Problem Set #5 Answer Key

Economics 435: Quantitative Methods

Fall 2011

1 Consequences of leaving out a quadratic term

a) Let $\epsilon = y - E(y|x)$. Applying the law of large numbers and Slutsky's theorem, we have:

$$\begin{aligned}
\gamma_1 &= \text{plim } \frac{c\hat{o}v(x,y)}{v\hat{a}r(x)} \\
&= \frac{cov(x,y)}{var(x)} \\
&= \frac{cov(x, \beta_0 + \beta_1 x + \beta_2 x^2 + \epsilon)}{var(x)} \\
&= \beta_1 + \beta_2 \frac{cov(x,x^2)}{var(x)} + \frac{cov(x,\epsilon)}{var(x)}
\end{aligned}$$

By the definition of ϵ, we know that $cov(x,\epsilon) = 0$. We also know that $var(x) = E(x^2) - E(x)^2$. All that's left is finding $cov(x, x^2)$:

$$\begin{aligned}
cov(x,x^2) &= E\left[(x - E(x))\left(x^2 - E(x^2)\right)\right] \\
&= E(x^3) - E(x^2)E(x)
\end{aligned}$$

Substituting in we get:

$$\text{plim}\hat{\gamma}_1 = \beta_1 + \beta_2 \frac{E(x^3) - E(x^2)E(x)}{E(x^2) - E(x)^2}$$

b) The marginal effect is given by:

$$\begin{aligned}
\delta(X) &= \frac{\partial E(y|x = X)}{\partial X} \\
&= \frac{\partial(\beta_0 + \beta_1 X + \beta_2 X^2)}{\partial X} \\
&= \beta_1 + 2\beta_2 X
\end{aligned}$$

c) We simply substitute our earlier results into the equation $\delta(X^*) = \gamma_1$, and solve:

$$\begin{aligned}
\delta(X^*) &= \gamma_1 \\
\beta_1 + 2\beta_2 X^* &= \beta_1 + \beta_2 \frac{E(x^3) - E(x^2)E(x)}{E(x^2) - E(x)^2} \\
X^* &= \frac{E(x^3) - E(x^2)E(x)}{2(E(x^2) - E(x)^2)}
\end{aligned}$$

2 Best linear predictors

a)

$$\begin{aligned} E\left((y-b_0-b_1x)^2\right) &= E(y^2-b_0y-b_1xy-b_0y+b_0^2+b_0b_1x-b_1xy+b_0b_1x+b_1^2x^2) \\ &= E(y^2)-2b_0E(y)-2b_1E(xy)+b_0^2+2b_0b_1E(x)+b_1^2E(x^2) \end{aligned}$$

Taking derivatives:

$$\begin{aligned} \partial E/\partial b_0 &= -2E(y)+2b_0+2b_1E(x)=0 \\ \partial E/\partial b_1 &= -2E(xy)+2b_0E(x)+2b_1E(x^2) \end{aligned}$$

Solving, we get:

$$\begin{aligned} b_0 &= E(y)-b_1E(x) \\ b_1 &= \frac{E(xy)-b_0E(x)}{E(x^2)} = \frac{E(xy)-E(x)E(y)}{E(x^2)-E(x)^2} = cov(x,y)/var(x) \end{aligned}$$

b) They are consistent estimators.

$$\begin{aligned} \text{plim } \hat{\beta}_1 &= \text{plim } \frac{c\hat{o}v(x,y)}{v\hat{a}r(x)} \\ &= \frac{\text{plim } c\hat{o}v(x,y)}{\text{plim } v\hat{a}r(x)} \\ &= \frac{cov(x,y)}{var(x)} \\ &= b_1 \\ \text{plim } \hat{\beta}_0 &= \text{plim } (\bar{y}-\hat{\beta}_1\bar{x}) \\ &= \text{plim } \bar{y} - \text{plim } \hat{\beta}_1 \text{plim } \bar{x} \\ &= E(y)-b_1E(x) \\ &= b_0 \end{aligned}$$

c) Based on the first order conditions:

$$\begin{aligned} a_1 &= \frac{cov(E(y|x),x)}{var(x)} \\ a_0 &= E(E(y|x))-a_1E(x) \end{aligned}$$

Now let $v=y-E(y|x)$. Note that $E(v|x)=0$, which implies $cov(v,x)=0$. Then:

$$\begin{aligned} cov(y,x) &= cov(E(y|x)+v,x) \\ &= cov(E(y|x),x)+cov(v,x) \\ &= cov(E(y|x),x) \end{aligned}$$

So this implies $a_1 = \frac{cov(E(y|x),x)}{var(x)} = \frac{cov(y,x)}{var(x)} = b_1$. Substituting this into our expression for a_0, we get

$$\begin{aligned} a_0 &= E(E(y|x))-b_1E(x) \\ &= E(y)-b_1E(x) \\ &= b_0 \end{aligned}$$

Since we have proved that $a_0=b_0$ and $a_1=b_1$, we have proved that the best linear predictor is also the best linear approximation to the true CEF.

3 Fitted values and residuals

a) No it isn't.

$$\begin{aligned}
\text{plim } \hat{y}_i &= \text{plim } (\hat{\beta}_0 + \hat{\beta}_1 x_i) \\
&= \text{plim } \hat{\beta}_0 + \text{plim } \hat{\beta}_1 \text{plim } x_i \\
&= \beta_0 + \beta_1 x_i \\
&= y_i - u_i \neq y_i
\end{aligned}$$

b) Yes it is:

$$\begin{aligned}
\text{plim } \hat{u}_i &= \text{plim } (y_i - \hat{y}_i) \\
&= \text{plim } y_i - \text{plim } \hat{y}_i \\
&= y_i - (y_i - u_i) \\
&= u_i
\end{aligned}$$

c) Yes. A proof was not required, just some sort of explanation, but here's a proof. Let $z_i = u_i^2$ and $w_i = w_i^2$:

$$\begin{aligned}
\hat{\alpha}_1 &= \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(w_i - \bar{w})}{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2} \\
&= \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})w_i - \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})\bar{w}}{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2} \\
&= \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})w_i}{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2} \\
&= \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})z_i + \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(w_i - z_i)}{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2} \\
&= \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(z_i - \bar{z}) + \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(w_i - z_i)}{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}
\end{aligned}$$

Taking probability limits

$$\begin{aligned}
\hat{\alpha}_1 &= \text{plim } \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(z_i - \bar{z}) + \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(w_i - z_i)}{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2} \\
&= \frac{\text{plim } \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(z_i - \bar{z}) + \text{plim } \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(w_i - z_i)}{\text{plim } \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2} \\
&= \frac{\text{plim } \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(z_i - \bar{z}) + \frac{1}{n}\sum_{i=1}^{n} \text{plim } (x_i - \bar{x})\text{plim } (w_i - z_i)}{\text{plim } \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2} \\
&= \frac{\text{plim } \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(z_i - \bar{z}) + \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})0}{\text{plim } \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2} \\
&= \frac{cov(x, z)}{var(x)} \\
&= \alpha_1
\end{aligned}$$

Your answer does not need to be this elaborate!