Introduction to Econometrics (4th Updated Edition)

by

James H. Stock and Mark W. Watson

Solutions to Odd-Numbered End-of-Chapter Exercises: Chapter 19

(This version September 18, 2018)

19.1. (a) The regression in the matrix form is

$$\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{U}$$

with

$$\mathbf{Y} = \begin{pmatrix} TestScore_1 \\ TestScore_2 \\ \vdots \\ TestScore_n \end{pmatrix}, \quad \mathbf{X} = \begin{pmatrix} 1 & Income_1 & Income_1^2 \\ 1 & Income_2 & Income_2^2 \\ \vdots & \vdots & \vdots \\ 1 & Income_n & Income_n^2 \end{pmatrix}$$

$$\mathbf{U} = \begin{pmatrix} U_1 \\ U_2 \\ \vdots \\ U_n \end{pmatrix}, \quad \beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix}.$$

(b) The null hypothesis is

$$\mathbf{R}\boldsymbol{\beta} = \mathbf{r}$$

versus $\mathbf{R}\boldsymbol{\beta} \neq \mathbf{r}$ with

$$\mathbf{R} = (\ 0 \quad 0 \quad 1\ ) \text{ and } \mathbf{r} = 0.$$

The heteroskedasticity-robust F-statistic testing the null hypothesis is

$$F = (\mathbf{R}\hat{\boldsymbol{\beta}} - \mathbf{r})'\left[\mathbf{R}\hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\beta}}}\mathbf{R}'\right]^{-1}(\mathbf{R}\hat{\boldsymbol{\beta}} - \mathbf{r})/q$$

With $q = 1$. Under the null hypothesis,

$$F \xrightarrow{d} F_{q,\infty}.$$

We reject the null hypothesis if the calculated F-statistic is larger than the critical value of the $F_{q,\infty}$ distribution at a given significance level.

19.3. (a)

$$\begin{aligned}\text{Var}(Q) &= E[(Q-\mu_Q)^2] \\ &= E[(Q-\mu_Q)(Q-\mu_Q)'] \\ &= E[(\mathbf{c}'\mathbf{W}-\mathbf{c}'\boldsymbol{\mu}_{\mathbf{W}})(\mathbf{c}'\mathbf{W}-\mathbf{c}'\boldsymbol{\mu}_{\mathbf{W}})'] \\ &= \mathbf{c}'E[(\text{W}-\boldsymbol{\mu}_{\mathbf{W}})(\mathbf{W}-\boldsymbol{\mu}_{\mathbf{W}})']\mathbf{c} \\ &= \mathbf{c}'\,\text{var}(\mathbf{W})\,\mathbf{c} = \mathbf{c}'\boldsymbol{\Sigma}_{\mathbf{w}}\mathbf{c}\end{aligned}$$

where the second equality uses the fact that Q is a scalar and the third equality uses the fact that $\mu_Q = \mathbf{c}'\boldsymbol{\mu}_{\mathbf{W}}$.

(b) Because the covariance matrix Σ_{W} is positive definite, we have $\mathbf{c}'\Sigma_{\mathbf{w}}\mathbf{c} > 0$ for every non-zero vector from the definition. Thus, var(Q) > 0. Both the vector **c** and the matrix Σ_{W} are finite, so var(Q) = $\mathbf{c}'\Sigma_{\mathbf{w}}\mathbf{c}$ is also finite. Thus, 0 < var(Q) < ∞.

19.5. $\mathbf{P_X} = \mathbf{X}\,(\mathbf{X'X})^{\square 1}\mathbf{X'}$, $\mathbf{M_X} = \mathbf{I}_n - \mathbf{P_X}$.

(a) $\mathbf{P_X}$ is idempotent because

$$P_X P_X = X(X'X)^{-1} X'X(X'X)^{-1} X' = X(X'X)^{-1} X'\,\hat{}= P_X.$$

$\mathbf{M_X}$ is idempotent because

$$\begin{aligned}\mathbf{M_X M_X} &= (\mathbf{I}_n - \mathbf{P_X})(\mathbf{I}_n - \mathbf{P_X}) = \mathbf{I}_n - \mathbf{P_X} - \mathbf{P_X} + \mathbf{P_X P_X} \\ &= \mathbf{I}_n - 2\mathbf{P_X} + \mathbf{P_X} = \mathbf{I_n} - \mathbf{P_X} = \mathbf{M_X}\end{aligned}$$

$\mathbf{P_X M_X} = 0_{nxn}$ because

$$\mathbf{P_X M_X} = \mathbf{P_X}\,(\mathbf{I}_\mathrm{n} - \mathbf{P_X}) = \mathbf{P_X} - \mathbf{P_X P_X} = \mathbf{P_X} - \mathbf{P_X} = 0_{n\times n}$$

(b) Because $\hat{\beta} = (\mathbf{X'X})^{-1}\mathbf{X'Y}$, we have

$$\hat{\mathbf{Y}} = \mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}(\mathbf{X'X})^{-1}\mathbf{X'Y} = \mathbf{P_X Y}$$

which is Equation (19.27). The residual vector is

$$\hat{\mathbf{U}} = \mathbf{Y} - \hat{\mathbf{Y}} = \mathbf{Y} - \mathbf{P_X Y} = (\mathbf{I}_n - \mathbf{P_X})\mathbf{Y} = \mathbf{M_X Y}.$$

We know that $\mathbf{M_X X}$ is orthogonal to the columns of $\mathbf{X}$:

$$M_X X = (I_n - P_X)\,X = X = P_X X = X + X\,(X'X)^{-1}\,X'X = X - X = \mathbf{0}$$

so the residual vector can be further written as

$$\hat{\mathbf{U}} = \mathbf{M_X Y} = \mathbf{M_X}(\mathbf{X}\boldsymbol{\beta} + \mathbf{U}) = \mathbf{M_X X}\boldsymbol{\beta} + \mathbf{M_X U} = \mathbf{M_X U}$$

which is Equation (19.28).

(c) From the hint, rank (P_X) = trace(P_X) = trace$[X(X'X)^{-1}X']$ = trace$[(X'X)^{-1}X'X]$ = trace$(I_{k+1}) = k+1$. The result for M_X follows from a similar calculation.

19.7. (a) We write the regression model, $Y_i = \beta_1 X_i + \beta_2 W_i + u_i$, in the matrix form as

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta}_1 + \boldsymbol{W}\boldsymbol{\beta}_2 + \boldsymbol{U}$$

with

$$\mathbf{Y} = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix}, \quad \mathbf{X} = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix}, \quad \mathbf{W} = \begin{pmatrix} W_1 \\ W_2 \\ \vdots \\ W_n \end{pmatrix}, \quad \mathbf{U} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix},$$

The OLS estimator is

$$\begin{aligned}
\begin{pmatrix} \hat\beta_1 \\ \hat\beta_2 \end{pmatrix} &= \begin{pmatrix} \mathbf{X'X} & \mathbf{X'W} \\ \mathbf{W'X} & \mathbf{W'W} \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{X'Y} \\ \mathbf{W'Y} \end{pmatrix} \\
&= \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} + \begin{pmatrix} \mathbf{X'X} & \mathbf{X'W} \\ \mathbf{W'X} & \mathbf{W'W} \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{X'U} \\ \mathbf{W'U} \end{pmatrix} \\
&= \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} + \begin{pmatrix} \frac{1}{n}\mathbf{X'X} & \frac{1}{n}\mathbf{X'W} \\ \frac{1}{n}\mathbf{W'X} & \frac{1}{n}\mathbf{W'W} \end{pmatrix}^{-1} \begin{pmatrix} \frac{1}{n}\mathbf{X'U} \\ \frac{1}{n}\mathbf{W'U} \end{pmatrix} \\
&= \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} + \begin{pmatrix} \frac{1}{n}\sum_{i=1}^n X_i^2 & \frac{1}{n}\sum_{i=1}^n X_i W_i \\ \frac{1}{n}\sum_{i=1}^n W_i X_i & \frac{1}{n}\sum_{i=1}^n W_i^2 \end{pmatrix}^{-1} \begin{pmatrix} \frac{1}{n}\sum_{i=1}^n X_i u_i \\ \frac{1}{n}\sum_{i=1}^n W_i u_i \end{pmatrix}
\end{aligned}$$

By the law of large numbers $\frac{1}{n}\sum_{i=1}^n X_i^2 \xrightarrow{p} E(X^2)$; $\frac{1}{n}\sum_{i=1}^n W_i^2 \xrightarrow{p} E(W^2)$; $\frac{1}{n}\sum_{i=1}^n X_i W_i \xrightarrow{p} E(XW) = 0$ (because X and W are independent with means of zero); $\frac{1}{n}\sum_{i=1}^n X_i u_i \xrightarrow{p} E(Xu) = 0$ (because X and u are independent with means of zero); $\frac{1}{n}\sum_{i=1}^n X_i u_i \xrightarrow{p} E(Xu) = 0$ Thus

$$\begin{aligned}
\begin{pmatrix} \hat\beta_1 \\ \hat\beta_2 \end{pmatrix} &\xrightarrow{p} \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} + \begin{pmatrix} E(X^2) & 0 \\ 0 & E(W^2) \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ E(Wu) \end{pmatrix} \\
&= \begin{pmatrix} \beta_1 \\ \beta_2 + \frac{E(Wu)}{E(W^2)} \end{pmatrix}.
\end{aligned}$$

(b) From the answer to (a) $\hat{\beta}_2 \xrightarrow{p} \beta_2 + \frac{E(Wu)}{E(W^2)} \neq \beta_2$ if $E(Wu)$ is nonzero.

(c) Consider the population linear regression u_i onto W_i:

$$u_i = \lambda W_i + a_i$$

where $\lambda = E(Wu)/E(W^2)$. In this population regression, by construction, $E(aW) = 0$. Using this equation for u_i rewrite the equation to be estimated as

$$\begin{aligned} Y_i &= X_i\beta_1 + W_i\beta_2 + u_i \\ &= X_i\beta_1 + W_i(\beta_2 + \lambda) + a_i \\ &= X_i\beta_1 + W_i\theta + a_i \end{aligned}$$

where $\theta = \beta_2 + \lambda$. A calculation like that used in part (a) can be used to show that

$$\begin{pmatrix} \sqrt{n}(\hat{\beta}_1 - \beta_1) \\ \sqrt{n}(\hat{\beta}_2 - \theta) \end{pmatrix} = \begin{pmatrix} \frac{1}{n}\sum_{i=1}^n X_i^2 & \frac{1}{n}\sum_{i=1}^n X_i W_i \\ \frac{1}{n}\sum_{i=1}^n W_i X_i & \frac{1}{n}\sum_{i=1}^n W_i^2 \end{pmatrix}^{-1} \begin{pmatrix} \frac{1}{\sqrt{n}}\sum_{i=1}^n X_i a_i \\ \frac{1}{\sqrt{n}}\sum_{i=1}^n W_u a_i \end{pmatrix}$$
$$\xrightarrow{d} \begin{pmatrix} E(X^2) & 0 \\ 0 & E(W^2) \end{pmatrix}^{-1} \begin{pmatrix} S_1 \\ S_2 \end{pmatrix}$$

where S_1 is distributed $N(0, \sigma_a^2 E(X_2))$. Thus by Slutsky's theorem

$$\sqrt{n}(\hat{\beta}_1 - \beta_1) \xrightarrow{d} N\left(0, \frac{\sigma_a^2}{E(X^2)}\right)$$

Now consider the regression that omits W, which can be written as:

$$Y_i = X_i\beta_1 + d_i$$

where $d_i = W_i\theta + a_i$. Calculations like those used above imply that

$$\sqrt{n}\left(\hat{\beta}_1^r - \beta_1\right) \xrightarrow{d} N\left(0, \frac{\sigma_d^2}{E(X^2)}\right).$$

Since $\sigma_d^2 = \sigma_a^2 + \theta^2 E(W^2)$, the asymptotic variance of $\hat{\beta}_1^r$ is never smaller than the asymptotic variance of $\hat{\beta}_1$.

19.9. (a)

$$\begin{aligned}\hat{\boldsymbol\beta} &= (\mathbf{X}'\mathbf{M_W}\mathbf{X})^{-1}\mathbf{X}'\mathbf{M_W}\mathbf{Y}\\ &= (\mathbf{X}'\mathbf{M_W}\mathbf{X})^{-1}\mathbf{X}'\mathbf{M_W}(\mathbf{X}\boldsymbol\beta + \mathbf{W}\boldsymbol\gamma + \mathbf{U})\\ &= \boldsymbol\beta + (\mathbf{X}'\mathbf{M_W}\mathbf{X})^{-1}\mathbf{X}'\mathbf{M_W}\mathbf{U}.\end{aligned}$$

The last equality has used the orthogonality $\mathbf{M_W W = 0}$. Thus

$$\hat{\boldsymbol\beta} - \boldsymbol\beta = (\mathbf{X}'\mathbf{M_W}\mathbf{X})^{-1}\mathbf{X}'\mathbf{M_W}\mathbf{U} = (n^{-1}\mathbf{X}'\mathbf{M_W}\mathbf{X})^{-1}(n^{-1}\mathbf{X}'\mathbf{M_W}\mathbf{U}).$$

(b) Using $\mathbf{M_W} = \mathbf{I}_n - \mathbf{P_W}$ and $\mathbf{P_W} = \mathbf{W}(\mathbf{W}'\mathbf{W})^{-1}\mathbf{W}'$ we can get

$$\begin{aligned}n^{-1}\mathbf{X}'\mathbf{M_W}\mathbf{X} &= n^{-1}\mathbf{X}'(\mathbf{I}_n - \mathbf{P_W})\mathbf{X}\\ &= n^{-1}\mathbf{X}'\mathbf{X} - n^{-1}\mathbf{X}'\mathbf{P_W}\mathbf{X}\\ &= n^{-1}\mathbf{X}'\mathbf{X} - (n^{-1}\mathbf{X}'\mathbf{W})(n^{-1}\mathbf{W}'\mathbf{W})^{-1}(n^{-1}\mathbf{W}'\mathbf{X}).\end{aligned}$$

First consider $n^{-1}\mathbf{X}'\mathbf{X} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_i\mathbf{X}_i'$. The (j, l) element of this matrix is $\frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_{ji}\mathbf{X}_{li}$. By Assumption (ii), $\mathbf{X}_i$ is i.i.d., so $X_{ji}X_{li}$ is i.i.d. By Assumption (iii) each element of $\mathbf{X}_i$ has four moments, so by the Cauchy-Schwarz inequality $X_{ji}X_{li}$ has two moments:

$$E(X_{ji}^2 X_{li}^2) \le \sqrt{E(X_{ji}^4)\cdot E(X_{li}^4)} < \infty.$$

Because $X_{ji}X_{li}$ is i.i.d. with two moments, $\frac{1}{n}\sum_{i=1}^{n}X_{ji}X_{li}$ obeys the law of large numbers, so

$$\frac{1}{n}\sum_{i=1}^{n}X_{ji}X_{li} \xrightarrow{p} E(X_{ji}X_{li}).$$

This is true for all the elements of $n^{-1}\mathbf{X}'\mathbf{X}$, so

$$n^{-1}\mathbf{X}'\mathbf{X} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_i\mathbf{X}_i' \xrightarrow{p} E(\mathbf{X}_i\mathbf{X}_i') = \Sigma_{\mathbf{XX}}.$$

Applying the same reasoning and using Assumption (ii) that $(\mathbf{X}_i, \mathbf{W}_i, Y_i)$ are i.i.d. and Assumption (iii) that $(\mathbf{X}_i, \mathbf{W}_i, u_i)$ have four moments, we have

$$n^{-1}\mathbf{W}'\mathbf{W} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{W}_i\mathbf{W}_i' \xrightarrow{p} E(\mathbf{W}_i\mathbf{W}_i') = \Sigma_{\mathbf{WW}},$$

$$n^{-1}\mathbf{X}'\mathbf{W} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_i\mathbf{W}_i' \xrightarrow{p} E(\mathbf{X}_i\mathbf{W}_i') = \Sigma_{\mathbf{XW}},$$

and

$$n^{-1}\mathbf{W}'\mathbf{X} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{W}_i\mathbf{X}_i' \xrightarrow{p} E(\mathbf{W}_i\mathbf{X}_i') = \Sigma_{\mathbf{WX}}.$$

From Assumption (iii) we know $\Sigma_{\mathbf{XX}}$, $\Sigma_{\mathbf{WW}}$, $\Sigma_{\mathbf{XW}}$, and $\Sigma_{\mathbf{WX}}$ are all finite non-zero, Slutsky's theorem implies

$$\begin{aligned} n^{-1}\mathbf{X}'\mathbf{M_W X} &= n^{-1}\mathbf{X}'\mathbf{X} - (n^{-1}\mathbf{X}'\mathbf{W})(n^{-1}\mathbf{W}'\mathbf{W})^{-1}(n^{-1}\mathbf{W}'\mathbf{X}) \\ &\xrightarrow{p} \boldsymbol{\Sigma}_{\mathbf{XX}} - \boldsymbol{\Sigma}_{\mathbf{XW}}\boldsymbol{\Sigma}_{\mathbf{WW}}^{-1}\boldsymbol{\Sigma}_{\mathbf{WX}} \end{aligned}$$

which is finite and invertible.

(c) The conditional expectation

$$E(\mathbf{U}|\mathbf{X}, \mathbf{W}) = \begin{pmatrix} E(u_1|\mathbf{X}, \mathbf{W}) \\ E(u_2|\mathbf{X}, \mathbf{W}) \\ \vdots \\ E(u_n|\mathbf{X}, \mathbf{W}) \end{pmatrix} = \begin{pmatrix} E(u_1|\mathbf{X}_1, \mathbf{W}_1) \\ E(u_2|\mathbf{X}_2, \mathbf{W}_2) \\ \vdots \\ E(u_n|\mathbf{X}_n, \mathbf{W}_n) \end{pmatrix}$$

$$= \begin{pmatrix} \mathbf{W}_1'\delta \\ \mathbf{W}_2'\delta \\ \vdots \\ \mathbf{W}_n'\delta \end{pmatrix} = \begin{pmatrix} \mathbf{W}_1' \\ \mathbf{W}_2' \\ \vdots \\ \mathbf{W}_n' \end{pmatrix}\delta = \mathbf{W}\delta.$$

The second equality used Assumption (ii) that $(\mathbf{X}_i, \mathbf{W}_i, Y_i)$ are i.i.d., and the third equality applied the conditional mean independence assumption (i).

(d) In the limit

$$n^{-1}\mathbf{X}'\mathbf{M_W}\mathbf{U} \xrightarrow{p} E(\mathbf{X}'\mathbf{M_W}\mathbf{U}|\mathbf{X},\ \mathbf{W}) = \mathbf{X}'\mathbf{M_W}E(\mathbf{U}|\mathbf{X},\ \mathbf{W}) = \mathbf{X}'\mathbf{M_W}\mathbf{W}\delta = \mathbf{0}_{k_1\times 1}$$

because $\mathbf{M_W}\mathbf{W} = \mathbf{0}$.

(e) $n^{-1}\mathbf{X}'\mathbf{M_W}\mathbf{X}$ converges in probability to a finite invertible matrix, and $n^{-1}\mathbf{X}'\mathbf{M_W}\mathbf{U}$ converges in probability to a zero vector. Applying Slutsky's theorem,

$$\hat{\beta} - \beta = (n^{-1}\mathbf{X}'\mathbf{M_W}\mathbf{X})^{-1}(n^{-1}\mathbf{X}'\mathbf{M_W}\mathbf{U}) \xrightarrow{p} \mathbf{0}.$$

This implies

$$\hat{\beta} \xrightarrow{p} \beta.$$

19.11. (a) Using the hint $\boldsymbol{C} = [\boldsymbol{Q}_1\ \boldsymbol{Q}_2]\begin{bmatrix} \mathbf{I}_r & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} \mathbf{Q}_1' \\ \mathbf{Q}_2' \end{bmatrix}$, where $\boldsymbol{Q}'\boldsymbol{Q} = \boldsymbol{I}$. The result follows with $\boldsymbol{A}=\boldsymbol{Q}_1$.

(b) $\boldsymbol{W} = \boldsymbol{A}'\boldsymbol{V} \sim \mathrm{N}(\boldsymbol{A}'\boldsymbol{0}, \boldsymbol{A}'\boldsymbol{I}_n\boldsymbol{A})$ and the result follows immediately.

(c) $\boldsymbol{V}'\boldsymbol{C}\boldsymbol{V} = \boldsymbol{V}'\boldsymbol{A}\boldsymbol{A}'\boldsymbol{V} = (\boldsymbol{A}'\boldsymbol{V})'(\boldsymbol{A}'\boldsymbol{V}) = \boldsymbol{W}'\boldsymbol{W}$ and the result follows from (b).

19.13. (a) This follows from the definition of the Lagrangian.

(b) The first order conditions are

$$(*)\ \boldsymbol{X}'(\boldsymbol{Y}-\boldsymbol{X}\tilde{\boldsymbol{\beta}}) + \boldsymbol{R}'\boldsymbol{\lambda} = 0$$

and

$$(**)\ \boldsymbol{R}\tilde{\boldsymbol{\beta}} - \boldsymbol{r} = 0$$

Solving (*) yields

$$(***)\ \tilde{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}} + (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}'\boldsymbol{\lambda}.$$

Multiplying by $\boldsymbol{R}$ and using (**) yields $\boldsymbol{r} = \boldsymbol{R}\hat{\boldsymbol{\beta}} + \boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}'\boldsymbol{\lambda}$, so that

$$\lambda = -[\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}']^{-1}(\boldsymbol{R}\hat{\boldsymbol{\beta}} - \boldsymbol{r}).$$

Substituting this into (***) yields the result.

(c) Using the result in (b), $\boldsymbol{Y} - \boldsymbol{X}\tilde{\boldsymbol{\beta}} = (\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}) - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}'[\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}']^{-1}(\boldsymbol{R}\hat{\boldsymbol{\beta}} - \boldsymbol{r})$, so that

$$(\boldsymbol{Y} - \boldsymbol{X}\tilde{\boldsymbol{\beta}})'(\boldsymbol{Y} - \boldsymbol{X}\tilde{\boldsymbol{\beta}}) = (\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\beta}})'(\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}) + (\boldsymbol{R}\hat{\boldsymbol{\beta}} - r)'[\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}']^{-1}(\boldsymbol{R}\hat{\boldsymbol{\beta}} - \boldsymbol{r})$$

$$+ 2(\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\beta}})'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}'[\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}']^{-1}(\boldsymbol{R}\hat{\boldsymbol{\beta}} - \boldsymbol{r}).$$

But $(\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\beta}})'\boldsymbol{X} = 0$, so the last term vanishes, and the result follows.

(d) The result in (c) shows that $(\boldsymbol{R}\hat{\boldsymbol{\beta}} - \boldsymbol{r})'[\boldsymbol{R}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}']^{-1}(\boldsymbol{R}\hat{\boldsymbol{\beta}} - \boldsymbol{r}) = SSR_{Restricted} - SSR_{Unrestricted}$. Also $s_u^2 = SSR_{Unrestricted}/(n - k_{Unrestricted} - 1)$, and the result follows immediately.

19.15. (a) This follows from exercise (19.6).

(b) $\tilde{\mathbf{Y}}_i = \tilde{\mathbf{X}}_i\beta + \tilde{\mathbf{u}}_i$, so that

$$\hat{\beta} - \beta = \left(\sum_{i=1}^{n} \tilde{X}_i{}'\tilde{X}_i\right)^{-1} \sum_{i=1}^{n} \tilde{X}_i{}'\tilde{u}_i$$

$$= \left(\sum_{i=1}^{n} \tilde{X}_i{}'\tilde{X}_{ii}\right)^{-1} \sum_{i=1}^{n} X_i{}'M'Mu_i$$

$$= \left(\sum_{i=1}^{n} \tilde{X}_{ii}{}'\tilde{X}_{ii}\right)^{-1} \sum_{i=1}^{n} X_i{}'M'u_i$$

$$= \left(\sum_{i=1}^{n} \tilde{X}_{ii}{}'\tilde{X}_{ii}\right)^{-1} \sum_{i=1}^{n} \tilde{X}_{ii}{}'u_i$$

(c) $\hat{Q}_{\tilde{X}} = \frac{1}{n}\sum_{i=1}^{n}(T^{-1}\sum_{t=1}^{T}(X_{it} - \bar{X}_i)^2)$, where $(T^{-1}\sum_{t=1}^{T}(X_{it} - \bar{X}_i)^2)$ are i.i.d. with mean $\mathbf{Q_X}$ and finite variance (because X_{it} has finite fourth moments). The result then follows from the law of large numbers.

(d) This follows the the Central limit theorem.

(e) This follows from Slutsky's theorem.

(f) η_i^2 are i.i.d., and the result follows from the law of large numbers.

(g) Let $\hat{\eta}_i = T^{-1/2}\tilde{\mathbf{X}}_i{}'\hat{\mathbf{u}}_i = \eta_i - T^{-1/2}(\hat{\beta} - \beta)\tilde{\mathbf{X}}_i{}'\tilde{\mathbf{X}}_i$. Then

$$\hat{\eta}_i^2 = T^{-1/2}\tilde{\mathbf{X}}_i{}'\hat{\mathbf{u}}_i = \eta_i^2 + T^{-1}(\hat{\beta} - \beta)^2(\tilde{\mathbf{X}}_i{}'\tilde{\mathbf{X}}_i)^2 - 2T^{-1/2}(\hat{\beta} - \beta)\eta_i\tilde{\mathbf{X}}_i{}'\tilde{\mathbf{X}}_i$$

and

$$\tfrac{1}{n}\textstyle\sum_{i=1}^{n}\hat{\eta}_i^2 - \frac{1}{n}\sum_{i=1}^{n}\eta_i^2 = T^{-1}(\hat{\beta} - \beta)^2\tfrac{1}{n}\sum_{i=1}^{n}(\tilde{\mathbf{X}}_i{}'\tilde{\mathbf{X}}_i)^2 - 2T^{-1/2}(\hat{\beta} - \beta)\tfrac{1}{n}\sum_{i=1}^{n}\eta_i\tilde{\mathbf{X}}_i{}'\tilde{\mathbf{X}}_i$$

Because $(\hat{\beta} - \beta)\xrightarrow{p}0$, the result follows from (a) $\frac{1}{n}\sum_{i=1}^{n}(\tilde{X}_i{}'\tilde{X}_i)^2 \xrightarrow{p} E[(\tilde{X}_i{}'\tilde{X}_i)^2]$ and (b) $\frac{1}{n}\sum_{i=1}^{n}\eta_i\tilde{X}_i{}'\tilde{X}_i \xrightarrow{p} E(\eta_i\tilde{X}_i{}'\tilde{X}_i)$. Both (a) and (b) follow from the law of large numbers; both (a) and (b) are averages of i.i.d. random variables. Completing the proof requires verifying that $(\tilde{\mathbf{X}}_i{}'\tilde{\mathbf{X}}_i)^2$ has two finite moments and $\eta_i \tilde{\mathbf{X}}_i ' \tilde{\mathbf{X}}_i$ has two finite moments. These in turn follow from 8-moment assumptions for (X_{it}, u_{it}) and the Cauchy-Schwartz inequality. Alternatively, a "strong" law of large numbers can be used to show the result with finite fourth moments.

19.17 The results follow from the hints and matrix multiplication and addition.