# Problem Set #6 Answer Key

Economics 835: Econometrics

Fall 2012

## 1 A preliminary result

$$
\begin{aligned}
\hat{cov}(x,y) &= \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y}) \\
&= \frac{1}{n}\sum_{i=1}^{n} x_i y_i - \bar{x}y_i - x_i\bar{y} + \bar{x}\bar{y} \\
&= \frac{1}{n}\sum_{i=1}^{n} x_i y_i - \frac{1}{n}\sum_{i=1}^{n}\bar{x}y_i - \frac{1}{n}\sum_{i=1}^{n}\bar{y}x_i + \frac{1}{n}\sum_{i=1}^{n}\bar{x}\bar{y} \\
&= \frac{1}{n}\sum_{i=1}^{n} x_i y_i - \bar{x}\frac{1}{n}\sum_{i=1}^{n} y_i - \bar{y}\frac{1}{n}\sum_{i=1}^{n} x_i + \bar{x}\bar{y} \\
&= \frac{1}{n}\sum_{i=1}^{n} x_i y_i - \bar{x}\bar{y} - \bar{y}\bar{x} + \bar{x}\bar{y} \\
&= \frac{1}{n}\sum_{i=1}^{n} x_i y_i - \bar{x}\bar{y} \\
\text{plim } \hat{cov}(x,y) &= \text{plim } \left(\frac{1}{n}\sum_{i=1}^{n} x_i y_i - \bar{x}\bar{y}\right) \\
&= \text{plim } \frac{1}{n}\sum_{i=1}^{n} x_i y_i - \text{plim } \bar{x}\text{plim } \bar{y} \qquad \text{by Slutsky's theorem} \\
&= E(xy) - E(x)E(y) \qquad \text{by the LLN} \\
&= cov(x,y)
\end{aligned}
$$

# 2 OLS with a single explanatory variable

**a**) We can show this by:

$$\begin{aligned}
\frac{cov(x,y)}{var(x)} &= \frac{cov(x,\beta_0+\beta_1 x+u)}{var(x)} \quad \text{by substitution} \\
&= \frac{\beta_1 var(x)+cov(x,u)}{var(x)} \quad \text{covariance of sums} \\
&= \frac{\beta_1 var(x)}{var(x)} \quad \text{since } cov(x,u)=0 \\
&= \beta_1 \\
E(y)-\beta_1 E(x) &= E(\beta_0+\beta_1 x+u)-\beta_1 E(x) \quad \text{by substitution} \\
&= \beta_0+\beta_1 E(x)+E(u)-\beta_1 E(x) \quad \text{by linearity of expectations} \\
&= \beta_0
\end{aligned}$$

**b**) First note that:

$$\begin{aligned}
\mathbf{X}'\mathbf{y} &= \begin{bmatrix} \sum_{i=1}^n y_i \\ \sum_{i=1}^n x_i y_i \end{bmatrix} \\
\mathbf{X}'\mathbf{X} &= \begin{bmatrix} n & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & \sum_{i=1}^n x_i^2 \end{bmatrix} \\
(\mathbf{X}'\mathbf{X})^{-1} &= \frac{1}{n\sum_{i=1}^n x_i^2-\left(\sum_{i=1}^n x_i\right)^2}\begin{bmatrix} \sum_{i=1}^n x_i^2 & -\sum_{i=1}^n x_i \\ -\sum_{i=1}^n x_i & n \end{bmatrix}
\end{aligned}$$

So:

$$\begin{aligned}
\hat\beta = \begin{bmatrix} \hat\beta_0 \\ \hat\beta_1 \end{bmatrix} &= \frac{1}{n\sum_{i=1}^n x_i^2-\left(\sum_{i=1}^n x_i\right)^2}\begin{bmatrix} \sum_{i=1}^n x_i^2 & -\sum_{i=1}^n x_i \\ -\sum_{i=1}^n x_i & n \end{bmatrix}\begin{bmatrix} \sum_{i=1}^n y_i \\ \sum_{i=1}^n x_i y_i \end{bmatrix} \\
&= \frac{1}{n\sum_{i=1}^n x_i^2-\left(\sum_{i=1}^n x_i\right)^2}\begin{bmatrix} \sum_{i=1}^n x_i^2\sum_{i=1}^n y_i-\sum_{i=1}^n x_i\sum_{i=1}^n x_i y_i \\ -\sum_{i=1}^n x_i\sum_{i=1}^n y_i+n\sum_{i=1}^n x_i y_i \end{bmatrix} \\
&= \frac{1}{n^2\left(\frac{1}{n}\sum_{i=1}^n x_i^2-\left(\frac{1}{n}\sum_{i=1}^n x_i\right)^2\right)}\begin{bmatrix} n^2\left(\frac{1}{n}\sum_{i=1}^n x_i^2\frac{1}{n}\sum_{i=1}^n y_i-\frac{1}{n}\sum_{i=1}^n x_i\frac{1}{n}\sum_{i=1}^n x_i y_i\right) \\ n^2\left(\frac{1}{n}\sum_{i=1}^n x_i y_i-\frac{1}{n}\sum_{i=1}^n x_i\frac{1}{n}\sum_{i=1}^n y_i\right) \end{bmatrix} \\
&= \frac{1}{\left(\frac{1}{n}\sum_{i=1}^n x_i^2-\left(\frac{1}{n}\sum_{i=1}^n x_i\right)^2\right)}\begin{bmatrix} \left(\frac{1}{n}\sum_{i=1}^n x_i^2\frac{1}{n}\sum_{i=1}^n y_i-\frac{1}{n}\sum_{i=1}^n x_i\frac{1}{n}\sum_{i=1}^n x_i y_i\right) \\ \left(\frac{1}{n}\sum_{i=1}^n x_i y_i-\frac{1}{n}\sum_{i=1}^n x_i\frac{1}{n}\sum_{i=1}^n y_i\right) \end{bmatrix} \\
&= \frac{1}{\hat{var}(x)}\begin{bmatrix} \left(\frac{1}{n}\sum_{i=1}^n x_i^2\frac{1}{n}\sum_{i=1}^n y_i-\frac{1}{n}\sum_{i=1}^n x_i\frac{1}{n}\sum_{i=1}^n x_i y_i+\frac{1}{n}\sum_{i=1}^n x_i\frac{1}{n}\sum_{i=1}^n y_i-\frac{1}{n}\sum_{i=1}^n x_i\frac{1}{n}\sum_{i=1}^n y_i\right) \\ \hat{cov}(x,y) \end{bmatrix} \\
&= \frac{1}{\hat{var}(x)}\begin{bmatrix} \bar y\,\hat{var}(x)-\bar x\,\hat{cov}(x,y) \\ \hat{cov}(x,y) \end{bmatrix} \\
&= \begin{bmatrix} \bar y-\hat\beta_1\bar x \\ \frac{\hat{cov}(x,y)}{\hat{var}(x)} \end{bmatrix}
\end{aligned}$$

**c**)

$$\begin{aligned}
\text{plim } \hat{\beta}_1 &= \text{plim } \frac{\hat{cov}(x,y)}{\hat{var}(x)} && \text{by substitution} \\
&= \frac{\text{plim } \hat{cov}(x,y)}{\text{plim } \hat{var}(x)} && \text{by Slutsky's theorem} \\
&= \frac{cov(x,y)}{var(x)} && \text{previously shown} \\
&= \beta_1 && \text{previously shown} \\
\text{plim } \hat{\beta}_0 &= \text{plim } \left(\bar{y} - \hat{\beta}_1 \bar{x}\right) && \text{by substitution} \\
&= \text{plim } \bar{y} - \text{plim } \hat{\beta}_1 \text{plim } \bar{x} && \text{by Slutsky's theorem} \\
&= E(y) - \text{plim } \hat{\beta}_1 E(x) && \text{by LLN} \\
&= E(y) - \beta_1 E(x) && \text{previously shown} \\
&= \beta_0 && \text{previously shown}
\end{aligned}$$

# 3 Measurement error

**a**)

$$\begin{aligned}
\text{plim } \hat{\beta}_1 &= \frac{cov(\tilde{y},\tilde{x})}{var(\tilde{x})} \\
&= \frac{cov(\beta_0 + \beta_1 x + u + v, x + w)}{var(x+w)} \\
&= \frac{\beta_1 var(x) + \beta_1 cov(x,w) + cov(u,x) + cov(u,w) + cov(v,x) + cov(v,w)}{var(x) + var(w) + 2cov(x,w)} \\
&= \frac{\beta_1 var(x)}{var(x) + var(w)} \\
&= \beta_1 \frac{1}{1+\epsilon_x}
\end{aligned}$$

**b**) The sign is the same (since $\frac{1}{1+\epsilon_x} > 0$) but the magnitude is smaller since $\frac{1}{1+\epsilon_x} < 1$. This kind of bias is called *attenuation bias*.

**c**) No effect at all.

# 4 Omitted variables

**a**)

$$\begin{aligned}
\text{plim } \hat{\beta}_1 &= \text{plim } \frac{c\hat{o}v(x_1, y)}{v\hat{a}r(x_1)} \\
&= \frac{cov(x_1, y)}{var(x_1)} \\
&= \frac{cov(x_1, \beta_0 + \beta_1 x_1 + \beta_2 x_2 + u)}{var(x_1)} \\
&= \frac{\beta_1 var(x_1) + \beta_2 cov(x_1, x_2)}{var(x_1)} \\
&= \beta_1 + \beta_2 \frac{cov(x_1, x_2)}{var(x_1)}
\end{aligned}$$

**b**) My guess is that $\beta_2$ is positive (higher ability leads to higher earnings) and that $cov(x_1, x_2)$ is also positive (higher ability leads to higher educational attainment). If my guess is correct, then $\hat{\beta}_1$ will be biased upwards, i.e., it will overstate the earnings benefit from education.

# 5 Choice of units: The simple version

**a**)

$$\begin{aligned}
\tilde{\beta}_1 &= \frac{c\hat{o}v(\tilde{x}, \tilde{y})}{v\hat{a}r(\tilde{x})} && \text{by substitution} \\
&= \frac{c\hat{o}v(ax + b, cy + d)}{v\hat{a}r(ax + b)} && \text{by substitution} \\
&= \frac{ac\, c\hat{o}v(x, y)}{a^2\, v\hat{a}r(x)} && \text{see below} \\
&= \frac{c}{a} \frac{c\hat{o}v(x, y)}{v\hat{a}r(x)} \\
&= \frac{c}{a} \hat{\beta}_1 && \text{by substitution}
\end{aligned}$$

The third step makes use of the fact that $c\hat{o}v(ax + b, cy + d)$ obeys the same rules as $cov(ax + b, cy + d)$ - remember that analog estimators (like the sample covariance) can be written as population moments for the empirical distribution, so they obey all the same rules that population moments do. If you didn't know this (or think to apply it here), you can still establish the result, it just takes a few more steps of algebra.

# 6 Choice of units: The complicated version

**a**)

$$\begin{aligned}
\tilde{\beta} &= (\tilde{\mathbf{X}}'\tilde{\mathbf{X}})^{-1}\tilde{\mathbf{X}}'\tilde{\mathbf{y}} \\
&= (A'\mathbf{X}'\mathbf{X}A)^{-1}A'\mathbf{X}'c\mathbf{y} \\
&= A^{-1}(\mathbf{X}'\mathbf{X})^{-1}(A')^{-1}A'\mathbf{X}'c\mathbf{y} \\
&= cA^{-1}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \\
&= cA^{-1}\hat{\beta}
\end{aligned}$$

**b)** Since $\tilde{\beta} = cA^{-1}\hat{\beta}$,

$$\begin{aligned} cov(\tilde{\beta}) &= (cA^{-1})'\Sigma cA^{-1} \\ &= c^2 A^{-1'}\Sigma A^{-1} \end{aligned}$$

**c)** The coefficient on that variable is multiplied by $\frac{1}{10}$ and the other coefficients will not be affected. To see this, suppose that there are four explanatory variables and that we have multiplied the second variable by 10. In this case, $c = 1$ (since we haven't changed $y$) and:

$$\begin{aligned} A &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 10 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \\ cA^{-1} = A^{-1} &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0.1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \end{aligned}$$

**d)** In this case $A$ is the identity matrix and $c = 10$. So $\tilde{\beta} = 10\hat{\beta}$. All coefficients are multiplied by 10.

**e)** Let $\imath_n$ be an $n \times 1$ matrix of ones.

The FWL theorem tells us that $\hat{\beta}_1$ will be equal to the coefficient from a linear regression of (the residual from a linear regression of $y_i$ on 1) on (the residual from a linear regression of $\mathbf{x}_i$ on 1). The residual from a linear regression of $y_i$ on 1 is:

$$\begin{aligned} y_i - 1 * (\imath_n'\imath_n)^{-1}(\imath_n'\mathbf{y}) &= y_i - (n)^{-1}\sum_{i=1}^{n} y_i \\ &= y_i - \frac{1}{n}\sum_{i=1}^{n} y_i \\ &= y_i - \bar{y} \end{aligned} \tag{1}$$

By the same reasoning the residual from a linear regression of $\mathbf{x}_i$ on 1 is just: $\mathbf{x}_i - \bar{\mathbf{x}}$. So $\hat{\beta}_1$ is equal to the coefficient from a linear regression of $y_i - \bar{y}$ on $\mathbf{x}_i - \bar{\mathbf{x}}$.

The FWL theorem also tells us that $\tilde{\beta}_1$ will be equal to the coefficient from a linear regression of (the residual from a linear regression of $\tilde{y}_i$ on 1) on (the residual from a linear regression of $\tilde{\mathbf{x}}_i$ on 1). The residual from a linear regression of $\tilde{y}_i$ on 1 is:

$$\begin{aligned} \tilde{y}_i - 1 * (\imath_n'\imath_n)^{-1}(\imath_n'\tilde{\mathbf{y}}) &= \tilde{y}_i - (n)^{-1}\sum_{i=1}^{n} \tilde{y}_i \\ &= y_i + d - (n)^{-1}\sum_{i=1}^{n}(y_i + d) \\ &= y_i - (n)^{-1}\sum_{i=1}^{n} y_i \\ &= y_i - \frac{1}{n}\sum_{i=1}^{n} y_i \\ &= y_i - \bar{y} \end{aligned}$$

By the same reasoning the residual from a linear regression of $\tilde{\mathbf{x}}_i$ on 1 is just $\mathbf{x}_i - \bar{\mathbf{x}}$. So $\tilde{\beta}_1$ is equal to the coefficient from a linear regression of $y_i - \bar{y}$ on $\mathbf{x}_i - \bar{\mathbf{x}}$, and thus equal to $\hat{\beta}_1$.

**f**) The covariance matrix of $\tilde{\beta}_1$ is $\Sigma_1$.

**g**) Nothing.

# 7 An application

**a**) The null is:

$$H_0 : \beta_1 = 0$$

The alternative is:

$$H_1 : \beta_1 \neq 0$$

The test statistic is:

$$t = \frac{\hat{\beta}_1 - 0}{\hat{\sigma}_1} = \frac{-0.00013509}{0.00203867} = -0.06626379$$

Under the null, this test statistic has the $t$ distribution with 45 degrees of freedom. To get a 5% test, the critical values should be the 2.5 and 97.5 percentiles of this distribution:

$$\begin{aligned} c_L &= -2.014103 \\ c_H &= 2.014103 \end{aligned}$$

Since our test statistic lies between the critical values we do not reject the null.

**b**) Let $\hat{\beta}_i$ be the $i$th coefficient estimate listed in the table (e.g., $\hat{\beta}_1$ is the coefficient on "% urban", etc.) let $\beta_i$ be its associated probability limit, and let $\hat{\beta}$ and $\beta$ be the vector of all of the coefficients. Then our null hypothesis can be written:

$$H_0 : R\beta - r = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_5 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

and the alternative is:

$$H_1 : R\beta - r \neq 0$$

The F statistic is:

$$\begin{aligned} F &= \frac{(R\hat{\beta} - r)'[R\hat{\sigma}^2(\mathbf{X}'\mathbf{X})^{-1}R']^{-1}(R\hat{\beta} - r)}{q} \\ &= \frac{\begin{bmatrix} .048356 \\ .14154 \\ .340245 \end{bmatrix}' \begin{bmatrix} .00737400 & .00431739 & .00406416 \\ .00431739 & .00674424 & .00401494 \\ .00406416 & .00401494 & .00710384 \end{bmatrix}^{-1} \begin{bmatrix} .048356 \\ .14154 \\ .340245 \end{bmatrix}}{3} \\ &= 6.85 \end{aligned}$$

Under the null, this statistic has the $F_{3,45}$ distribution. The 95th percentile of this distribution is $c_H \approx 2.81$, so we reject the null.

**c**) The null is:

$$H_0 : \beta_1 = 0$$

The alternative is:

$$H_1 : \beta_1 \neq 0$$

The test statistic is:

$$t = \frac{\hat{\beta}_1 - 0}{\hat{\sigma}_1} = \frac{-0.00013509}{0.00203867} = -0.06626379$$

Under the null, this test statistic has the $N(0,1)$ distribution. To get a 5% test, the critical values should be the 2.5 and 97.5 percentiles of this distribution:

$$\begin{aligned} c_L &= -1.959964 \\ c_H &= 1.959964 \end{aligned}$$

Since our test statistic lies between the critical values we do not reject the null.

Note that the only difference between this test and the finite sample test that assumes normality is that the critical values are a little lower.

**d**) Our null hypothesis can be written:

$$H_0 : g(\beta) = \begin{bmatrix} \beta_2 \\ \beta_3 \\ \beta_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

and the alternative is:

$$H_1 : g(\beta) \neq 0$$

The Jacobian is:

$$G(\beta) = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

so the Wald statistic is:

$$\begin{aligned} W &= \sqrt{n} g(\hat{\beta})' \left[ G(\hat{\beta}) \hat{\sigma}^2 \left( \frac{\mathbf{X}'\mathbf{X}}{n} \right)^{-1} G(\hat{\beta})' \right]^{-1} \sqrt{n} g(\hat{\beta}) \\ &= g(\hat{\beta})' \left[ G(\hat{\beta}) \hat{\sigma}^2 (\mathbf{X}'\mathbf{X})^{-1} G(\hat{\beta})' \right]^{-1} g(\hat{\beta}) \\ &= \begin{bmatrix} .048356 \\ .14154 \\ .340245 \end{bmatrix}' \begin{bmatrix} .00737400 & .00431739 & .00406416 \\ .00431739 & .00674424 & .00401494 \\ .00406416 & .00401494 & .00710384 \end{bmatrix}^{-1} \begin{bmatrix} .048356 \\ .14154 \\ .340245 \end{bmatrix} \\ &= 20.55223 \end{aligned}$$

Under the null, this statistic has the $\chi^2$ distribution with 3 degrees of freedom. The 95th percentile of this distribution is $c_H \approx 7.81$, so we reject the null.

**e**)

1. The coefficient on the urbanization rate, its standard error, and its t-statistic would be unchanged.
2. The coefficient on North Central and its standard error would both be multiplied by 0.01. The t-statistic would be unchanged.