Of Nickell Bias, Cross-Sectional Dependence, and Their Cures: Reply

Khusrav Gaibulloev
Department of Economics
American University of Sharjah
Sharjah, UAE
kgaibulloev@aus.edu
and
Todd Sandler
Donggyu Sul
School of Economic, Political & Policy Sciences
University of Texas at Dallas, GR 31
800 W. Campbell Road
Richardson, TX 75080 USA
tsandler@utdallas.edu
d.sul@utdallas.edu

Final Draft: February 2014

In Gaibulloev, Sandler, and Sul (GSS) (2014), we discussed estimation of the fixed-effects dynamic regression model under cross-sectional dependence. Although Nickell bias becomes less of a problem for reasonably large *T*, we showed that its impact on statistical inference remains serious as long as *N* is greater then *T*. We suggested dividing data into subsamples to ensure that *N* is sufficiently smaller than *T* for correct statistical inference, and applying the factor augmented regression method to address the cross-sectional dependence.

Beck, Katz, and Mignozzetti (BKM) (2014) claimed that with sufficiently large *T* "there is essentially no harm" in using the standard fixed-effects estimator "regardless of the size of *N*". For those who think that the standard fixed-effect estimator is problematic, BKM strongly argued against using our subsample suggestion on the ground that it leads to the loss of efficiency. Instead, BKM suggested using methods such as bootstrap or other standard procedures to take advantage of the efficiency gains of using the full sample.

BKM's suggestion (of using alternative methods) is valid *only when there is no cross-sectional dependence*. Of course, under the assumption of zero cross-sectional dependence, typical GMM/IV or mean unbiased estimators can eliminate the Nickell bias; the limiting distributions of their *t*-statistics become a standard normal distribution regardless of the size of *N* and *T*. Hence, the use of the full sample leads to more efficient testing and estimation. However, typical GMM/IV estimators do not deal with cross-sectional dependence. Therefore, the *t*-statistics based on standard GMM/IV estimators are generally unknown under the presence of cross-sectional dependence.

The simplest solution for this problem is the use of the nonparametric sieve bootstrap combined with GMM/IV or other unbiased estimators: Drawing *N*-columns of pseudo variables jointly maintains the information of the cross-sectional dependence. We discussed the validity of this sieve bootstrap in GSS (2014) under a strong exogeneity assumption. That is, the regressors should be independent from the regression errors; otherwise, as we suggested, the factor-augmented estimator should be used. As mentioned in GSS (2014) just prior to equation (7), the bootstrap fails when the limiting distribution is dependent on a nuisance parameter (not a pivotal statistic). That is, when Nickell bias is present in the limiting distribution (when $N > T$), the bootstrap fails. Therefore, we suggested using the subsamples so that $N < T$ holds. The well-known paper by Horowitz (2001) has a detailed discussion of the validity of the bootstrap.

In sum, BKM did not understand the main point of GSS (2014), which is clear from BKM's footnote 1. This footnote incorrectly asserted that the cross-sectional dependence issue is orthogonal to Nickell bias. The correct answers to BKM's second and third questions, "(2) is the problem serious in applied work; and (3) does the proposed solution do more good than harm," are both yes, contrary to their answers.

Next, we want to make few additional points for a situation when there is no cross-sectional dependence. First, BKM's argument can be summarized by the following single equation. As $N, T \to \infty$, the limiting distributions of the within-group estimator $\hat{\boldsymbol{\beta}}$, both under the null and alternative hypotheses, can be written as

$$\sqrt{NT}\hat{\boldsymbol{\beta}} \to^{d} \sqrt{NT}\boldsymbol{\beta} + B\sqrt{N/T} + N\left(0, \Sigma_{\boldsymbol{\beta}}^{2}\right), \tag{1}$$

where B is the Nickell bias term, $\boldsymbol{\beta}$ is the vector of the true values, and $\Sigma_{\boldsymbol{\beta}}^{2}$ is the covariance matrix of $\hat{\boldsymbol{\beta}}$. BKM argued that using the full sample leads to a more efficient estimator. We, of course, agree because the convergence rate is $\sqrt{NT}$, so that the variance shrinks at this rate. Hence, the estimators become more efficient as N and T increase. Second, when BKM discussed statistical inference, they should have used the limiting distribution of the *t*-statistics, which can be obtained by dividing both sides of equation (1) by $\Sigma_{\boldsymbol{\beta}}$. That is,

$$t_{\hat{\boldsymbol{\beta}}} = \sqrt{NT}\hat{\boldsymbol{\beta}} / \Sigma_{\boldsymbol{\beta}} \to^{d} \sqrt{NT}\boldsymbol{\beta} / \Sigma_{\boldsymbol{\beta}} + (B / \Sigma_{\boldsymbol{\beta}})\sqrt{N / T} + N\left(0,1\right). \tag{2}$$

Keep in mind that, here, we are assuming no cross-sectional dependence. From a back of the envelope calculation, let $\boldsymbol{\beta} = 0.3$ (for a single regressor case and BKM's experiment). With $T = 40$ and $N = 20$, the *t*-statistic for $\hat{\boldsymbol{\beta}}$ will be $8.5 + N\left(0,1\right)$ when B can be ignored and $\Sigma_{\boldsymbol{\beta}} = 1$. That is, the null hypothesis will be rejected almost always. In other words, the power of the test will be almost perfect even with small *N*. When $\boldsymbol{\beta} = 0.1$, then $\sqrt{800} \times 0.1 = 2.82$, so that the asymptotic power of the test is approximately 0.88, which is close to one. Third, this simple exercise leads to a somewhat important practical guideline. Suppose that the Nickell bias is small and *T* is large. Then, in practice one does not need to correct the bias because *B* can be ignored. This is true especially when a smaller *N* is used so that the $\sqrt{N / T}$ term reduces the bias further. Next, consider applying the standard GMM/IV estimator on the full sample. The standard GMM/IV estimator may correct the tiny bias but it usually leads to an increase in the variance in the finite sample. Denote $\tilde{\Sigma}_{\boldsymbol{\beta}}$ as the standard deviation of the GMM/IV estimator. Then $\boldsymbol{\beta} / \tilde{\Sigma}_{\boldsymbol{\beta}}$ is usually smaller than $\boldsymbol{\beta} / \Sigma_{\boldsymbol{\beta}}$ in the finite sample (not asymptotically). In other words, the *t*-statistic of the GMM/IV estimator depends on the magnitude of $\tilde{\Sigma}_{\boldsymbol{\beta}}$. Generally, one should worry first about statistical inference and then about efficiency even when the bias is small. Many studies in political economy are policy oriented and making policy recommendation based on incorrect statistical inferences can have significant practical consequences.

References

Beck, N., J. N. Katz, and U. Mignozzetti. 2014. Of Nickell bias and its cures: Comment on Gaibulloev, Sandler, and Sul. *Political Analysis* 22:forthcoming.

Gaibulloev, K., T. Sandler, and D. Sul. 2014. Dynamic panel analysis under cross-sectional dependence. *Political Analysis* 22:forthcoming.

Horowitz, J. L. 2001. The bootstap. In *Handbook of econometrics, Vol. 5*, eds. J. J. Heckman and E. E. Leamer, 3159–3228. Amsterdam: North-Holland.