**Economics 628** **Term 1 2017/2018**

**Topics in Applied Econometrics I** **Hiro Kasahara**

# Homework 3

(Due: Monday, October 16 at the *start* of the class)

Note: Study groups *discussing* the problems are strongly encouraged. But please write your own answers and submit your own programs (no copy and paste from your classmate's program!).

## 1 Bond (2002)

Read Bond (2002) and replicate Table 4 of Bond (2002) using STATA. The reference by Steve Bond on how to replicate Table 4 is available here.

Install xtabond2 by typing the following in STATA command:

```
ssc install xtabond2
```

Go through help files of xtset, xtreg, and xtabond2. Namely, type

```
help xtset
help xtreg
help xtabond2
```

Following the reference while referring to the relevant part of Bond (2002) to replicate Table 4. You should learn:

- how to use "xtset".
- how to use "xtreg" to estimate fixed effects model. The option of "fe vce(cluster id)" means that a fixed effects model based on within transformation is estimated while the standard errors are estimated clustering at firm-level (see equation (10.59), page 311 of Wooldridge).
- how to use "xtabond2" to estimate first-differenced GMM using lagged variables as instruments.
- how to use "xtabond2" to estimate System GMM.

## 2 Table 11.2 of Wooldridge (2010, page 373)

Replicate Table 11.2 of Wooldridge (2010, page 373) using airfare.dta and add two additional columns: the one with fixed effects and the other with System GMM. Also add rows that report the result of test for overidentifying restrictions (so called, J-test or Sargan test). Also, report the Difference Sargan statistic and discuss whether the additional moment conditions used in the levels equations are valid or not.

## 3 Measurement errors in panel data

Consider a panel data model

$$y_{it} = \beta x_{it}^* + \alpha_i + \epsilon_{it} \tag{1}$$

with $x_{it}^*$ is scalar, $E[x_{it}^*] = E[\alpha_i] = 0$, $E[\epsilon_{it}|x_{i1}^*, ..., x_{iT}^*] = 0$, and $E[\epsilon_{it}\epsilon_{is}|x_{i1}^*, ..., x_{iT}^*] = 0$ when $t \neq s$.

Furthermore, we do not observe $x_{it}^*$ but we observe its proxy $x_{it}$ which is subject to a measurement error: $x_{it} = x_{it}^* + \omega_{it}$, where we assume that $E[\omega_{it}|\alpha_i, x_{it}^*] = E[\omega_{it}\epsilon_{it}|\alpha_i, x_{it}^*] = 0$. We have a random sample of $\{\{y_{it}, x_{it}\}_{t=1}^T\}_{i=1}^n$ from (1), where we consider fixed $T$ while $n \to \infty$. For simplicity, we assume that $\sigma_{x^*\alpha} := E[\alpha_i x_{it}^*]$, $\sigma_\omega^2 := E[\omega_{it}^2]$, and $\sigma_{x^*}^2 := E[(x_{it}^*)^2]$ for all $t$ although both $x_{it}^*$ and $\omega_{it}$ are potentially serially correlated.

1. Suppose that we estimate $\beta$ by pooled OLS, denoted by $\hat{\beta}_{OLS}$, by regressing $y_{it}$ on $x_{it}$ (without intercept). If you are confused, consider the case when $T = 1$.

   (a) Show that $\hat{\beta}_{OLS} \to_p \beta + \frac{\sigma_{x^*\alpha} - \beta\sigma_\omega^2}{\sigma_{x^*}^2 + \sigma_\omega^2}$.

   (b) What would you conclude the direction of OLS bias when $\sigma_{x^*\alpha} = 0$?

   (c) Characterize the asymptotic distribution of $\hat{\beta}_{OLS}$.

2. Now, suppose we estimate $\beta$ by fixed effects estimator, denoted by $\hat{\beta}_{FE}$, by regressing $y_{it} - \bar{y}_i$ on $x_{it} - \bar{x}_i$ where $\bar{y}_i = (1/T)\sum_{t=1}^T y_{it}$ etc.. Derive the asymptotic bias of this fixed effects estimator and characterize the asymptotic distribution of $\hat{\beta}_{FE}$. Discuss whether you recommend using the fixed effects estimator over the pooled OLS estimator. Please explain your reasoning.