Brief Solutions

1. (a) $\hat{\beta} = (X'M_0X)^{-1}X'M_0y$ and $\hat{\alpha} = \frac{\iota' y}{N} - \frac{\iota' X\hat{\beta}}{N}$, where ι is an $N \times 1$ vector and $M_0 = I_N - \frac{\iota\iota'}{N}$.

   (b) The F statistic can be calculated as $F = \frac{y'(M_0-M_z)y/K}{y'M_zy/(N-K-1)}$, where $M_z = I_N - Z(Z'Z)^{-1}Z'$ and $Z = [\iota, X]$. (Students can also calculate Wald, Lagrange multiplier (LM) or likelihood ration (LR) statistic to answer this question.)

   (c) Under $\beta = 0$, $F = \frac{\epsilon'(M_0-M_z)\epsilon/K}{\epsilon' M_z\epsilon/(N-K-1)}$. Since $\epsilon|X \sim N(0, \sigma^2 I_N)$, we have $\frac{\epsilon' M_z \epsilon}{\sigma^2} \sim \chi^2_{N-K-1}$ and $\epsilon'(M_0 - M_z)\epsilon \sim \chi^2_K$, which are independent since $(M_0 - M_z)M_z = 0$. Therefore F follows an F distribution with degrees of freedom K and $N - K - 1$.

2. Maximum likelihood estimation is to choose the parameter values which maximize the log likelihood function, where the likelihood function is the joint density function of the dependent variable observations. There is no guarantee that maximum likelihood estimator is unbiased, though under certain regularity conditions, maximum likelihood estimator is consistent. It is possible to use Wald, LR and LM statistic for hypothesis tests, which all follow a chi-squared distribution with the degrees of freedom equal to the number of restrictions.

3. (a) $R^2 = \frac{\left[\sum_{i=1}^N (x_i-\bar{x})(y_i-\bar{y})\right]^2}{\sum_{i=1}^N (x_i-\bar{x})^2 \sum_{i=1}^N (y_i-\bar{y})^2}$. It is called R-squared because it is the square of the sample correlation coefficient between the dependent variable and the model fitted value in general. In this case, it is the square of the sample correlation coefficient between the dependent variable and the explanatory variable.

   (b) If we keep including irrelevant regressors, R^2 will increase in the absence of perfect collinearity. The maximum value one can obtain is 1 since the model will have perfect fit when the number of parameters in the model is equal to the number of observations.

   (c) The probability limit should be $\frac{\beta^2\sigma_X^4}{\sigma_X^2(\sigma^2+\beta^2\sigma_X^2)} = \frac{\beta^2\sigma^2}{\sigma^2+\beta^2\sigma_X^2}$.

4. It would be possible due to data unavailability. Students are free to give any practical examples. If the omitted variable is correlated with the included explanatory variable(s), the estimates of the model parameter(s) would be inconsistent. We can find proxy or instrument variables to remedy such

problems. For instrument variable, students should describe the procedure of two stage least squares or GMM estimation.

5. (a) The coefficient for $log(dist)$ is elasticity, i.e. if $dist$ increases by one percent, $price$ will decrease approximately by 0.1343 percentage points. The coefficient for $rooms$ is semi-elasticity, i.e. if there is one more room, house price will increase by about 25.45%.

   (b) The exact percentage change should be $100[\exp(0.2545) - 1] = 28.98\%$.

   (c) Though the OLS estimator $\hat{\beta}$ is unbiased, the exact percentage change estimator is biased since $E\left(\exp(\hat{\beta}) - 1\right) \neq \exp E(\hat{\beta}) - 1 = \exp\beta - 1$. However, it is consistent since $plim\left(\exp(\hat{\beta}) - 1\right) = \exp plim(\hat{\beta}) - 1 = \exp\beta - 1$.

   (d) Given the Jarque-Bera statistic, which follows a chi-squared distribution with two degrees of freedom, we can say that the error term does not follow normal distribution and house prices do not follow log-normal distribution. We can estimate the house price as

$$\begin{aligned} 19175.7 = & \frac{\sum_{i=1}^{506} \exp(\hat{\epsilon}_i)}{506} \exp\Big[11.084 - 0.9535 \cdot log(5.54) - 0.1343 \cdot log(3.21) \\ & + 0.2545 \cdot 6 - 0.05245 \cdot 19 \Big]. \end{aligned}$$

6. (a) We require the additional condition of $Var(Z_i) < \infty$ or $E(Z_i^2) < \infty$.

   (b) Define x_i as the ith row of X, $w_i = (1, x_i)'$, $\gamma = (\alpha, \beta')'$ and ϵ_i as the ith element in ϵ. The conditions are $E(\epsilon_i w_i) = 0$, $E(\epsilon_i \epsilon_j w_i w_j') = 0$, $E(x_i x_j') = E(x_i)E(x_j')$ for $i \neq j$, $E(\epsilon_i^2 x_{ik}^2) < \infty$, $E(x_{ik}^4) < \infty$ for $k = 1, 2, \ldots, K$ and $E(w_i w_i') = M$ is positive definite. Note that $\hat{\gamma}_{OLS} = \gamma + \left(\frac{\sum_{i=1}^N w_i w_i'}{N}\right)^{-1} \frac{\sum_{i=1}^N w_i \epsilon_i}{N}$. Under the conditions, we can use law of large numbers to prove $\frac{\sum_{i=1}^N w_i w_i'}{N} \xrightarrow{p} M$ and $\frac{\sum_{i=1}^N w_i \epsilon_i}{N} \xrightarrow{p} 0$. Note that the k, lth element in $x_i x_i'$ is $x_{ik} x_{il}$. Using Cauchy-Schwarz inequality, we can have $E(|x_{ik} x_{il}|^2) \leq \sqrt{E(|x_{ik}|^4) E(|x_{il}|^4)} \leq \infty$. By continuous mapping theorem, $\hat{\gamma}_{OLS} \xrightarrow{p} \gamma + M^{-1} 0 = \gamma$.