol{\lambda}$:

$$
\begin{aligned}
SSR_R - SSR_U &= \boldsymbol{\lambda}'[\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}']\boldsymbol{\lambda} \\
&= \widehat{\boldsymbol{\epsilon}}'\boldsymbol{P}\widehat{\boldsymbol{\epsilon}}
\end{aligned}
$$

where $\boldsymbol{P}$ is the projection matrix.

Question 2 Note that $\mathbb{E}(y_i|x_i) = x_i\beta_1 + x_i^2\beta_2$ then $\mathbb{E}(y_i|x_i = 40) = 40\beta_1 + 40^2\beta_2$ and the hypothesis is:

$$
H_0 \;:\; 40\beta_1 + 40^2\beta_2 = 20 \tag{3}
$$

which is a linear restriction. If desired, this can be rewritten as,

$$
H_0 \;:\; 2\beta_1 + 80\beta_2 = 1 \tag{4}
$$

At this stage, we have one linear constraint on parameters which we can, first vectorize using $\boldsymbol{R}$ and $\boldsymbol{r}$ and then feed into the F-test setup.

Question 3

(3.1) In order to show that $(\widehat{\gamma}_1,\ \widehat{\gamma}_2)' = (\overline{y}_1,\ \overline{y}_2)'$, I re-write the regression model in matrix form and compute the inverse of the quadratic form for explanatory variables. Next, note that because $\boldsymbol{d}_1$ and $\boldsymbol{d}_2$ are binary variables and are perfectly orthogonal, then their inner product is zero, the inverse form simplifies to a inverse of diagonal elements inside the matrix:

$$
\begin{aligned}
\begin{pmatrix}\widehat{\gamma}_1\\ \widehat{\gamma}_2\end{pmatrix} &= \begin{pmatrix}\boldsymbol{d}_1'\boldsymbol{d}_1 & \boldsymbol{d}_1'\boldsymbol{d}_2\\ \boldsymbol{d}_2'\boldsymbol{d}_1 & \boldsymbol{d}_2'\boldsymbol{d}_2\end{pmatrix}^{-1}\begin{pmatrix}\boldsymbol{d}_1'\\ \boldsymbol{d}_2'\end{pmatrix}\boldsymbol{y} \\
&= \begin{pmatrix}\sum_{i\in n_1} d_{1,i}^2 & 0\\ 0 & \sum_{i\in n_1} d_{2,i}^2\end{pmatrix}^{-1}\begin{pmatrix}\boldsymbol{d}_1'\\ \boldsymbol{d}_2'\end{pmatrix}\boldsymbol{y} \\
&= \begin{pmatrix}n_1 & 0\\ 0 & n_2\end{pmatrix}^{-1}\begin{pmatrix}\boldsymbol{d}_1'\boldsymbol{y}\\ \boldsymbol{d}_2'\boldsymbol{y}\end{pmatrix} = \begin{pmatrix}\frac{1}{n_1} & 0\\ 0 & \frac{1}{n_2}\end{pmatrix}\begin{pmatrix}\sum_{i\in n_1} y_i\\ \sum_{i\in n_2} y_i\end{pmatrix} = \begin{pmatrix}\frac{1}{n_1}\sum_{i\in n_1} y_i\\ \frac{1}{n_2}\sum_{i\in n_2} y_i\end{pmatrix} = \begin{pmatrix}\overline{y}_1\\ \overline{y}_2\end{pmatrix}
\end{aligned}
$$

(3.2) Under the definition of the model, we have $\widehat{\gamma}_1 = \widetilde{\gamma}_1$ since $\boldsymbol{d}_1$ and $\boldsymbol{d}_2$ are perfectly orthogonal, indicating that we can split the regression for each sample and run individual regressions. Alternatively, we note that because of orthogonality, each original coefficient $\widehat{\gamma}_i$ only measures the impact from its respective variable. Algebraically, running the regression without either of the variables, amounts to no estimation bias due to omitted variable.

Question 4 No, the estimates are different, except in the special case that the regressors are uncorrelated in-sample: $\boldsymbol{X}_1'\boldsymbol{X}_2 = 0$. We can calculate that:

$$
\begin{aligned}
\overline{\boldsymbol{\beta}}_1 &= (\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1'\boldsymbol{Y} \\
\tilde{\boldsymbol{u}} &= \boldsymbol{M}_1\boldsymbol{Y} \\
\boldsymbol{M}_1 &= I - \boldsymbol{X}_1(\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1' \\
\overline{\boldsymbol{\beta}}_2 &= (\boldsymbol{X}_2'\boldsymbol{X}_2)^{-1}\boldsymbol{X}_2'\tilde{\boldsymbol{u}} \\
&= (\boldsymbol{X}_2'\boldsymbol{X}_2)^{-1}\boldsymbol{X}_2'\boldsymbol{M}_1\boldsymbol{Y}
\end{aligned}
$$

then[1]:

$$\widehat{\boldsymbol{\beta}}_2 = (\boldsymbol{X}_2'\boldsymbol{M}_1\boldsymbol{X}_2)^{-1}\boldsymbol{X}_2'\boldsymbol{M}_1\boldsymbol{Y}$$

These two expressions are different, except when:

$$\begin{aligned}\boldsymbol{X}_2'\boldsymbol{X}_2 &= \boldsymbol{X}_2'\boldsymbol{M}_1\boldsymbol{X}_2 \\ &= \boldsymbol{X}_2'\boldsymbol{X}_2 - \boldsymbol{X}_2'\boldsymbol{X}_1(\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1'\boldsymbol{X}_2\end{aligned}$$

which happens when $\boldsymbol{X}_1'\boldsymbol{X}_2 = 0$. Another way to see the difference is to take the expression for $\widetilde{\boldsymbol{\beta}}$ and substitute in the OLS estimates on the full model,

$$\boldsymbol{Y} = \boldsymbol{X}_1\widehat{\boldsymbol{\beta}}_1 + \boldsymbol{X}_2\widehat{\boldsymbol{\beta}}_2 + \widehat{\boldsymbol{e}}$$

to find:

$$\begin{aligned}\overline{\boldsymbol{\beta}} &= (\boldsymbol{X}_2'\boldsymbol{X}_2)^{-1}\boldsymbol{X}_2'\boldsymbol{M}_1\boldsymbol{Y} \\ &= (\boldsymbol{X}_2'\boldsymbol{X}_2)^{-1}\boldsymbol{X}_2'\boldsymbol{M}_1(\boldsymbol{X}_1\widehat{\boldsymbol{\beta}}_1 + \boldsymbol{X}_2\widehat{\boldsymbol{\beta}}_2 + \widehat{e}) \\ &= (\boldsymbol{X}_2'\boldsymbol{X}_2)^{-1}\boldsymbol{X}_2'\boldsymbol{M}_1\boldsymbol{X}_1\widehat{\boldsymbol{\beta}}_1 + (\boldsymbol{X}_2'\boldsymbol{X}_2)^{-1}\boldsymbol{X}_2'\boldsymbol{M}_1\boldsymbol{X}_2\widehat{\boldsymbol{\beta}}_2 + (\boldsymbol{X}_2'\boldsymbol{X}_2)^{-1}\boldsymbol{X}_2'\boldsymbol{M}_1\widehat{\boldsymbol{e}}) \\ &= (\boldsymbol{X}_2'\boldsymbol{X}_2)^{-1}\boldsymbol{X}_2'\boldsymbol{M}_1\boldsymbol{X}_2\widehat{\boldsymbol{\beta}}_2 \\ &= \widehat{\boldsymbol{\beta}}_2 - (\boldsymbol{X}_2'\boldsymbol{X}_2)^{-1}\boldsymbol{X}_2'\boldsymbol{X}_1(\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1'\boldsymbol{X}_2\widehat{\boldsymbol{\beta}}_2 \\ &\neq \widehat{\boldsymbol{\beta}}_2\end{aligned}$$

Noting that $\boldsymbol{M}_1\boldsymbol{X}_1 = \mathbf{0}$ and $\boldsymbol{M}_1\widehat{e} = \boldsymbol{M}_1\boldsymbol{M}\boldsymbol{Y} = \mathbf{0}$.

[1] The following results is called the Frisch-Waugh-Lovell Theorem (FWL): Pre-multiply $y_i = \boldsymbol{x}_{1i}'\widehat{\boldsymbol{\beta}}_1 + \boldsymbol{x}_{2i}'\widehat{\boldsymbol{\beta}}_2 + \widehat{e}_i$ by $\boldsymbol{M}_1$ which leaves us with $\boldsymbol{M}_1 y_i = \boldsymbol{M}_1\boldsymbol{x}_{2i}'\widehat{\boldsymbol{\beta}}_2 + \boldsymbol{M}_1\widehat{e}_i$, then apply usual OLS estimator's formula.