

Econometrics 1 - Homework 5
Spring 2017

The due date is Thursday April 26

1. Consider the panel data model,

$$y_{it} = x'_{it}\beta + c_i + e_{it}$$

Assume that the residual e_{it} is serially uncorrelated and homoskedastic. The “first difference” estimator is OLS of $y_{it} - y_{it-1}$ on $x_{it} - x_{it-1}$.

- (a) Under what conditions will this estimator be consistent for β ?
 - (b) Derive the asymptotic properties of this estimator as $n \rightarrow \infty$ with T fixed imposing the conditions in part a.
 - (c) How would one estimate the variance covariance matrix under homoskedasticity?
 - (d) Show that this estimator is the same as fixed effects estimator when $T = 2$.
 - (e) Assuming homoskedasticity of the residual e_{it} and that $T = 3$ compare the variance of the first difference estimator with that of the fixed effects estimator.
2. Using the data in wagepan.dta estimate the wage equation using the explanatory variables: exper, expersq, educ, union, black, poohlth. Use each of the estimators considered in class (OLS, FE, RE) and compute appropriate standard errors. Also compute the Hausman test that tests the assumption that the regressors are uncorrelated with the “fixed effect”. Re-do the results using year dummy variables.

3. Consider the latent variable,

$$\begin{aligned} y_i^* &= x_i'\beta + \varepsilon_i \\ \varepsilon_i &\sim N(0, \sigma^2) \end{aligned}$$

As noted in class we get Probit, Truncated and Censored models when we observe each of the following dependent variables,

$$\begin{aligned} y_i &= 1(y_i^* > 0) \\ y_i &= y_i^* \text{ if } y_i^* > 0 \\ y_i &= \max\{0, y_i^*\} \end{aligned}$$

If the above model is correct what would it imply about the coefficients in each case. ie.

- (a) Should the coefficients in each model look similar (note in the Probit model you will have to normalize $\sigma = 1$).
- (b) Using the data in MROZ.DTA estimate each of these models for hours (you pick covariates that you think are appropriate) and see if the estimates satisfy the conditions that you have derived.

- (c) Estimate the (unconditional) distribution of hours for positive hours only and see if you think the censored or truncated model would be appropriate.
4. Using the cps data estimate each of the following:
- (a) The density of hourly wage using a normal distribution
 - (b) The density of hourly wage using the default kernel estimator in STATA
 - (c) Repeat a. and b. for log wages and determine whether the normality assumption is reasonable for either situation (ie for wages or log wages)
 - (d) Do a non-parametric regression of wage (and then log wage) on experience separately for white males and white females and discuss the appropriateness of the quadratic model that we typically use as well as any differences between these two regressions.