Exam #1 Answer Key

Economics 435: Quantitative Methods

Fall 2008

1 A few warmup questions

a) First note that:

$$\begin{aligned} E(xu) &= E(E(xu|x)) && \text{(by the law of iterated expectations)} \\ &= E(xE(u|x)) && \text{(by the conditioning rule)} \\ &= E(xE(u)) && \text{(since we are given } E(u|x) = E(x)\text{)} \\ &= E(x)E(u) && \text{(by linearity of expectations)} \end{aligned}$$

Also note that

$$\begin{aligned} cov(x,u) &= E\left((x - E(x))(u - E(u))\right) && \text{(by definition)} \\ &= E\left(xu - uE(x) - xE(u) + E(u)E(x)\right) && \text{(by algebra)} \\ &= E(xu - E(u)E(x) - E(x)E(u) + E(u)E(x) && \text{(by the linearity of expectations)} \\ &= E(xu) - E(x)E(u) && \text{(by algebra)} \\ &= E(x)E(u) - E(x)E(u) && \text{(by our earlier result that } E(xu) = E(x)E(u)\text{)} \\ &= 0 \end{aligned}$$

b) The way to prove this is by finding one example where $cov(x,u) = 0$ but $E(u|x) \neq E(u)$. There are a lot of ways of doing this, here's my example:

$$\begin{aligned} x &= \begin{cases} -1 & \text{with probability } 0.25 \\ 0 & \text{with probability } 0.5 \\ 1 & \text{with probability } 0.25 \end{cases} \\ u &= x^2 \end{aligned}$$

Obviously $E(u|x) = x^2$, which depends on x. To find the covariance:

$$\begin{aligned} cov(x,u) &= E(xu) - E(x)E(u) \\ &= [0.25(-1*1) + 0.5(0*0) + 0.25(1*1)] - [0.25*(-1) + 0.5*0 + 0.25*1]\,[0.25(1) + 0.5*0 + 0.25*1] \\ &= 0 - 0*0.5 \\ &= 0 \end{aligned}$$

c) The slope is $\frac{dy}{dx} = \beta_1 e^{\beta_0 + \beta_1 x}$ and the elasticity is $\frac{dy}{dx}\frac{x}{y} = \beta_1 x$.

d) Let A be the event "snow in the air" and let B be the event "2 cm of snow on the ground". The first column in the table gives $\Pr(B)$ and the second gives $\Pr(A \cap B)$. We are looking for $\Pr(A|B) = \Pr(A \cap B)/\Pr(B)$, which can be found by just plugging in the numbers

- Quebec: $\Pr(A|B) = 0.50/1.00 = 0.50 \approx 50\%$.
- Vancouver: $\Pr(A|B) = 0.04/0.11 = 0.3636 \approx 36\%$.

e) We have

$$\Pr(A) = \Pr(A \cap B) + \Pr(A \cap B^c) = \Pr(A \cap B) + \Pr(A|B^c)(1 - \Pr(B))$$

We know everything in this expression but $\Pr(A|B^c)$. We know that $0 \leq \Pr(A|B^c) \leq 1$, so:

$$\Pr(A \cap B) \leq \Pr(A) \leq \Pr(A \cap B) + (1 - \Pr(B))$$

Plugging in the numbers:

- Quebec: $0.50 \leq \Pr(A) \leq 0.50$, or more simply: $\Pr(A) = 0.5$.
- Vancouver: $0.04 \leq \Pr(A) \leq 0.93$.

2 The relationship between least squares prediction and the expected value

a)

$$\begin{aligned} ESPE &\equiv E\left[(x-m)^2\right] \qquad \text{(by definition)} \\ &= E(x^2 - 2mx + m^2) \qquad \text{(by algebra)} \\ &= E(x^2) - 2mE(x) + m^2 \qquad \text{(by linearity of expectations)} \end{aligned} \tag{1}$$

Taking derivatives:

$$\frac{\partial ESPE}{\partial m} = -2E(x) + 2m$$

So ESPE is minimized where $-2E(x) + 2m = 0$. Solving for m we get $m = E(X)$

b)

$$\begin{aligned} \frac{\partial ASPE}{\partial \hat{m}} &\equiv \frac{1}{n}\sum_{i=1}^{n}\frac{\partial(x_i - \hat{m})^2}{\partial \hat{m}} \qquad \text{(by differentiation rules)} \\ &= \frac{1}{n}\sum_{i=1}^{n} -2(x_i - \hat{m}) \qquad \text{(by differentiation rules)} \\ &= \frac{-2}{n}\sum_{i=1}^{n}(x_i - \hat{m}) \qquad \text{(by summation rules)} \\ &= \frac{-2}{n}\left(\left(\sum_{i=1}^{n} x_i\right) - \left(\sum_{i=1}^{n}\hat{m}\right)\right) \qquad \text{(by summation rules)} \end{aligned} \tag{2}$$

This quantity is zero if $(\sum_{i=1}^n x_i) - (\sum_{i=1}^n \hat{m}) = 0$, or equivalently:

$$\begin{aligned}
\sum_{i=1}^n \hat{m} &= \sum_{i=1}^n x_i \\
\frac{1}{n}\sum_{i=1}^n \hat{m} &= \frac{1}{n}\sum_{i=1}^n x_i \\
\frac{1}{n} n\hat{m} &= \frac{1}{n}\sum_{i=1}^n x_i \\
\hat{m} &= \frac{1}{n}\sum_{i=1}^n x_i
\end{aligned}$$

3 The education production function

a) This is a standard omitted variables problem:

$$\begin{aligned}
\text{plim } \hat{\beta}_1^A &= \frac{cov(c,q)}{var(q)} \\
&= \frac{cov(\beta_0 + \beta_1 q + \beta_2 s + u, q)}{var(q)} \\
&= \beta_1 + \beta_2 \frac{cov(s,q)}{var(q)}
\end{aligned}$$

b) The bias is $\beta_2 \frac{cov(s,q)}{var(q)}$. I would guess that students with high ability are likely to have both high initial achievement and high current achievement ($\beta_2 > 0$). I would also guess that students with high initial achievement are likely to be in higher quality schools ($cov(s,q) > 0$). This implies that the bias is positive, i.e., $\hat{\beta}_1$ overstates the true β_1.

c) We have:

$$\begin{aligned}
\text{plim } \hat{\beta}_1^B &= \frac{cov(g,q)}{var(q)} \\
&= \frac{cov(\beta_0 + \beta_1 q + (\beta_2 - 1) s + u, q)}{var(q)} \\
&= \beta_1 + (\beta_2 - 1) \frac{cov(s,q)}{var(q)}
\end{aligned}$$

d) Yes.

e) When $\beta_2 < 1$, then $\beta_2 - 1 < 0$. Since we earlier assumed that $cov(s,q) > 0$, this implies that the bias is negative.

f) First we note that:

$$\begin{aligned}
E(c|q,\tilde{s}) &= \beta_0 + \beta_1 q + \beta_2 E(s|q,\tilde{s}) \\
&= \beta_0 + \beta_1 q + \beta_2 (a_0 + a_1 q + a_2 \tilde{s}) \\
&= (\beta_0 + \beta_2 a_0) + (\beta_1 + \beta_2 a_1) q + \beta_2 a_2 \tilde{s})
\end{aligned}$$

This implies that:

$$\begin{aligned} \text{plim } \hat{\beta}_1^C &= \beta_1 + \beta_2 a_1 \\ &= \beta_1 + \beta_2 var(\epsilon) cov(q,s) var(s) var(q) \left(1 - corr(q,s)^2\right) \end{aligned} \tag{3}$$

g) We already assumed that $\beta_2 > 0$, and were told that $a_1 > 0$, so the bias is positive.

h)

$$\begin{aligned} \text{plim } \hat{\beta}_1 &= \frac{cov(\tilde{g}, q)}{var(q)} \\ &= \frac{cov(\beta_0 + \beta_1 q + (\beta_2 - 1)s + u + \epsilon, q)}{var(q)} \\ &= \beta_1 + (\beta_2 - 1)\frac{cov(s,q)}{var(q)} \end{aligned}$$

i) The gain score approach has lower asymptotic bias (in absolute value) whenever

$$(1 - \beta_2)\frac{cov(s,q)}{var(q)} < \beta_2 var(\epsilon) cov(q,s) var(s) var(q) \left(1 - corr(q,s)^2\right)$$

In other words, when measurement error $var(\epsilon)$ is relatively large, and the amount of decay $(1 - \beta_2)$ is relatively small.

j) Since plim $\hat{\beta}_1^C > \beta_1$ and plim $\hat{\beta}_1^D < \beta_1$, then the interval $[\hat{\beta}_1^D, \hat{\beta}_1^C]$ will contain the true value of β_1 with probability approaching one as n approaches infinity.

k)

1. This is an example of random selection. OLS regression will consistently estimate β_1.

2. This is an example of selection on observables, or exogenous selection. OLS regression will consistently estimate β_1.

3. This is an example of selection on unobservables, or endogenous selection. OLS regression will not consistently estimate β_1.

4. This is an example of selection on an omitted variable. OLS regression will not consistently estimate β_1.