Political Analysis (2011) 19:170–172
doi:10.1093/pan/mpr011

Reply to Rejoinder by Plümper and Troeger

William Greene
Department of Economics, Stern School of Business, New York University, New York, NY 10012
e-mail: wgreene@stern.nyu.edu

The remarks in my paper, ''Fixed Effects Vector Decomposition: A Magical Solution to the Problem of Time Invariant Variables in Fixed Effects Models?,'' were a comment on the econometric methods proposed in Plümper and Troeger, ''Efficient Estimation of Time-Invariant and Rarely Changing Variables in Finite Sample Panel Analyses with Unit Fixed Effects,'' *Political Analysis*, 15, 2007, which I will refer to as PT in the following. The remarks below will respond to the authors' comments published in this issue, which I will refer to as P&T.

PT's Equation (1) and the explanatory text define the model they intend to analyze:

> Recall the data-generating process (DGP) of a FE model with time-invariant variables:
>
> $$y_{it} = \alpha + \sum_{k=1}^{K} \beta_k x_{kit} + \sum_{m=1}^{M} \gamma_m z_{mi} + u_i + \varepsilon_{it}, \quad (1)$$
>
> where the x variables are time varying and the z variables are assumed to be time invariant, u_i denotes the $N - 1$ unit-specific effects (FE) of the DGP, is the independent and identically distributed error term, α is the intercept of the base unit, and β and γ are the parameters to be estimated (127).

After a series of steps, PT arrive at the following two paragraphs that conclude their Section 3:

> In stage 3, we rerun the full model without the unit effects but include the unexplained part h_i of the decomposed unit FE vector obtained in stage 2. *This stage is estimated by pooled OLS.* [Emphasis added.]
>
> $$y_{it} = \alpha + \sum_{k=1}^{K} \beta_k x_{kit} + \sum_{m=1}^{M} \gamma_m z_{mi} + \delta h_i + \varepsilon_{it}. \quad (7)$$
>
> By design, h_i is no longer correlated with the vector of the z variables. If the time-invariant variables are assumed to be orthogonal to the unobserved unit effects—i.e., if the assumption underlying our estimator is correct—the estimator is consistent. If this assumption is violated, the estimated coefficients for the time-invariant variables are biased, but this bias is of course just the normal omitted variable bias. Yet, given that the estimated unit effects $\hat{ui}$ consist of much more than the real unit effect u_i and since we cannot disentangle the true elements of u_i from the between variation of the observed and included variables, researchers necessarily face a choice between using as much information as possible and using an unbiased estimator. The fevd procedure thus gives as much power as possible to the available variables unless the within variation is sufficiently large to guarantee efficient estimation.
>
> The estimation of stage 3 proves necessary for various reasons. First of all, *only the third stage allows obtaining the correct SEs.* [Emphasis added.] Not correcting the degrees of freedom leads to a potentially serious underestimation of SEs and overconfidence in the results ... (129)

(Some additional irrelevant commentary about robust estimation at the end of the second paragraph is omitted.)

The authors recommend a computation used to compute the third step of the FEVD estimator. The second paragraph is the sole (albeit, only implicit) mention in PT about how one should compute SEs for the FEVD estimator. Upon reading this prescription, I asked what the analyst would obtain if they carried out the recommended computation and whether that computation would, indeed, produce the ''correct standard errors,'' as claimed by PT. Pooled ordinary least squares (OLS) at stage 3 produces estimates of (α, β, γ, δ) and what would appear to be a covariance matrix for the full set of estimates. PT state next that ''only the third stage allows obtaining the correct SEs.'' They go on to suggest that the third stage is

© The Author 2011. Published by Oxford University Press on behalf of the Society for Political Methodology.
All rights reserved. For Permissions, please email: journals.permissions@oup.com

https://doi.org/10.1093/pan/mpr011 Published online by Cambridge University Press

making a correction for degrees of freedom, but it is not clear what that correction is, as will be evident shortly. PT do not state that the pooled OLS SEs for $\hat{\alpha}, \hat{\beta}$, and $\hat{\delta}$ computed at stage 3 should be discarded. They make no further comment about the covariance matrix computed at step 3, so I was left to assume that they meant for pooled OLS estimation of equation (7) to be what one would normally understand pooled OLS to be.

The remainder of my paper explored this ''pooled OLS estimator.'' My first derivation proved that the estimator of β computed at step 3 is identical to the least squares dummy variable (LSDV) estimator at step 1. P&T claim this is obvious. Perhaps so, but that was not stated in PT and, in fact, proving it theoretically, though it can be done with a shorter proof than mine, does require some involved algebra. I also established that the covariance matrix computed for $\hat{\beta}$ at step 3 is smaller than the appropriate matrix that would be computed at step 1. This is not a matter of degrees of freedom. There is hint of an issue with degrees of freedom noted above—the pooled OLS estimator appears to be based on $NT - 1 - K - M - 1$ degrees of freedom. But, this neglects the fact that the dummy variable coefficients have been computed at step 1, so at a minimum, the degrees of freedom should be reduced by $N - 1$. But, this correction does not redeem the matrix computation based on the pooled OLS estimator in equation (7). (Nor, is it necessarily what PT had in mind for ''correcting the degrees of freedom''—that is not made clear in PT.) The matrix, itself, is smaller than its appropriate counterpart, by construction. The appropriate matrix actually appears in PT's equation (8), but that is not the matrix that is produced by pooled OLS estimation of equation (7).[1] The matrix produced by equation (7) is, by construction, always smaller, possibly far smaller than the one in equation (8), as I demonstrated in a numerical example in my paper. It was in this sense that I argued that the efficiency gains are ''illusory''—the FEVD estimator of β at step 3 seems to be more efficient than it really is because of this algebraic result. The standard errors computed for $\hat{\beta}$ at step 3 are too small.

That leaves the FEVD estimator of γ. I also proved that the estimator of γ computed at step 2 is algebraically identical to that at step 3, so there is an ambiguity as to which one should use for the covariance matrix of the estimator of γ, step 2 or step 3. Here, it is unclear what is ''right,'' but it was also clear that the result at step 3 was not ''correct.'' I suggested two possibilities to resolve the ambiguity, one based on a known result that appears in my textbook and others and a second, ''robust'' covariance matrix along the lines of the White estimator.

My final conclusion based on the preceding was that *in the context of their model in equation (1) and under their assumptions that we all agree on*, step 3 of the FEVD estimator was worse than merely a waste of effort. Although it produces the right coefficient estimates, it produces the wrong SEs for the estimator of β and an ambiguous result for the SEs for the estimator of γ. My final recommendation was that researchers who desire to use this technique should not carry out step 3, but, rather should rely entirely on step 1 for β and step 2 plus a side calculation for γ. Along the way, I explored the theoretical implications for identification and estimation of a parameter, γ, on time-invariant variables in a FE model, in the presence of a complete set of ''unit effects'' represented by a set of dummy variables.

In reply to P&T, I note the following: I did not suggest that the FEVD estimator is inconsistent. Not a single word of my paper was about statistical properties of the FEVD estimator, finite sample or asymptotic. My use of the term ''illusory'' was explained above and was entirely appropriate—a practitioner who carries out the recommended computation will be misled at step 3 by the pooled OLS results. The SEs computed for $\hat{\beta}$ at step 3 are indeed ''too small'' in comparison to the long known result that even appears (after the aforementioned correction) in PT's equation (8). P&T have not refuted any of my results.

P&T assert that some of my comments are ''obsolete'' because they updated their *Stata* program in 2010. None of my discussion has anything to do with their *Stata* program (which to this day, I have still never seen). At one point, they also make reference to an ''option,'' apparently a feature of their computer program. But, a fair amount of mechanical detail, including the crucial statement about how to compute SEs, is simply omitted from PT; presumably the reader will find it in the *Stata* program (or will be asked [mistakenly] to trust the program to do the right calculations and not find it at all). The econometric methodology advocated in PT was proposed to the community by publication in this journal. *Stata* is not econometric methodology, and new econometric methods are not established or advanced simply by posting

[1] PT's equation (8) contains an error. The $\hat{\sigma}_u^2$ in that expression should be $\hat{\sigma}_\varepsilon^2$. The error is substantive; $\hat{\sigma}_u^2$ and $\hat{\sigma}_\varepsilon^2$ could differ by orders of magnitude.

https://doi.org/10.1093/pan/mpr011 Published online by Cambridge University Press

code online in a user written .ado file. Practitioners (and theorists) who are not *Stata* users have only the results in *Political Analysis* (and, now, this exchange) to inform them of the relevant theory of this estimator (such as it is). If the claims made for this estimator are to stand up to scrutiny (and reasonable people may doubt they will), and the estimator survives the contest with Hausman and Taylor (H&T) and other candidates, the analysis should be shown to the community of researchers in a more broadly viewed and accepted arena than buried in a *Stata* .ado file.

I have been shown an application of PT's newer version of xtfevd to the data I used in my paper. It does, indeed, faithfully reproduce the LSDV estimator and estimated covariance matrix for β. So, in practical terms, *at least for Stata users of this version of the program*, that part of my paper is indeed moot or merely pedagogical. For the rest of the world that does not use *Stata* (or xtfevd, whatever the current version is), I stand my ground. This leaves several issues on the table as regards estimation of γ. P&T and I still differ on the appropriate covariance matrix to compute for the step 2 estimator, and on my assertion that their step 3 estimator is incorrect. (My opinion that step 3 should not be carried out at all is unchanged.) Then, there are their claims of asymptotic and finite sample efficiency, about which I have thus far had no comment. There are grandiose claims made for the FEVD estimator in both PT and P&T. The statement: ''there cannot be any doubt that FEVD is more efficient than the fixed-effects and HT models (for time invariant, rarely changing and exogenous time-varying variables) and less biased than pooled OLS and random effects (for endogenous time-varying variables in finite samples)'' [P&T, 2–3] is optimistic in the extreme. None of these claims are defended theoretically; they are based on a few Monte Carlo studies. In fact, they are largely speculative. The H&T estimator that figures prominently in this discussion is a straw man. In PT/P&T, there is an explicit assumption that *every TIV is uncorrelated with* u_i. Part of my discussion demonstrated that these M moment equations are essential to identification of γ (and its meaning) in their model, and the entire apparatus must be viewed differently if this assumption is not met. H&T, on the other hand, explicitly assume that *some elements of* x_{it} *and some elements of* z_i *are correlated with* u_i. H&T and PT/P&T are discussing different models. Directly to the point, in H&T's own framework, the first step is LSDV, exactly the same as PT. But, if it were assumed at the outset that z_i and u_i were orthogonal, one would generally not want to follow through with H&T—the availability of the consistent estimator of all of β and the full set of orthogonality conditions for estimation of γ by least squares makes H&T's second to fourth steps moot. It is designed for a different purpose. As noted above, contrary to the grandiose claims made by P&T, there can be doubt.

https://doi.org/10.1093/pan/mpr011 Published online by Cambridge University Press