Econometrics 140B
Fall 2017 Midterm 3

1. What is the definition of panel data?

   Answer: Panel data contains observations of the same individuals or firms over time.

2. How can the fixed effects estimator be used to overcome endogeneity bias? Be careful to explain when this method will, and will not, work.

   Answer: A fixed effects estimator using individual fixed effects can remove all observed and unobserved time invariant individual characteristics from the regression.

3. Consider the model

$$y_i = \beta_0 + \beta_1 x_i + u_i,$$

   where $\mathbb{E}(u_i|x_i) = 0$.

   (a) Under the assumption that $Var[u_i|x] = \sigma^2$, derive the $Var[\widehat{\beta}_1|x]$.

   Answer:

$$Var\left(\widehat{\beta}^{OLS}\right) = Var\left(\frac{\sum_i (x_i - \overline{x}_i)(y_i - \overline{y}_i)}{\sum_i (x_i - \overline{x}_i)^2}\right)$$

$$= Var\left(\frac{\sum_i (x_i - \overline{x}_i)(\beta_0 + \beta_1 x_i + \varepsilon_i - \frac{1}{n}\sum_i (\beta_0 + \beta_1 x_i + \varepsilon_i))}{\sum_i (x_i - \overline{x}_i)^2}\right) \quad \text{(Plug in the model)}$$

$$= Var\left(\beta_1 + \frac{\sum_i (x_i - \overline{x}_i)(\varepsilon_i - \overline{\varepsilon}_i)}{\sum_i (x_i - \overline{x}_i)^2}\right)$$

$$= Var\left(\frac{\sum_i (x_i - \overline{x}_i)(\varepsilon_i - \overline{\varepsilon}_i)}{\sum_i (x_i - \overline{x}_i)^2}\right) \quad \text{(the variance of a constant is zero)}$$

$$= \frac{1}{\left(\sum_i (x_i - \overline{x}_i)^2\right)^2} Var\left(\sum_i (x_i - \overline{x}_i)(\varepsilon_i - \overline{\varepsilon}_i)\right) \quad \text{(pull constant out and square)}$$

$$= \frac{1}{\left(\sum_i (x_i - \overline{x}_i)^2\right)^2}$$

$$\left[\sum_i (x_i - \overline{x}_i)^2 Var(\varepsilon_i - \overline{\varepsilon}_i) + \sum_{i \neq j} (x_i - \overline{x}_i)(x_j - \overline{x}_j) Cov(\varepsilon_i - \overline{\varepsilon}_i, \varepsilon_j - \overline{\varepsilon}_j)\right]$$

$$= \frac{1}{\left(\sum_i x_i - \overline{x}_i^2\right)^2}\left[\sigma^2 \sum_i (x_i - \overline{x}_i)^2\right] \quad (Var[u_i|x] = \sigma^2 \;\&\; Cov(\varepsilon_i, \varepsilon_j) = 0)$$

$$= \frac{\sigma^2}{\sum_i (x_i - \overline{x}_i)^2}$$

Under the assumption $Cov(\varepsilon_i, \varepsilon_j) \neq 0$

$$Var\left(\widehat{\beta}^{OLS}\right) = \frac{1}{\left(\sum_i x_i - \overline{x_i}^2\right)^2}\left[\sigma^2 \sum_i (x_i - \overline{x}_i)^2 + \sum_{i \neq j}(x_i - \overline{x}_i)(x_j - \overline{x}_j)\sigma_{ij}\right] \quad (Var[u_i|x] = \sigma^2)$$

$$= \frac{\sigma^2}{\sum_i (x_i - \overline{x}_i)^2} + \frac{1}{\left(\sum_i x_i - \overline{x_i}^2\right)^2}\left[\sum_{i \neq j}(x_i - \overline{x}_i)(x_j - \overline{x}_j)\sigma_{ij}\right]$$

(b) Under the assumption that $Var[u_i|x] = \sigma_i^2$, derive the $Var[\widehat{\beta}_1|x]$.

Answer:

$$Var\left(\widehat{\beta}^{OLS}\right) = Var\left(\frac{\sum_i (x_i - \overline{x}_i)(y_i - \overline{y}_i)}{\sum_i (x_i - \overline{x}_i)^2}\right)$$

$$= Var\left(\frac{\sum_i (x_i - \overline{x}_i)(\beta_0 + \beta_1 x_i + \varepsilon_i - \frac{1}{n}\sum_i (\beta_0 + \beta_1 x_i + \varepsilon_i))}{\sum_i (x_i - \overline{x}_i)^2}\right)$$

(Plug in the model)

$$= Var\left(\beta_1 + \frac{\sum_i (x_i - \overline{x}_i)(\varepsilon_i - \overline{\varepsilon}_i)}{\sum_i (x_i - \overline{x}_i)^2}\right)$$

$$= Var\left(\frac{\sum_i (x_i - \overline{x}_i)(\varepsilon_i - \overline{\varepsilon}_i)}{\sum_i (x_i - \overline{x}_i)^2}\right) \quad \text{(the variance of a constant is zero)}$$

$$= \frac{1}{\left(\sum_i (x_i - \overline{x}_i)^2\right)^2} Var\left(\sum_i (x_i - \overline{x}_i)(\varepsilon_i - \overline{\varepsilon}_i)\right) \quad \text{(pull constant out and square)}$$

$$= \frac{1}{\left(\sum_i (x_i - \overline{x}_i)^2\right)^2}$$

$$\left[\sum_i (x_i - \overline{x}_i)^2 Var(\varepsilon_i - \overline{\varepsilon}_i) + \sum_{i \neq j}(x_i - \overline{x}_i)(x_j - \overline{x}_j)Cov(\varepsilon_i - \overline{\varepsilon}_i, \varepsilon_j - \overline{\varepsilon}_j)\right]$$

$$= \frac{1}{\left(\sum_i x_i - \overline{x_i}^2\right)^2}\left[\sum_i (x_i - \overline{x}_i)^2 \sigma_i^2\right] \quad (Var[u_i|x] = \sigma_i^2 \ \& \ Cov(\varepsilon_i, \varepsilon_j) = 0)$$

Under the assumption $Cov(\varepsilon_i, \varepsilon_j) \neq 0$

$$Var\left(\widehat{\beta}^{OLS}\right) = \frac{1}{\left(\sum_i x_i - \overline{x_i}^2\right)^2}\left[\sum_i (x_i - \overline{x}_i)^2 \sigma_i^2 + \sum_{i \neq j}(x_i - \overline{x}_i)(x_j - \overline{x}_j)\sigma_{ij}\right] \quad (Var[u_i|x] = \sigma_i^2)$$