Brief Solutions

1. (a) $\hat{\beta}_1 = (X_1'X_1)^{-1}X_1'(y - x_2\hat{\beta}_2)$ and $\hat{\beta}_2 = \frac{x_2'M_{X_1}'y}{x_2'M_{X_1}x_2}$.

   (b) Note that $\underset{\hat{\beta}}{min}(y-X\hat{\beta})'(y-X\hat{\beta})$ and $\underset{\hat{\beta}_1,\hat{\beta}_2}{min}(y-X_1\hat{\beta}_1-X_2\hat{\beta}_2)'(y-X_1\hat{\beta}_1-X_2\hat{\beta}_2)$ yield the same minimum. The identity can be derived therein.

   (c) To prove the adjusted R-squared will rise, we need to prove $\frac{(N-K)y'M_Xy}{N-K-1} > y'M_{X_1}y = y'M_Xy + \frac{(x_2'M_{X_1}y)^2}{x_2'M_{X_1}x_2}$, i.e. $1 > \frac{(x_2'M_{X_1}y)^2}{y'M_Xyx_2'M_{X_1}x_2/(N-K-1)}$, which is the squared t statistic.

2. GMM estimation is to choose the parameter values to minimize a criterion function, which is dervived from some moment conditions. There is no guarantee that GMM estimator is unbiased. All the properties related to the estimator hold only asymptotically. It is possible to use Wald and LM statistic for hypothesis tests, which all follow a chi-squared distribution with the degrees of freedom equal to the number of restrictions. Students can use IV estimation as an example to illustrate the points.

3. (a) $plim\hat{\beta}_1 = \beta_1 + \frac{Cov(\epsilon_i,x_i)}{Var(x_i)} = \beta_1 + \frac{\alpha_2\sigma^2}{\alpha_1^2\gamma^2+\alpha_2^2\sigma^2+\omega^2}$ and $plim\hat{\sigma^2} = Var(\epsilon|x_i) = \sigma^2 - \frac{Cov^2(\epsilon_i,x_i)}{Var(x_i)} = \sigma^2 - \frac{\alpha_2^2\sigma^4}{\alpha_1^2\gamma^2+\alpha_2^2\sigma^2+\omega^2}$.

   (b) By stacking up the observations, we can rewrite [3.3] as

$$y = X\beta + M_Zx\beta_2 + v.$$

   If we regress X on M_Zx, the estimated residual will be $(I-M_Zx(x'M_Zx)^{-1}x'M_Z)X = X-M_Zx[0,1] = P_ZX$. By Frisch-Waugh Theorem, $\tilde{\beta} = (X'P_ZX)^{-1}X'P_Zy$, which is the same as the TSLS estimator.

   (c) Define $w_i = \alpha_2\epsilon_i + u_i$. The SSR can be written as $y'M_Xy - y'M_XM_ZX(X'M_ZM_XM_ZX)^{-1}X'M_ZM_Xy$. Hence $plim\frac{SSR}{N} = \frac{\epsilon'M_X\epsilon}{N} - \frac{\epsilon'M_XM_ZX(X'M_ZM_XM_ZX)^{-1}X'M_ZM_X\epsilon}{N} = Var(\epsilon|x_i) - \frac{E^2[(\epsilon_i-E(\epsilon_i|x_i))w_i]}{Var(w_i|x_i)} = Var(\epsilon_i|x_i,w_i) = \sigma^2 - \frac{\alpha_2^2\sigma^4}{\alpha_1^2\gamma^2+\alpha_2^2\sigma^2+\omega^2} - \frac{\alpha_2^2\sigma^4\left(1+\frac{\alpha_2^2\sigma^2+\omega^2}{\alpha_1^2\gamma^2+\alpha_2^2\sigma^2+\omega^2}\right)^2}{\alpha_2^2\sigma^2+\omega^2-\frac{(\alpha_2^2\sigma^2+\omega^2)^2}{\alpha_1^2\gamma^2+\alpha_2^2\sigma^2+\omega^2}}$

4. (a) The statistic follows a student t distribution with q degrees of freedom.

(b) The statistic follows a student t distribution with $N - K - 1$ degrees of freedom. For the proof, see Theorem 4.3 of Lecture 2. Note that the $(2, 2)$ element in $(X'X)^{-1}$ is the same as the SSR obtained by regressing x_1 on all the other columns of X.

(c) Even if ϵ_i is not normal, it is still possible for $\hat{\beta}_1$ and its t statistic to follow normal distribution asymptotically. Given a large degrees of freedom, there is little differene between t and normal distribution. For the assumptions, please refer to Theorem 8.1 and Theorem 8.2.

5. (a) We require $E(|Z_i|) < \infty$.

(b) See Page 3 in the notes for Lecture 6.

(c) $\hat{\theta}_{MLE}$ will follow a normal distribution asymptotically. See Theorem 3.2 in Lecture 6.