Suggested Answers Problem Set 1 ECOE 60303

Bill Evans Spring 2014

1. Using partitioned inverses,

 $\hat{\beta}_1 = [X_1'(I - X_2(X_2'X_2)^{-1}X_2')X_1]^{-1}X_1'(I - X_2(X_2'X_2)^{-1}X_2')Y.$

Consider a regression of X₂ on X₁ where $X_1 = X_2 \gamma + \nu$ and $\hat{\gamma} = (X_2 X_2)^{-1} X_2 X_1$. If $\hat{\gamma} = 0$, the X₂ on X₁ are uncorrelated. Work with the term in brackets in the portioned inverse, $X_1 (I - X_2 (X_2 X_2)^{-1} X_2) X_1$ can be re-written to read $X_1 X_1 - X_2 (X_2 X_2)^{-1} X_2 X_1$. The second term is really just $X_2 (X_2 X_2)^{-1} X_2 X_1 = X_2 \hat{\gamma}$ which equals zero because X₂ on X₁ are uncorrelated. This means the first term collapses to $(X_1 X_1)^{-1}$.

The second term in the portioned inverse $X_1'(I - X_2(X_2'X_2)^{-1}X_2')Y$ can be re-written to read $X_1'Y - X_1'X_2(X_2'X_2)^{-1}X_2'Y$ which can be re-written as $X_1'Y - \hat{\gamma}'X_2'Y$ which collapses to $X_1'Y$ because $\hat{\gamma} = 0$.

- 2. This is a dumb idea. The suggested regression of $\hat{\varepsilon} = X\pi + v$ will produce an estimate of $\hat{\pi} = (X'X)^{-1}X'\hat{\varepsilon}$ but from the initial estimate of $\hat{\beta} = (X'X)^{-1}X'Y$ we know that by construction, $X'\hat{\varepsilon} = 0$ in the sample, so for all models $\hat{\pi} = 0$. The suggested regression is not informative.
- 3. Given $\hat{\beta}^* = (x'x)^{-1}x'y^*$ substitute in the true value of y^* which equals $y^* = y + v = x\beta + \varepsilon + v$ which produces $\hat{\beta}^* = (x'x)^{-1}x'y^* = (x'x)^{-1}x'(x\beta + \varepsilon + v)$ which simplifies to read $\hat{\beta}^* = \beta + (x'x)^{-1}x'(\varepsilon + v)$. Taking expectation, we obtain

$$E[\hat{\beta}^* | x] = \beta + E[(x'x)^{-1}x'(\varepsilon + v) | x] = \beta + (x'x)^{-1}x'(E[\varepsilon | x] + E[v | x]) = \beta$$

Because E[v | x] = 0 $E[\varepsilon | x] = 0$

.

Even with measurement error in y, the estimate is still unbiased.

We know that $Var(\hat{\beta}) = \sigma_{\varepsilon}^{2} (x'x)^{-1}$ For the estimate with measurement error, note that $V[\hat{\beta}^{*} | x] = E[(\hat{\beta}^{*} - E[\hat{\beta}^{*}])(\hat{\beta}^{*} - E[\hat{\beta}^{*}])]' = E[(\hat{\beta}^{*} - \beta)(\hat{\beta}^{*} - \beta)'] = E[(x'x)^{-1}x'(\varepsilon + v)(\varepsilon' + v')x(x'x)^{-1}]$

Note that $E[(\varepsilon + v)(\varepsilon' + v')] = E[\varepsilon \varepsilon' + vv' + 2\varepsilon v']$ and by assumption $E[2\varepsilon v'] = 0$ While $E[\varepsilon\varepsilon' + vv'] = (\sigma_{\varepsilon}^{2} + \sigma_{v}^{2})(x'x)^{-1}$ Therefore $Var(\hat{\beta}^{*}) = (\sigma_{\varepsilon}^{2} + \sigma_{v}^{2})(x'x)^{-1} > Var(\hat{\beta}) = \sigma_{\varepsilon}^{2}(x'x)^{-1}$

4. The correlation coefficient between y and \hat{y} is $\hat{\rho}(y, \hat{y}) = \frac{[y'M\hat{y}/(n-1)]}{[y'My/(n-1)]^{0.5}[\hat{y}'M\hat{y}/(n-1)]^{0.5}}$

Note that the (n-1)'s in the numerator and denominator cancel so the squared correlation coefficient is then $\hat{\rho}(y, \hat{y})^2 = \frac{[y'M\hat{y}][y'M\hat{y}]}{[y'My][\hat{y}'M\hat{y}]}$. Note further that because $y = x\hat{\beta} + \hat{\varepsilon}$

$$y'M\hat{y} = (\hat{y} + \hat{\varepsilon})'M\hat{y} = \hat{y}'M\hat{y} + \hat{\varepsilon}'M\hat{y} \text{ and since } \hat{y} = x\hat{\beta} \text{ then } \hat{\varepsilon}'M\hat{y} = \hat{\varepsilon}'Mx\hat{\beta} = \hat{\varepsilon}'x\hat{\beta} = 0$$

because $X'\hat{\varepsilon} = 0$. Therefore $\hat{\rho}(y, \hat{y})^2 = \frac{[y'M\hat{y}][y'M\hat{y}]}{[y'My][\hat{y}'M\hat{y}]} = \frac{[\hat{y}'M\hat{y}][\hat{y}'M\hat{y}]}{[y'My][\hat{y}'M\hat{y}]} = \frac{[\hat{y}'M\hat{y}]}{[y'My]} = R^2$

5. This is trivial. Reduce this to a bivariate regression model. $y_{it} = \alpha + x_{it}\beta + u_i + \varepsilon_{it}$. Note that in first difference $y_{i2} - y_{i1} = \Delta y_i = (x_{i2} - x_{i1})\beta + (\varepsilon_{i2} - \varepsilon_{i1}) = \Delta x_i\beta + \Delta \varepsilon_i$. The OLS estimate is

then $\hat{\beta} = \frac{\sum_{i=1}^{n} \Delta y_i \Delta x_i}{\sum_{i=1}^{n} \Delta x_i^2}$. If there is no within-panel variation in x, then $\Delta x_i = 0$ and by including the

observation in the model, neither the numerator or the denominator is altered at all.

6. If the data is sorted by year then group, the first N observations in D will be an identify matrx I_n . This will be repeated T times, do as a result, the matrix D will look like $D = i_t \otimes I_n$

$$M = I_{nt} - D(D'D)^{-1}D'$$

$$(D'D)^{-1} = ((i_t \otimes I_n)(i_t \otimes I_n))^{-1} = (i_t \otimes I_n)^{-1} = (t \otimes I_n)^{-1} = \frac{1}{t}I_n$$

$$D(D'D)^{-1}D' = (i_t \otimes I_n)\left[\frac{1}{t} \otimes I_n\right](i_t \otimes I_n) = \left(\frac{i_t}{t} \otimes I_n\right)(i_t \otimes I_n) = \left(\frac{i_ti_t}{t} \otimes I_n\right)$$

$$M = I_{nt} - \left(\frac{i_ti_t}{t} \otimes I_n\right)$$

Which can be re-written to read

$$M = \left(I_t - \frac{i_t i_t'}{t}\right) \otimes I_n$$

7. A program is available on the web page $- ps1_q7.do$

b. Note that the sample means of within panel deviations in means are all zero out to about 6 decimal placed. Since all the x's and y's have zero mean, you should not need to include an intercept in these regressions. The standard errors in this part are so much smaller because they od not take into consideration that we SHOULD be estimating dummy variables for all groups. The DOF are NT-K but they should be NT-N-K.

c. The coefficient is 0.048 when means that workers make 4.8% more when then move from a non-union to a union job.

d. Note that the coefficient stays the same – the fixed effect estimate does not utilize information from people who do not change union status – which is the majority of people.

e. Although we delete people without within panel variation in union status the estimates from this question and part b differ because union status is correlated with other variables in the model