Peter Hull June 26, 2017

Fixed Effects Identify a Weighted Sum of First- and Long-Differences Only Under a "Parallel Trends" Assumption

It is now well-known that the fixed effect (FE) and first-difference (FD) estimators produce identical results in panel data with only two periods, but not otherwise (Griliches and Hausman, 1986; Angrist and Pischke, 2009). Nevertheless, applied researchers often motivate FE regressions with two-period logic, perhaps reflecting an intuition that multi-period models should estimate a weighted average of constituent two-period comparisons. This note shows that this is not the case generally, but is under a "parallel trends" assumption of the kind frequently motivating panel data techniques. Intuitively, FE estimation over long panels imposes a constant individual effect, while different FDs allow for different effects at various points in time. Although parallel trends likely rules out this heterogeneity in large samples, in practice FE coefficients may lie quite far outside of the convex hull of component FD estimates.

Formally, consider the balanced panel specification

$$y_{it} = \alpha_i + \tau_t + x_{it}'\beta + \epsilon_{it} \tag{1}$$

for $t = 0, \ldots, T$. It is straightforward to verify that a FE regression of (1) produces the same estimates as that of

$$y_{itc} = \alpha_i + \tau_t + x_{itc}'\beta + \epsilon_{itc}, \tag{2}$$

where we relabel $x_{it0} = x_{it}$ and $y_{it0} = x_{it}$ and generate x_{it1} and y_{it1} as duplicate copies of x_{it} and y_{it}, respectively. This specification is in turn nested by the richer model

$$y_{itc} = \widetilde{\alpha}_i + \widetilde{\tau}_t + x_{itc}'\widetilde{\beta} + \sum_{s=1}^{T} FD_{tcs}(\widetilde{\alpha}_{is} + \widetilde{\tau}_{ts} + x_{it}'\widetilde{\beta}_s) + \widetilde{\epsilon}_{it}, \tag{3}$$

where

$$FD_{tcs} = \mathbf{1}\{(t = s-1 \text{ and } c = 0) \text{ or } (t = s \text{ and } c = 1)\} \tag{4}$$

is a "first difference sample" indicator for disjoint sets of sequential two-period cuts of the original data. That is, FD_{tc1} identifies the original $t = 0$ observations and the copied $t = 1$ observations, FD_{tc2} identifies the original $t = 1$ observations and the copied $t = 2$ observations, and so on.

By this construction, OLS of equation (3) will produce estimates $\widehat{\widetilde{\alpha}}_i$, $\widehat{\widetilde{\tau}}_t$, and $\widehat{\widetilde{\beta}}$ that are numerically equivalent to what would be obtained from running OLS of equation (1) on just the observations with $t = 0$ or $t = T$. Call these $\widehat{\alpha}_{i0}$, $\widehat{\tau}_{t0}$, and $\widehat{\beta}_0$. Furthermore, the OLS estimates $\widehat{\widetilde{\beta}}_s$ of each $\widetilde{\beta}_s$ in equation (3) will numerically equal the difference between $\widehat{\beta}_s$ and $\widehat{\beta}_0$, where $\widehat{\beta}_s$ denotes the OLS estimate of β in equation (1) using just the observations with $t = s-1$ or $t = s$, and similarly for $\widehat{\widetilde{\alpha}}_{is}$ and $\widehat{\widetilde{\tau}}_{ts}$:

$$\begin{aligned} \widehat{\widetilde{\beta}}_s &= \widehat{\beta}_s - \widehat{\beta}_0 \\ \widehat{\widetilde{\alpha}}_{is} &= \widehat{\alpha}_{is} - \widehat{\alpha}_{i0} \\ \widehat{\widetilde{\tau}}_{ts} &= \widehat{\tau}_{ts} - \widehat{\tau}_{t0}. \end{aligned} \tag{5}$$

By the well-known result mentioned at the outset, each $\widehat{\beta}_s$ is numerically the coefficient obtained from the first-differenced regression of $y_{is} - y_{i,s-1}$ on $x_{is} - x_{i,s-1}$, while $\widehat{\beta}_0$ is the coefficient from the "long-differenced" regression of $y_{iT} - y_{i1}$ on $x_{iT} - x_{i1}$.

Next, note that we can use the usual omitted-variables bias formula to write the regression coefficient estimate $\widehat{\beta}$ from equation (2) in terms of the coefficient estimates of the richer model (3). Namely, we have

$$\widehat{\beta} = \widehat{\widehat{\beta}} + \sum_{s=1}^{T} R_s, \tag{6}$$

where R_s denotes the coefficient from regressing $FD_{tcs}(\widehat{\widehat{\alpha}}_{is} + \widehat{\widehat{\tau}}_{ts} + x'_{it}\widehat{\widehat{\beta}}_s)$ on x_{itc}, controlling for i and t main effects. In a balanced panel, the regression of each $FD_{tcs}\widehat{\tau}^L_{ts}$ on x_{itc} will be zero after controlling for t effects, so we can drop these terms from the R_s. Thus, with

$$\overline{x}_{it} = x_{it} - \frac{1}{N}\sum_i x_{it} - \frac{1}{T+1}\sum_t x_{it} - \frac{1}{N(T+1)}\sum_{it} x_{it} \tag{7}$$

denoting the residuals from regressing x_{it} on i and t effects, we have by applying (5) that

$$\begin{aligned}
\widehat{\beta} =& \widehat{\widehat{\beta}} + \sum_{s=1}^{T}\left(\left[\sum_{it}\overline{x}_{it}\overline{x}'_{it}\right]^{-1}\left[\sum_{it}\overline{x}_{it}\left(FD_{tcs}\widehat{\widehat{\alpha}}_{is} + FD_{tcs}x'_{it}\widehat{\widehat{\beta}}_s\right)\right]\right)\\
=& \widehat{\beta}_0 + \sum_{s=1}^{T}\left(\left[\sum_{it}\overline{x}_{it}\overline{x}'_{it}\right]^{-1}\left[\sum_{it}\overline{x}_{it}\left(FD_{tcs}(\widehat{\alpha}_{is} - \widehat{\alpha}_{i0}) + FD_{tcs}x'_{it}(\widehat{\beta}_s - \widehat{\beta}_0)\right)\right]\right)\\
=& \sum_{s=1}^{T}\widehat{W}_s\widehat{\beta}_s + \left(I - \sum_{s=1}^{T}\widehat{W}_s\right)\widehat{\beta}_0 + \left[\sum_{it}\overline{x}_{it}\overline{x}'_{it}\right]^{-1}\sum_{s=1}^{T}\left[\sum_{it}\overline{x}_{it}FD_{tcs}(\widehat{\alpha}_{is} - \widehat{\alpha}_{i0})\right],
\end{aligned} \tag{8}$$

where I is the identity matrix and $\widehat{W}_s = \left[\sum_{it}\overline{x}_{it}\overline{x}'_{it}\right]^{-1}\left[\sum_{it}\overline{x}_{it}FD_{tcs}x'_{it}\right]$.

Equation (8) shows that the FE estimates of β in equation (1) can be written as a matrix-weighted sum of first- and long-difference estimates $\widehat{\beta}_s$ and $\widehat{\beta}_0$, with weights summing to the identity matrix, plus a "bias" term which is not in general equal to zero. However, note that if

$$\frac{1}{N(T+1)}\sum_{s=1}^{T}\sum_{it}\overline{x}_{it}FD_{tcs}(\widehat{\alpha}_{is} - \widehat{\alpha}_{i0}) \xrightarrow{p} 0, \tag{9}$$

we will have

$$\widehat{\beta} \xrightarrow{p} \sum_{s=1}^{T} W_s\beta_s + \left(I - \sum_{s=1}^{T} W_s\right)\beta_0, \tag{10}$$

where β_s and β_0 are the population short- and long-difference coefficients, and W_s is the probability limit of the weighting matrix $\widehat{W}_s$.

We have thus shown that (9) is the key condition for a FE regression to estimate a weighted average of two-period model coefficients. In general, unless the set of $plim \frac{1}{N(T+1)} \sum_{it} \overline{x}_{it} FD_{tcs}(\widehat{\alpha}_{is} - \widehat{\alpha}_{i0})$ for $s = 1, \ldots, T$ exactly offset each other, each will have to be zero for equation (9) to hold. This will be the case when differences in individual coefficients across the long-difference and each short-difference model are uncorrelated with the regressors x_{it} in the short-differenced sample, conditional on individual and time main effects. Intuitively, we can think of this as a kind of parallel trends assumption, stating that the residual variation of interest in x_{it} is uncorrelated with the growth in sample-mean y_{is} over two-period samples s.

Finally, it is worth noting that this derivation suggests a lightly-modified FE specification that will always produce a weighted average of first- and long-difference coefficients, whether or not (9) holds. Observe that the "bias" term in equation (8) comes from the fact that we omit the sample-specific intercepts $\widetilde{\alpha}_{is}$ effects from equation (3) in estimating equation (2). If instead we allow sample-mean y_{is} to systematically vary by including sample-specific individual effects in this regression, $\widehat{\beta}$ will numerically equal a weighted sum of the $\widehat{\beta}_0$ and $\widehat{\beta}_s$ coefficients. With parallel trend assumptions like equation (9) typically underlying the validity of panel data specifications, it is unclear from what circumstances such a specification would arise. Nevertheless a researcher could use this specification to estimate the population bias term and test whether it is statistically or practically significant enough to raise concerns.

References

Angrist, J. and J.-S. Pischke, Mostly Harmless Econometrics: An Empiricist's Companion, Princeton University Press (2009).

Griliches, Z and J. Hausman, "Errors in Variables in Panel Data," *Journal of Econometrics,* 31 (1986), 93-118.