EMET2007/4007/6007 — Econometric Methods
Practice Final Exam — Answer Key

Semester 1, 2026

Question 1

(a) **Sign.** The negative sign for the `male` coefficient estimates indicates that males *tend* to score lower than females, *holding everything else constant*.

Here's an example of a low quality answer:
The negative sign for the `male` coefficient estimates indicates that males score lower than females.

Size. The magnitude is extremely small (about 0.02 points on a 600-point scale), which is economically and practically negligible.

Statistical significance. The t-statistic is -2.942 and thus smaller than the critical value -1.96. We therefore reject $H_0\colon \beta_{\texttt{male}} = 0$ at the 5% significance level.

Another correct answer:
The confidence interval is $[-0.039, -0.088]$ and thus does not include zero. We therefore reject $H_0\colon \beta_{\texttt{male}} = 0$ at the 5% significance level.

Another correct answer:
The p-value is 0.003 and thus smaller than 5%. We therefore reject $H_0\colon \beta_{\texttt{male}} = 0$ at the 5% significance level.

(b) We use the reported p-value to recover the statistic. The p-value for `mediahrs` is 0.050, and the test is two-sided, so the absolute value of the *t*-statistic needs to be 1.96. The sign of the *t*-statistic is the same as that of the coefficient estimate, therefore the *t*-statistic must have been -1.96.

(c) The interaction term allows for the effect of `mediahrs` to differ by gender. The estimated effect of increasing `mediahrs` by one unit, while holding everything else constant, is -2.24 for females and $-2.24 - 1.28$ for males. While overall the effect of additional `mediahrs` tends to be negative, it is especially negative for males. This means that males lose out disproportionately from media usage.

Aside: no need to discuss statistical significance as it wasn't asked.

(d) A candidate for an omitted variable is parental socioeconomic status (SES).

Why it might be correlated with `mediahrs`. Children from lower-SES households tend to spend more time on unsupervised screen activities because parents may have less time for monitoring or fewer resources for alternative activities. Thus $\text{Corr}(\texttt{SES}, \texttt{mediahrs}) < 0$.

Why it might be correlated with `mathscr`. Higher-SES families provide better educational resources, attend better-resourced schools, and benefit from positive peer effects. Thus $\text{Corr}(\texttt{SES}, \texttt{mathscr}) > 0$.

Direction of bias. The omitted-variable bias (OVB) formula states that, when SES is omitted from the regression,

$$\underbrace{\hat{\beta}_{\texttt{mediahrs}}}_{\text{biased estimate}} = \underbrace{\beta_{\texttt{mediahrs}}}_{\text{true effect}} + \text{ovb}$$

$$\text{where ovb} = \underbrace{\beta_{\texttt{SES}}}_{>0} \cdot \underbrace{\frac{\text{Cov}(\texttt{mediahrs}, \text{SES})}{\text{Var}(\texttt{SES})}}_{<0},$$

The bias term is therefore *negative*: the estimated coefficient on `mediahrs` is more negative than the true causal effect. In other words, omitting SES causes us to overstate (in absolute value) how harmful media consumption actually is.

Expected change if SES is included. Including SES would absorb part of the negative association currently attributed to media consumption. We would therefore expect the coefficient on `mediahrs` to increase toward zero (become less negative in absolute value).

Question 2

(a) **True.** The generic omitted-variable bias (OVB) formula is

$$\text{ovb} = \beta_2 \cdot \frac{\text{Cov}(X_1, X_2)}{\text{Var}(X_2)}$$

where X_1 is the included regressor and X_2 is the omitted one.

We see that ovb is zero if either $\beta_2 = 0$ (the omitted regressor doesn't play a role in the model), or if the covariance between X_1 and X_2 is zero (the omitted regressor cannot spill over into the estimation of β_1).

Therefore, omitting a variable does *not necessarily* introduce bias.

(b) **True.** Adding more lags to an autoregression is equivalent to adding more regressors to a linear model. We've learned that the the RSS can only decrease (or, in a degenerate case, remain the same), and since $R^2 = 1 - \frac{\text{RSS}}{\text{TSS}}$, R^2 never decreases when regressors are added.

(c) **False.** The two estimators are both unbiased, but they are *not* equally good: the simple average has strictly smaller variance.

Unbiasedness of both estimators. Since $Y_1, \ldots, Y_{10}$ is a random sample from the same distribution, $E(Y_k) = E(Y_7)$ for all k. Hence

$$E\left(\frac{Y_1+Y_{10}}{2}\right) = \frac{1}{2}E(Y_1) + \frac{1}{2}E(Y_{10}) = E(Y_7),$$

$$E\left(\frac{2}{3}Y_6 + \frac{1}{3}Y_9\right) = \frac{2}{3}E(Y_6) + \frac{1}{3}E(Y_9) = E(Y_7).$$

Variance comparison. Let $\sigma^2 = \mathrm{Var}(Y_7)$. By independence,

$$\mathrm{Var}\left(\tfrac{Y_1+Y_{10}}{2}\right) = \tfrac{1}{4}\sigma^2 + \tfrac{1}{4}\sigma^2 = \tfrac{1}{2}\sigma^2,$$
$$\mathrm{Var}\left(\tfrac{2}{3}Y_6 + \tfrac{1}{3}Y_9\right) = \tfrac{4}{9}\sigma^2 + \tfrac{1}{9}\sigma^2 = \tfrac{5}{9}\sigma^2.$$

Since $\frac{5}{9} > \frac{1}{2}$, the second estimator has larger variance. With equal bias, the estimator with smaller variance has smaller mean squared error (MSE), so the simple average is strictly better.

(d) **False.** The statement conflates the *true parameter* β_1 with the *OLS estimator* $\hat{\beta}_1$.

Recall that the OLS formula gives

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}.$$

For $\hat{\beta}_0 = \bar{Y}$ we need $\hat{\beta}_1 \bar{X} = 0$. Assuming $\bar{X} \neq 0$, this requires $\hat{\beta}_1 = 0$.

Extra intuition (not required):
Knowing that the true slope $\beta_1 = 0$ tells us nothing about the realised estimate $\hat{\beta}_1$. Even when $\beta_1 = 0$, the OLS estimator will in general be non-zero in any finite sample due to sampling variation. The OLS estimate $\hat{\beta}_1$ equals exactly zero only in the knife-edge (probability-zero) event that $\sum(X_i - \bar{X})(Y_i - \bar{Y}) = 0$, which does not follow from $\beta_1 = 0$.

Question 3

The model is $Y_i = \beta_1 X_{1i} + \beta_2 X_{2i} + u_i$ with $\beta_1 \neq 0$. Note that there is **no intercept**.

(a) **Setup.** When $\beta_2 = 0$ the model reduces to

$$Y_i = \beta_1 X_{1i} + u_i.$$

Definition of the OLS estimator. The OLS estimator minimises the sum of squared residuals over all candidate values b_1:

$$\hat{\beta}_1 := \arg\min_{b_1} \sum_{i=1}^{n} (Y_i - b_1 X_{1i})^2.$$

Derivation. Let $S(b_1) = \sum_{i=1}^{n} (Y_i - b_1 X_{1i})^2$. Taking the derivative and setting it to zero (the first-order condition, FOC):

$$\frac{dS}{db_1} = -2 \sum_{i=1}^{n} (Y_i - b_1 X_{1i})\, X_{1i} = 0.$$

Expanding and solving for b_1:

$$\sum_{i=1}^{n} X_{1i} Y_i - b_1 \sum_{i=1}^{n} X_{1i}^2 = 0 \quad \Longrightarrow \quad \hat{\beta}_1 = \frac{\sum_{i=1}^{n} X_{1i} Y_i}{\sum_{i=1}^{n} X_{1i}^2}.$$

(No need to check second order condition.)

Proof of unbiasedness. Substitute the model $Y_i = \beta_1 X_{1i} + u_i$ into the estimator:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n X_{1i}(\beta_1 X_{1i} + u_i)}{\sum_{i=1}^n X_{1i}^2} = \beta_1 \cdot \frac{\sum_{i=1}^n X_{1i}^2}{\sum_{i=1}^n X_{1i}^2} + \frac{\sum_{i=1}^n X_{1i} u_i}{\sum_{i=1}^n X_{1i}^2} = \beta_1 + \frac{\sum_{i=1}^n X_{1i} u_i}{\sum_{i=1}^n X_{1i}^2}.$$

Taking conditional expectations given $\{X_{1i}\}_{i=1}^n$ and using linearity of expectation:

$$E[\hat{\beta}_1 \mid X_{1i}] = \beta_1 + \frac{\sum_{i=1}^n X_{1i} \, E[u_i \mid X_{1i}]}{\sum_{i=1}^n X_{1i}^2}.$$

Assumption needed. If we additionally impose the conditional mean independence (CMI) assumption

$$E[u_i \mid X_{1i}] = 0 \quad \text{for all } i,$$

the numerator vanishes and we obtain $E[\hat{\beta}_1 \mid X_{1i}] = \beta_1$.

Aside (not required):
It's interesting that—in contrast to the unbiasedness derivation in the week 5 lecture notes—here it is NOT sufficient to say $E[u_i \mid X_{1i}] = \mu$. Here we actually do need $\mu = 0$ otherwise we don't get unbiasedness for $\hat{\beta}_1$. The reason for this, of course, is that there isn't a constant included here.

(b) **Setup.** Now $\beta_2 \neq 0$. The OLS estimators $(\hat{\beta}_1, \hat{\beta}_2)$ jointly minimise

$$S(b_1, b_2) = \sum_{i=1}^n (Y_i - b_1 X_{1i} - b_2 X_{2i})^2.$$

First-order condition wrt b_1. Setting the partial derivatives to zero:

$$\frac{\partial S}{\partial b_1} = 0: \quad \sum_{i=1}^n X_{1i}(Y_i - \hat{\beta}_1 X_{1i} - \hat{\beta}_2 X_{2i}) = 0. \tag{1}$$

Rearranging the FOC for b_1. Expanding (1):

$$\sum_{i=1}^n X_{1i} Y_i = \hat{\beta}_1 \sum_{i=1}^n X_{1i}^2 + \hat{\beta}_2 \sum_{i=1}^n X_{1i} X_{2i}.$$

Solving for $\hat{\beta}_1$:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n X_{1i} Y_i - \hat{\beta}_2 \sum_{i=1}^n X_{1i} X_{2i}}{\sum_{i=1}^n X_{1i}^2}.$$

When does this equal $\sum X_{1i}Y_i / \sum X_{1i}^2$? The formula in the question holds if and only if

$$\hat{\beta}_2 \sum_{i=1}^{n} X_{1i}X_{2i} = 0.$$

It follows that the result obtains either if $\hat{\beta}_2 = 0$, or if $\frac{1}{n}\sum_{i=1}^{n} X_{1i}X_{2i} = 0$.

Extra intuition (not required):
The condition $\frac{1}{n}\sum X_{1i}X_{2i} = 0$ is the sample analogue of *orthogonality* between the two regressors. When X_1 and X_2 are orthogonal in the data, including X_2 in the regression has no effect on the estimate of β_1: the estimator takes the same form as if X_2 were absent from the model entirely.