Test of Random vs Fixed Effects with Small Within Variation: 
Online Appendix

JINYONG HAHN, JOHN HAM, & HYUNGSIK ROGER MOON
A Normality of the Within Estimator

For simplicity of notation, we will assume that $x$ is a scalar and $T = 2$. The within estimator has the characterization

$$\bar{\beta} - \beta = \left( \sum_{i=1}^{N} \bar{x}_i^2 \right)^{-1} \sum_{i=1}^{N} \bar{x}_i \bar{\varepsilon}_i.$$ 

We will assume that $\sum_{i=1}^{N} \bar{x}_i^2 = o(N)$. For simplicity, assume that $E[\bar{\varepsilon}_i^2] = 1$. Let

$$Y_{N,i} = \left( \sum_{i=1}^{N} \bar{x}_i^2 \right)^{-1/2} \bar{x}_i \bar{\varepsilon}_i,$$

and we have

$$E[Y_{N,i}] = 0 \quad \text{and} \quad E\left[ \sum_{i=1}^{N} Y_{N,i}^2 \right] = 1.$$ 

Then, by the Lindeberg-Feller (LF) central limit theorem (CLT),

$$\left( \sum_{i=1}^{N} \bar{x}_i^2 \right)^{1/2} \left( \bar{\beta} - \beta \right) \Rightarrow \mathcal{N}(0, 1),$$

if

$$E \left[ \sum_{i=1}^{N} Y_{N,i}^2 \left\{ |Y_{N,i}| > \eta \right\} \right] = \sum_{i=1}^{N} E \left[ \frac{\bar{x}_i \bar{\varepsilon}_i}{\sum_{i=1}^{N} \bar{x}_i^2} \left\{ |\bar{x}_i \bar{\varepsilon}_i| > \left( \sum_{i=1}^{N} \bar{x}_i^2 \right)^{1/2} \eta \right\} \right] \rightarrow 0 \quad (13)$$

for any $\eta > 0$. In our case, a well known sufficient condition for (13) is the Lyapounov condition

$$E \sum_{i=1}^{N} |Y_{N,i}|^3 \rightarrow 0. \quad (14)$$

In the standard case where

$$\frac{1}{N} \sum_{i=1}^{N} \bar{x}_i^2 \rightarrow_p C > 0,$$

and

$$\frac{1}{N} \sum_{i=1}^{N} |\bar{x}_i|^3 = O_p(1) \quad \text{and} \quad E\left[ |\bar{x}_i|^3 \right] < \infty,$$
conditional on \( \tilde{x}_i \), we have

\[
\sum_{i=1}^{N} E \left[ \frac{\tilde{x}_i^2 e_i^2}{\sum_{i=1}^{N} \tilde{x}_i^2} 1 \left\{ |\tilde{x}_i \tilde{z}_i| > \left( \sum_{i=1}^{N} \tilde{x}_i^2 \right)^{1/2} \eta \right\} \right]
\]

\[
\leq \frac{1}{\eta} \sum_{i=1}^{N} E \left[ \frac{|\tilde{x}_i|^3 |\tilde{z}_i|^3}{\left( \sum_{i=1}^{N} \tilde{x}_i^2 \right)^{3/2}} \right]
\]

\[
= \frac{E |\tilde{z}_i|^3}{\eta} \frac{\sum_{i=1}^{N} |\tilde{x}_i|^3}{\left( \sum_{i=1}^{N} \tilde{x}_i^2 \right)^{3/2}}
\]

\[
= \frac{E |\tilde{z}_i|^3}{\eta} \frac{1}{N} \frac{\sum_{i=1}^{N} |\tilde{x}_i|^3}{\left( \frac{1}{N} \sum_{i=1}^{N} \tilde{x}_i^2 \right)^{3/2}} = O_p \left( \frac{1}{\sqrt{N}} \right).
\]

Thus, conditional on \( \tilde{x}_i \), the LF condition is satisfied and

\[
\left( \sum_{i=1}^{N} \tilde{x}_i^2 \right)^{1/2} \left( \tilde{\beta} - \beta \right) \Rightarrow N(0, 1).
\]

Since the limit does not depends on \( \tilde{x}_i \), the CLT follows unconditionally.

In our case, we have \( \sum_{i=1}^{N} \tilde{x}_i^2 = o(N) \), but even in this case the CLT will follow as long as

\[
\frac{\sum_{i=1}^{N} |\tilde{x}_i|^3}{\left( \sum_{i=1}^{N} \tilde{x}_i^2 \right)^{3/2}} \to 0.
\]

Then,

\[
\sum_{i=1}^{N} E \left[ \frac{\tilde{x}_i^2 e_i^2}{\sum_{i=1}^{N} \tilde{x}_i^2} 1 \left\{ |\tilde{x}_i \tilde{z}_i| > \left( \sum_{i=1}^{N} \tilde{x}_i^2 \right)^{1/2} \eta \right\} \right]
\]

\[
\leq \frac{1}{\eta} \sum_{i=1}^{N} E \left[ \frac{|\tilde{x}_i|^3 |\tilde{z}_i|^3}{\left( \sum_{i=1}^{N} \tilde{x}_i^2 \right)^{3/2}} \right]
\]

\[
= \frac{E |\tilde{z}_i|^3}{\eta} \frac{\sum_{i=1}^{N} |\tilde{x}_i|^3}{\left( \sum_{i=1}^{N} \tilde{x}_i^2 \right)^{3/2}} \to 0,
\]

and

\[
\left( \sum_{i=1}^{N} \tilde{x}_i^2 \right)^{1/2} \left( \tilde{\beta} - \beta \right) \Rightarrow N(0, 1).
\]

This condition is expected to be violated, for example, when \( x_{it} \) is a dummy variable and there are very few observations where the binary regressors change over time. Suppose that the number of such observations is fixed at \( r \). Then, since \( \tilde{x}_i \in \{-1, 0, 1\} \),

\[
\sum_{i=1}^{N} \tilde{x}_i^2 = r, \quad \sum_{i=1}^{N} |\tilde{x}_i|^3 = r,
\]
and
\[ \frac{\sum_{i=1}^{N} |\tilde{\varepsilon}_i|^3}{\left( \sum_{i=1}^{N} \tilde{x}_i^2 \right)^{3/2}} = \frac{r}{r} = 1. \]

B  Proofs

Proof of Lemma 1  Because the \( \tilde{\varepsilon}_i \) are iid with finite variance across \( i \), it follows that
\[ \left( \sum_{i=1}^{N} \tilde{x}_i' \tilde{x}_i \right)^{1/2} (\tilde{\beta} - \beta) = \sum_{i=1}^{N} \left( \sum_{i=1}^{N} \tilde{x}_i' \tilde{x}_i \right)^{-1/2} \tilde{x}_i' \tilde{\varepsilon}_i \]
is an \( L_2 \)–martingale with respect to the natural filtration. Also notice that
\[
\sup_N E \left\| \sum_{i=1}^{N} \left( \sum_{i=1}^{N} \tilde{x}_i' \tilde{x}_i \right)^{-1/2} \tilde{x}_i' \tilde{\varepsilon}_i \right\|^2 = \sup_N E \left[ \text{trace} \left( (\tilde{\beta} - \beta)' \left( \sum_{i=1}^{N} \tilde{x}_i' \tilde{x}_i \right) (\tilde{\beta} - \beta) \right) \right]
\]
\[ = \sup_N \text{trace} \left\{ \left( \sum_{i=1}^{N} \tilde{x}_i' \tilde{x}_i \right) \cdot \sigma^2 \left( \sum_{i=1}^{N} \tilde{x}_i' \tilde{x}_i \right)^{-1} \right\}
\]
\[ = \dim (\beta) = K < \infty. \]

Then, by the martingale convergence theorem (e.g., Theorem 10.5.4 in Dudley(1989)), the martingale
\[ \left( \sum_{i=1}^{N} \tilde{x}_i' \tilde{x}_i \right)^{1/2} (\tilde{\beta} - \beta) \]
converges almost surely to a finite limit random variable, from which we deduce the required result for the lemma. The second claim follows easily from \( \tilde{\beta} = \beta + o_a.s. (1) \).

Proof of Lemma 2  Note first that
\[ \frac{1}{N} \sum_{i=1}^{N} \left( \tilde{\varepsilon}_i - \bar{\varepsilon} \right)' (\tilde{\varepsilon}_i - \bar{\varepsilon}) = \sigma^2 + o_a.s (1) \]
by the strong law of large numbers. Also note that
\[ \frac{1}{N} \sum_{i=1}^{N} \sum_{t=1}^{T} \left( (\tilde{x}_{it} - \bar{x}_t)' (\tilde{\beta} - \beta) \right)^2 = \left( \tilde{\beta} - \beta \right)' \left( \frac{1}{N} \sum_{i=1}^{N} \sum_{t=1}^{T} (\tilde{x}_{it} - \bar{x}_t)' (\tilde{x}_{it} - \bar{x}_t) \right) (\tilde{\beta} - \beta)
\]
\[ \leq \left( \tilde{\beta} - \beta \right)' \left( \frac{1}{N} \sum_{i=1}^{N} \sum_{t=1}^{T} \tilde{x}_{it} \tilde{x}_t' \right) (\tilde{\beta} - \beta)
\]
\[ = \left( \tilde{\beta} - \beta \right)' \left( \frac{1}{N} \sum_{i=1}^{N} \tilde{x}_i' \tilde{x}_i \right) (\tilde{\beta} - \beta)
\]
\[ = o_a.s. (1), \]
where the last equality is based on Lemma 1. These two facts, combined with the equality \( \tilde{e}_i = \tilde{\varepsilon}_i - \bar{\varepsilon} - (\tilde{x}_i - \bar{x}) (\tilde{\beta} - \beta) \), imply the desired conclusion.

\[ ^1 \text{See Lai, Robbins, and Wei (1979) - references in this appendix are contained in the main references of the paper.} \]
References