FINM3123 Introduction to Econometrics

Chapter 9 Exercises

Solutions

1. The sample selection in this case is arguably endogenous. Because prospective students may look at campus crime as one factor in deciding where to attend college, colleges with high crime rates have an incentive not to report crime statistics. If this is the case, then the chance of appearing in the sample is negatively related to *u* in the crime equation. (For a given school size, higher *u* means more crime, and therefore a smaller probability that the school reports its crime figures.)

2.

(a) We could use the law of iterated expectations or, because we will need it in part (iii), we can plug $x_2 = \delta_0 + \delta_1 x_1 + r$ into the equation for *y* to get

$$\begin{aligned} y &= \beta_0 + \beta_1 x_1 + \beta_2(\delta_0 + \delta_1 x_1 + r) + u \\ &= (\beta_0 + \beta_2\delta_0) + (\beta_1 + \beta_2\delta_1)x_1 + \beta_2 r + u \end{aligned}$$

Given the assumptions on *u* and *r*, each has a zero mean conditional on x_1. Therefore,

$$\begin{aligned} \mathrm{E}(y \mid x_1) &= (\beta_0 + \beta_2\delta_0) + (\beta_1 + \beta_2\delta_1)x_1 + \mathrm{E}(\beta_2 r + u \mid x_1) \\ &= (\beta_0 + \beta_2\delta_0) + (\beta_1 + \beta_2\delta_1)x_1 + \beta_2\mathrm{E}(r \mid x_1) + \mathrm{E}(u \mid x_1) \\ &= (\beta_0 + \beta_2\delta_0) + (\beta_1 + \beta_2\delta_1)x_1 \end{aligned}$$

Given random sampling we can read the probability limits from regressing *y* on *x* from the conditional expectation. The plim of the slope coefficient is $\beta_1 + \beta_2\delta_1$. Unless $\beta_2 = 0$ or $\delta_1 = 0$ the simple regression estimator is inconsistent for β_1.

(b) Part (i) shows that $\mathrm{E}(y \mid x_1)$ is a linear function of x_1. That means no other functions of x_1 appear in $\mathrm{E}(y \mid x_1)$, including x_1^2. So the plim of the OLS coefficient is zero.

(c) We already did the substitution in the solution to (i). For completeness, note that the zero conditional mean assumption on *u* means that *u* is uncorrelated with *r* conditional on x_1. This is why we can write

$$\text{Var}(v \mid x_1) = \text{Var}(v) = \beta_2^2 \text{Var}(u) + \text{Var}(r)$$

This means we can write

$$\begin{aligned} &y = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_1^2 + v \\ &\text{E}(v \mid x_1) = 0 \\ &\text{Var}(v \mid x_1) = \eta^2 \end{aligned}$$

where $\alpha_2 = 0$. Along with the assumption that x_1 is not constant, and assuming random sampling, the Gauss-Markov assumptions hold. That means that the *t* statistic for testing $\text{H}_0 : \alpha_2 = 0$ has an asymptotic normal distribution. So, whether or not x_2 has been improperly excluded, the test for significance of x_1^2 will almost never reject. (If we perform the test at the 5% level then the test will reject in about 5% of all random samples.)

(d) The suggesting of adding x_1^2 to test for omission of another variable, x_2, presumes that x_1^2 serves as a suitable proxy for x_2. There are a couple of problems with this approach. First, we just saw that if $\text{E}(y \mid x_1)$ is linear in x_1, then x_1^2 makes a terrible proxy for x_2: it has no extra predictive power. Second, if x_1^2 is significant it simply could be that the effect of x_1 on *y* is quadratic, not linear. The bottom line is that tests for functional forms are exactly that; they should not be expected to do anything else.

3.

(i) The mean of *stotal* is .047, its standard deviation is .854, the minimum value is –3.32, and the maximum value is 2.24.

(ii) In the regression *jc* on *stotal*, the slope coefficient is .0111 (se = .0110). Therefore, while the estimated relationship is positive, the *t* statistic is only one: the correlation between *jc* and *stotal* is weak at best. In the regression *univ* on *stotal*, the slope coefficient is 1.170 (se = .029), for a *t* statistic of 39.68. Therefore, *univ* and *stotal* are positively correlated (with correlation = .435).

(iii) When we add *stotal* to the equation and estimate the resulting equation by OLS, we get

$$\log(\widehat{wage}) = \underset{(.021)}{1.495} + \underset{(.0068)}{.0631}\,jc + \underset{(.0026)}{.0686}\,univ + \underset{(.00016)}{.00488}\,exper + \underset{(.0068)}{.0494}\,stotal$$

$n = 6{,}763$, $R^2 = .228$

For testing $\beta_{jc} = \beta_{univ}$, we can use the same trick as in Chapter 4 to get the standard error of the difference: replace *univ* with *totcoll* = *jc* + *univ*, and then the coefficient on *jc* is the difference in the estimated returns, along with its standard error. Let $\theta_1 = \beta_{jc} - \beta_{univ}$. Then $\widehat{\theta_1} = -.0056\ (se = .0069)$. Compared with what we found without *stotal*, the evidence is even weaker against H_1: $\beta_{jc} < \beta_{univ}$. The t statistic from equation without *stotal* is about –1.48, while here we have obtained only −.81.

(iv) When $stotal^2$ is added to the equation, its coefficient is .0019 (t statistic = .40). Therefore, there is no reason to add the quadratic term.

(v) The F statistic for testing exclusion of the interaction terms $stotal \cdot jc$ and $stotal \cdot univ$ is about 1.96; with 2 and 6,756 *df*, this gives *p*-value = .141. So, even at the 10% level, the interaction terms are jointly insignificant. It is probably not worth complicating the basic model estimated in part (iii).

(vi) I would just use the model from part (iii), where *stotal* appears only in level form. The other embellishments were not statistically significant at small enough significance levels to warrant the additional complications.

4. (i) The regression gives $\hat{\beta}_{exec} = .085$ with $t = .30$. The positive coefficient means that there is no deterrent effect, and the coefficient is not statistically different from zero.

(ii) Texas had 34 executions over the period, which is more than three times the next highest state (Virginia with 11). When a dummy variable is added for Texas, its t statistic is −.32, which is not unusually large. (The coefficient is large in magnitude, −8.31, but the studentized residual is not large.) We would not characterize Texas as an outlier.

(iii) When the lagged murder rate is added, $\hat{\beta}_{exec}$ becomes −.071 with $t = -2.34$. The coefficient changes sign and becomes nontrivial: each execution is estimated to reduce the murder rate by .071 (murders per 100,000 people).

(iv) When a Texas dummy is added to the regression from part (iii), its t is only $-.37$ (and the coefficient is only -1.02). So, it is not an outlier here, either. Dropping TX from the regression reduces the magnitude of the coefficient to $-.045$ with $t = -0.60$. Texas accounts for much of the sample variation in *exec*, and dropping it gives a very imprecise estimate of the deterrent effect.