Problem Set II: Answer key

Problem 1

1. OLS estimates

$$\begin{aligned} \widehat{\log(\text{earnings})}_i &= 5.0455 + 0.0667 \cdot \text{educ}_\text{i} \\ &\quad (0.0849) \quad (0.0062) \end{aligned}$$

The coefficient is the return to education: one additional year of schooling increases weekly wage by 6.7%.

2. The OLS estimates

$$\begin{aligned} \widehat{\log(\text{earnings})}_i &= 4.7050 + 0.0443 \cdot \text{educ}_\text{i} + 0.00633 \cdot \text{iq}_\text{i} \\ &\quad (0.1003) \quad (0.0071) \quad (0.0010) \end{aligned}$$

The coefficient on education is now 0.0224 smaller. A linear regression of iq on educ gives

$$\begin{aligned} \widehat{\text{iq}}_i &= 53.6872 + 3.5388 \cdot \text{educ}_\text{i}, \\ &\quad (2.6229) \quad (0.1922) \end{aligned}$$

Now $0.0224 = .00633 \cdot 3.5388$.

Problem 2

1. In the relation

$$y = \beta_0 + \beta_1 x + \beta_2 z + \beta_3 w + \varepsilon$$

we have that β_1 is the true partial effect of x and $\beta_3 = 0$. We also have that in

$$z = \delta_0 + \delta_1 x + \delta_2 w + \zeta$$

$\delta_1 = 0$. Now consider the linear relation

$$y = \gamma_0 + \gamma_1 x + \gamma_2 w + \eta$$

If we estimate γ_1 and γ_2 by OLS we have by the omitted variable (z omitted) that the OLS estimators converge in probability to

$$\hat{\gamma}_1 \xrightarrow{p} \beta_1 + \delta_1 \beta_2 = \beta_1$$

$$\hat{\gamma}_2 \xrightarrow{p} \delta_2 \beta_2 \neq 0$$

2. The asymptotic bias of omitting z is $\beta_z \kappa_x$.

3. The asymptotic bias of omitting z is now $\beta_z \lambda_x$.

4. The bias with the proxy is smaller if and only if $\lambda_x < \kappa_x$.

5. We have

$$\begin{array}{rcl} \widehat{\text{iq}}_i & = & 53.6872 \; + \; 3.5388 \cdot \text{educ}_\text{i}. \\ & & (2.6229) \quad (0.1922) \end{array}$$

so that $\hat{\kappa}_x = 3.5300$. Further

$$\begin{array}{rcl} \widehat{\text{iq}}_i & = & 44.9921 \; + \; 2.8657 \cdot \text{educ}_\text{i} + \; 0.4950 \cdot \text{kww}_\text{i}, \\ & & (2.7229) \quad (0.2009) \qquad\quad (0.0578) \end{array}$$

so that $\hat{\lambda}_x = 2.8657$. We conclude that that it is better to use the proxy. Actually the situation is slightly more complicated because in his example assumption (i) for a perfect proxy also does not hold.