Oxford Bulletin of Economics and Statistics

WILEY

ORIGINAL ARTICLE OPEN ACCESS

Time Averaged CCE

Yousef Kaddoura[1] | Yiannis Karavias[2,3] | Joakim Westerlund[4,5]

[1]School of Business, Örebro University, Örebro, Sweden | [2]Brunel University of London, London, UK | [3]University of Birmingham, Birmingham, UK | [4]Lund University, Lund, Sweden | [5]Deakin University, Melbourne, Australia

Correspondence: Yousef Kaddoura (yousef.kaddoura@oru.se)

Received: 23 December 2025 | **Revised:** 20 July 2026 | **Accepted:** 22 July 2026

Keywords: common correlated effects | interactive fixed effects | panel data

ABSTRACT

A popular approach to interactive effects panel data models is the common correlated effects (CCE) estimator of Pesaran (Estimation and inference in large heterogeneous panels with a multifactor error structure. *Econometrica* **74**, 967–1012, 2006). The current paper proposes a modified version of this estimator that is useful in a number of cases where the original is not expected to work, such as when the number of cross-sectional units is small. The idea is to use time instead of cross-sectional averages of the observables to purge the interactive effects.

JEL Classification: C13, C33, C01

1 | Introduction

Consider the scalar panel variable $y_{i,t}$, observed for cross-sectional units $i = 1, \dots, N$ over time periods $t = 1, \dots, T$. A major advantage of such data when compared to pure time series or cross-sectional data is the ability to control for some form of unobserved heterogeneity. For a very long time it was thought that conventional fixed effects could do this job. However, recently it has become clear that such effects will not be enough as the effect of common shocks is unlikely to be the same for all cross-sectional units but is instead expected to depend on their individual exposure to such shocks (see Ditzen and Karavias, [1], for a discussion). This has prompted the development of econometric techniques for models with interactive fixed effects, in which time and cross-section unit effects enter multiplicatively instead of additively. The particular interactive effects model that we will be considering is given by

$$y_{i,t} = \mathbf{x}'_{i,t}\beta + \boldsymbol{\gamma}'_i\mathbf{f}_t + \varepsilon_{i,t}, \tag{1}$$

where $\mathbf{x}_{i,t}$ is a $p \times 1$ vector of regressors and $\varepsilon_{i,t}$ is an error term.[1] The interactive effects are here given by the product $\boldsymbol{\gamma}'_i\mathbf{f}_t$, where $\mathbf{f}_t$ is an $r \times 1$ vector of unobserved common factors and $\boldsymbol{\gamma}_i$ is a conformable vector of loadings. Because $\boldsymbol{\gamma}'_i\mathbf{f}_t = \alpha_i + \theta_t$ in the special case when $\boldsymbol{\gamma}_i = [\alpha_i, 1]'$ and $\mathbf{f}_t = [1, \theta_t]'$, interactive effects nest regular fixed effects. They are therefore more general when it comes to the types of unobserved heterogeneity that can be accommodated. The main concern from an econometric point of view is if the interactive effects are correlated with the regressors, as this would render the ordinary least squares (OLS) estimator of $\boldsymbol{\beta}$ inconsistent.

Despite their generality and detrimental effect on OLS, interactive effects models are simply estimated. The simplest approach by far is the common correlated effects (CCE) estimator of Pesaran ([2]), which consists of two steps. The first step is to “defactor” the data, which means that $y_{i,t}$ and $\mathbf{x}_{i,t}$ are regressed onto their cross-sectional averages, and residuals are formed. The second step estimates $\boldsymbol{\beta}$ by applying OLS to the first-step defactored data. CCE is not only general and computationally simple, but has also been shown to perform well in small samples (see, for example, Chudik and Pesaran, [3], and Westerlund and Urbain, [4]). It is therefore very popular. In fact, the approach is so popular that it has given rise to a separate CCE literature, which explains our interest in it.

Defactoring is often an effective way to control for interactive effects, but not always. One example of a situation in which it fails completely is when N is small. This is so because the factors are estimated using only the cross-sectional variation of the panel, which cannot work unless N is large. Another instance when defactoring is problematic is if the correlation between the interactive effects and the regressors is driven not by the factors but by the loadings, which is likely to be the case when the cross-sectional unit of observation is economically distinct, as in the case of aggregated cross-country panels. Defactoring is problematic also in the presence of structural change, because the time-varying slope coefficients in the model for the defactored data are not the same as for the raw data (see Kaddoura and Westerlund, [5]). Another issue is that defactoring makes the estimated slopes dependent on the properties of the factors (see Westerlund et al. [6]). Other examples of situations when defactoring can be problematic are in models with lagged dependent variables or when wanting to estimate the effect of regressors that only vary over time.

The present paper is motivated by the above discussion. The purpose is to develop a new estimator, henceforth referred to as "time averaged CCE (TACCE)", that retains the advantages of regular CCE yet does not require explicit defactoring. This is accomplished by modifying the first step of the CCE estimation procedure. Instead of defactoring $y_{i,t}$ and $\mathbf{x}_{i,t}$, we "deload" them by regressing onto their time averages. While reasonably straightforward, it appears that this modification of CCE has not been considered before in the literature, and there are, as already pointed out, many situations in which it is expected to be useful.

The rest of the paper is organized as follows: Section 2 introduces the model and the TACCE estimator that we will use to estimate it. Section 3 is concerned with the asymptotic properties of TACCE. It shows that the new estimator is consistent and asymptotically normal regardless of whether N is small or large, provided only that T is large enough. The small-sample accuracy of this asymptotic prediction is evaluated by means of Monte Carlo simulations in Section 4. Section 5 presents the results of an empirical illustration using as an example the Solow model of growth. Both the simulation and empirical studies are carried out in Stata using the `xttacce` command of Chen ([7]). Section 6 concludes. All proofs are provided in an online appendix.

2 | Model and Estimator

It is useful to denote by $\mathbf{y}_t = [y_{1,t}, \dots, y_{N,t}]'$ the $N \times 1$ vector of stacked cross-sectional observations on $y_{i,t}$, and by $\mathbf{X}_t = [\mathbf{x}_{1,t}, \dots, \mathbf{x}_{N,t}]'$ the corresponding $N \times p$ matrix of stacked observations on $\mathbf{x}_{i,t}$. The stacked version of (1) can be written as follows:

$$\mathbf{y}_t = \mathbf{X}_t \boldsymbol{\beta} + \boldsymbol{\Gamma} \mathbf{f}_t + \boldsymbol{\varepsilon}_t, \tag{2}$$

where $\boldsymbol{\Gamma} = [\boldsymbol{\gamma}_1, \dots, \boldsymbol{\gamma}_N]'$ is $N \times r$, while $\boldsymbol{\varepsilon}_t = [\varepsilon_{1t}, \dots, \varepsilon_{Nt}]'$ is $N \times 1$. Again, the researcher only observes $y_{i,t}$ and $\mathbf{x}_{i,t}$, and the main challenge is how to consistently estimate $\boldsymbol{\beta}$ when $\mathbf{x}_{i,t}$ is potentially correlated with $\boldsymbol{\gamma}_i' \mathbf{f}_t$ and hence endogenous. In order to allow for this possibility, following the classical approach of Mundlak ([8]) (see also Hansen and Liao, [9]), we assume that $\boldsymbol{\gamma}_i$ lives also in $\mathbf{x}_{i,t}$;

$$\mathbf{x}_{i,t} = \mathbf{G}_t' \boldsymbol{\gamma}_i + \mathbf{v}_{i,t}. \tag{3}$$

where $\mathbf{G}_t$ is an $r \times p$ matrix of common factors that may include some or indeed all of the elements of $\mathbf{f}_t$, and $\mathbf{v}_{i,t}$ is a $p \times 1$ vector of error terms. This last equation can be stacked to obtain the following model for $\mathbf{X}_t$:

$$\mathbf{X}_t = \boldsymbol{\Gamma} \mathbf{G}_t + \mathbf{V}_t, \tag{4}$$

where $\mathbf{X}_t$ and $\mathbf{V}_t = [\mathbf{v}_{1,t}, \dots, \mathbf{v}_{N,t}]'$ are $N \times p$ matrices.

The model in (2) and (4) forms the basis of our paper. It is similar to the one of Pesaran ([2]), but not identical. The main difference is that he assumes that the common component of $\mathbf{x}_{i,t}$ is not given by $\mathbf{G}_t' \boldsymbol{\gamma}_i$ but by $\boldsymbol{\Lambda}_i' \mathbf{f}_t$, where $\boldsymbol{\Lambda}_i$ is $r \times p$. Hence, while Pesaran assumes that it is the factors that drive any correlation between $y_{i,t}$ and $\mathbf{x}_{i,t}$, here we assume that it is the loadings. Which condition is most suitable will have to depend on the data at hand. However, we note that our specification is consistent with common practice in the fixed effects literature. Here the starting point is almost always cross-section unit fixed effects, as there are often good reasons for expecting such effects to be present and to correlate with $\mathbf{x}_{i,t}$. Time fixed effects are sometimes added but not always. The reason is that $y_{i,t}$ and $\mathbf{x}_{i,t}$ often co-move not because they are subject to the same shocks over time but because they originate from the same cross-sectional unit. And cross-sectional units are distinct, especially in the type of aggregate cross-country panels that we have in mind in the present paper. In the empirical illustration of Section 5, we elaborate on this point within the context of the Solow growth model.

As already pointed out, (2) and (4) are also consistent with the so-called "Mundlak device", which models unit-specific unobserved heterogeneity as a linear function of time-averaged variables. Another point worth considering is that while the CCE literature seems to have embraced Pesaran's ([2]) specification, the little empirical evidence that exists is not particularly supportive of the idea that $y_{i,t}$ and $\mathbf{x}_{i,t}$ are driven by the same set of factors (see, for example, Westerlund, [10], and Westerlund et al. [6]). Our specification allows for the possibility that the factors affecting $y_{i,t}$ and $\mathbf{x}_{i,t}$ are the same, but it does not require it. Moreover, while (2) and (4) are stated in terms of the same loading $\boldsymbol{\gamma}_i$, loadings that are specific to $y_{i,t}$ or $\mathbf{x}_{i,t}$, or both, can be permitted by simply setting the corresponding factors to zero.

Because $\boldsymbol{\gamma}_i$ enters both (2) and (4), the estimation of $\boldsymbol{\beta}$ is non-trivial. Before we write down our proposed TACCE estimator, we introduce the $N \times (p+1)$ matrix $\mathbf{Z}_t = [\mathbf{y}_t, \mathbf{X}_t]$, which by (2) and (4) has the following common factor representation:

$$\mathbf{Z}_t = \boldsymbol{\Gamma} \mathbf{C}_t + \mathbf{U}_t, \tag{5}$$

where $\mathbf{C}_t = [\mathbf{f}_t + \mathbf{G}_t \boldsymbol{\beta}, \mathbf{G}_t]$ and $\mathbf{U}_t = [\boldsymbol{\varepsilon}_t + \mathbf{V}_t \boldsymbol{\beta}, \mathbf{V}_t]$ are $r \times (p+1)$ and $N \times (p+1)$, respectively. The proposed TACCE estimator $\widehat{\boldsymbol{\Gamma}}$ of the space spanned by $\boldsymbol{\Gamma}$ is given by

$$\widehat{\boldsymbol{\Gamma}} = \overline{\mathbf{Z}} = \boldsymbol{\Gamma} \overline{\mathbf{C}} + \overline{\mathbf{U}}, \tag{6}$$

where $\overline{\mathbf{A}} = T^{-1} \sum_{t=1}^{T} \mathbf{A}_t$ is the time average of any generic matrix-valued variable $\mathbf{A}_t$. Note that because $\overline{\mathbf{U}} \to_p \mathbf{0}_{N \times (p+1)}$ as $T \to \infty$ for fixed N under general conditions on $\mathbf{U}_t$, we can show that $\widehat{\boldsymbol{\Gamma}} \to_p \boldsymbol{\Gamma}\overline{\mathbf{C}}$. We say that $\widehat{\boldsymbol{\Gamma}}$ is "rotationally consistent" for $\boldsymbol{\Gamma}$. The idea now is to deload $\mathbf{X}_t$ and $\mathbf{y}_t$ using $\widehat{\boldsymbol{\Gamma}}$. Because $\widehat{\boldsymbol{\Gamma}}$ is rotationally consistent, the deloading should asymptotically eliminate the interactive effects, so that $\boldsymbol{\beta}$ can be estimated by simply applying OLS to the deloaded data. The TACCE estimator of $\boldsymbol{\beta}$ is therefore given by

$$\widehat{\boldsymbol{\beta}}_{\text{TACCE}} = \left(\sum_{t=1}^{T} \mathbf{X}_t' \mathbf{M}_{\widehat{\boldsymbol{\Gamma}}} \mathbf{X}_t \right)^{-1} \sum_{t=1}^{T} \mathbf{X}_t' \mathbf{M}_{\widehat{\boldsymbol{\Gamma}}} \mathbf{y}_t, \tag{7}$$

where $\mathbf{M}_{\mathbf{A}} = \mathbf{I}_N - \mathbf{A}(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'$ for any N-rowed matrix $\mathbf{A}$.[2]

3 | Asymptotic Analysis

3.1 | Assumptions

The conditions we will be working under in this section are given in Assumptions 1–3 below. Here and throughout, $\text{tr}(\mathbf{A})$, $\text{rank}(\mathbf{A})$ and $\|\mathbf{A}\| = \sqrt{\text{tr}(\mathbf{A}'\mathbf{A})}$ denote the trace, the rank, and the Frobenius (Euclidean) norm of the matrix $\mathbf{A}$, respectively.

Assumption 1. $\varepsilon_{i,t}$ and $\mathbf{v}_{i,t}$ are covariance stationary processes that are independent over i with absolutely summable autocovariances, zero mean, $N^{-1}\mathbb{E}(\mathbf{V}_t'\mathbf{V}_t)$ positive definite, and finite fourth-order cumulants. They are also independent of one another.

Assumption 2. $\boldsymbol{\gamma}_i$, $\mathbf{G}_t$ and $\mathbf{f}_t$ are non-random such that $\sup_i \|\boldsymbol{\gamma}_i\| < \infty$, $\sup_t \|\mathbf{G}_t\| < \infty$, $\sup_t \|\mathbf{f}_t\| < \infty$ and $\text{rank}(\overline{\mathbf{C}}) = r \leq p+1 < N$ for all T.

Assumption 3. $\widehat{\boldsymbol{\Phi}} = (NT)^{-1}\sum_{t=1}^{T} \mathbf{X}_t' \mathbf{M}_{\widehat{\boldsymbol{\Gamma}}} \mathbf{X}_t$ and $N^{-1}\boldsymbol{\Gamma}'\boldsymbol{\Gamma}$ are positive definite for all N and T.

Assumption 1 is standard in the literature. We therefore do not elaborate on it. We do note, however, that the cross-sectional independence assumption is only needed when N is large, and that we place no restrictions on the cross-sectional properties of $\varepsilon_{i,t}$ and $\mathbf{v}_{i,t}$ if N is fixed. This last allowance is partly expected because the fixed-N case is the complete opposite of the fixed-T case considered by Westerlund et al. ([6]), and in that case there is no need to restrict the temporal dependence of the errors.

Assumption 2 does for TACCE what Assumptions 1 and 3, and condition (21) in Pesaran ([2]) do for regular CCE. One difference here is that while Pesaran assumes that loadings and factors follow certain probability laws, here we follow, for example, Bai ([11]), and Moon and Weidner ([12]), and treat these parameters as non-random, which is a more general consideration in the sense that it is nonparametric. Random loadings can still be accommodated by interpreting $\boldsymbol{\gamma}_i$ as a single realization from the underlying random loading process, and the same applies to random factors. Under this interpretation, Assumption 2 is more restrictive than existing CCE studies, which typically assume that factors and loadings are drawn from distributions with a certain number of bounded moments (see Pesaran, [2]). The bounded support condition in Assumption 2 is tantamount to requiring that all the moments of both factors and loadings are finite. This condition is not necessary, but it is used in the proofs. It can be relaxed at the cost of additional high-level conditions. In Section 4, we use Monte Carlo simulations as a means to evaluate the small-sample behaviour of TACCE under random factors and loadings. We consider both stationary and non-stationary factors.[3] The results confirm that Assumption 2 is stronger than needed.

The condition that $\overline{\mathbf{C}}$ has full row rank (Assumption 2) ensures that the r columns of $\boldsymbol{\Gamma}$ can be recovered asymptotically from the $p+1$ columns of $\widehat{\boldsymbol{\Gamma}}$, and is analogous to condition (21) in Pesaran ([2]). It is a mild restriction in the sense that it only requires that the number of loadings is not underspecified. This contrasts with much of the existing literature, which supposes that the number of interactive effects is either known or can be accurately estimated (see Moon and Weidner, [12], for a discussion).[4]

Assumption 3 is similar to Assumption 5 in Pesaran ([2]) in the sense that it rules out "low-rank" regressors, such as constants or dummy variables, as it is almost always done in models with interactive effects (see Moon and Weidner, [12]). The difference is that while Pesaran's ([2]) Assumption 5 requires that the defactored regressors have sufficient variation, here it is the deloaded regressors that must have enough variation. If there are low-rank regressors present in (2), then these should be treated as observed loadings that can be appended to $\widehat{\boldsymbol{\Gamma}}$. Hence, while we cannot estimate their coefficients, we can still allow for low-rank regressors.

3.2 | Results

Before reporting the asymptotic distribution of the new estimator, we define the following matrices:

$$\boldsymbol{\Omega} = \lim \frac{1}{NT} \sum_{t=1}^{T} \mathbb{E}(\mathbf{V}_t' \mathbf{M}_{\boldsymbol{\Gamma}^0} \boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t' \mathbf{M}_{\boldsymbol{\Gamma}^0} \mathbf{V}_t | C), \tag{8}$$

$$\boldsymbol{\Phi} = \lim \frac{1}{NT} \sum_{t=1}^{T} \mathbb{E}(\mathbf{V}_t' \mathbf{M}_{\boldsymbol{\Gamma}^0} \mathbf{V}_t | C), \tag{9}$$

where C is the sigma-algebra generated by $\boldsymbol{\Gamma}^0$, which we now describe. Assumption 2 supposes that $r \leq p+1$, so that the number of loadings is not larger than the number of observables. The definition of $\boldsymbol{\Gamma}^0$ depends on whether $r = p+1$ or $r < p+1$. If $r = p+1$, then $\boldsymbol{\Gamma}^0 = \boldsymbol{\Gamma}$, whereas if $r < p+1$, then $\boldsymbol{\Gamma}^0 = [\boldsymbol{\Gamma}, \sqrt{T}\overline{\mathbf{U}}\overline{\mathbf{B}}_{-r}]$, where $\overline{\mathbf{B}}_{-r}$ is a certain $(p+1) \times (p+1-r)$ matrix defined in the online appendix. The limit arguments in $\boldsymbol{\Omega}$ and $\boldsymbol{\Phi}$ are not written out, as they apply in general under the conditions of Theorem 1.

Theorem 1. *Under Assumptions 1–3, as $T \to \infty$ with $N < \infty$ or $N,\ T \to \infty$ with $N/T \to 0$,*

$$\sqrt{NT}\left(\widehat{\boldsymbol{\beta}}_{\text{TACCE}} - \boldsymbol{\beta}\right) \to_d MN\left(\mathbf{0}_{p\times 1}, \boldsymbol{\Phi}^{-1}\boldsymbol{\Omega}\boldsymbol{\Phi}^{-1}\right),$$

where $MN(\cdot,\cdot)$ signifies a mixed normal distribution that is here conditional on C.

Allowing N to be finite is an advantage in the sense that this is the only relevant scenario in practice. But then we do not want to assume that $N < \infty$ because if we did, the results would be silent about the effect of increasing N. Theorem 1 is motivated by these considerations. It allows N to be fixed without requiring it.

Suppose for sake of argument that $r = p + 1$, so that $\boldsymbol{\Gamma}^0 = \boldsymbol{\Gamma}$. In the fixed-$T$ case considered by Westerlund et al. ([6]), the asymptotic covariance matrix of regular CCE has the same basic structure as in Theorem 1. The main difference is that projections (and conditioning) are taken not with respect to $\boldsymbol{\Gamma}$ but with respect to the factors. Hence, unlike TACCE, the asymptotic distribution of CCE depends on the factors, and the literature has suggested that the properties of these can vary considerably (see Westerlund, [10], for a discussion). Of course, as pointed out in Section 3.1, this does not mean that the present paper is assumption-free when it comes to the factors; however, it does mean that if the conditions of Theorem 1 are met, the asymptotic distribution of TACCE is independent of the factors.

Inference based on Theorem 1 requires consistent estimators of $\boldsymbol{\Phi}$ and $\boldsymbol{\Omega}$. The first matrix can be estimated using $\widehat{\boldsymbol{\Phi}}$ in Assumption 3. The second matrix can be estimated analogously to Westerlund et al. ([6]) as

$$\widehat{\boldsymbol{\Omega}} = \frac{1}{NT}\sum_{t=1}^{T}\mathbf{X}_t'\mathbf{M}_{\widehat{\boldsymbol{\Gamma}}}\widehat{\varepsilon}_t\widehat{\varepsilon}_t'\mathbf{M}_{\widehat{\boldsymbol{\Gamma}}}\mathbf{X}_t, \tag{10}$$

where $\widehat{\varepsilon}_t = \mathbf{M}_{\widehat{\boldsymbol{\Gamma}}}(\mathbf{y}_t - \mathbf{X}_t\widehat{\boldsymbol{\beta}}_{\mathrm{TACCE}})$.

Theorem 2. *Under Assumptions 1–3, $\widehat{\boldsymbol{\Phi}} \to_p \boldsymbol{\Phi}$ and $\widehat{\boldsymbol{\Omega}} \to_p \boldsymbol{\Omega}$ as $T \to \infty$ with $N < \infty$ or $N,\ T \to \infty$ with $N/T \to 0$.*

According to Theorem 2, although the definitions of $\boldsymbol{\Phi}$ and $\boldsymbol{\Omega}$ depend on whether $r = p + 1$ or $r < p + 1$, we can estimate them consistently in a way that does not require knowledge of r. In order appreciate the implications of this for inference, consider testing the null hypothesis of $H_0 : \mathbf{R}\boldsymbol{\beta} = \mathbf{q}$, where $\mathbf{R}$ is a $s \times p$ matrix with $\mathrm{rank}(\mathbf{R}) = s \le p$ and $\mathbf{q}$ is a $s \times 1$ vector. This test can be carried out using the following Wald statistic:

$$W = NT(\mathbf{R}\widehat{\boldsymbol{\beta}}_{\mathrm{TACCE}} - \mathbf{q})'(\mathbf{R}\widehat{\boldsymbol{\Phi}}^{-1}\widehat{\boldsymbol{\Omega}}\widehat{\boldsymbol{\Phi}}^{-1}\mathbf{R}')^{-1}(\mathbf{R}\widehat{\boldsymbol{\beta}}_{\mathrm{TACCE}} - \mathbf{q}). \tag{11}$$

Theorems 1 and 2 imply that $W \to \chi^2(s)$ under H_0. If $s = 1$, the following t-statistic can be used:

$$t = \frac{\sqrt{NT}\left(\mathbf{R}\widehat{\boldsymbol{\beta}}_{\mathrm{TACCE}} - \mathbf{q}\right)}{\sqrt{\mathbf{R}\widehat{\boldsymbol{\Phi}}^{-1}\widehat{\boldsymbol{\Omega}}\widehat{\boldsymbol{\Phi}}^{-1}\mathbf{R}'}}, \tag{12}$$

which is asymptotically $N(0, 1)$ under the same null. It is important to note that while the asymptotic distribution of $\sqrt{NT}(\widehat{\boldsymbol{\beta}}_{\mathrm{TACCE}} - \boldsymbol{\beta})$ is dependent on C, the above-stated $\chi^2(s)$ and $N(0, 1)$ distributions are not. TACCE therefore enables standard inference regardless of whether $r = p + 1$ or $r < p + 1$.

4 | Monte Carlo Study

In this section, we present the results of a small-scale Monte Carlo study aimed at examining the small-sample performance of TACCE. The data-generating process used for this purpose is given by a restricted version of (2) and (4) that sets $r = 1$, $p = 2$, and $\boldsymbol{\beta} = [1, 2]'$. The elements of $\varepsilon_{i,t}$ and $\mathbf{v}_{i,t}$ are drawn from $N(0, 1)$. The loadings are generated as $\boldsymbol{\gamma}_i \sim N(1, 0.2)$. The factors are generated as

$$\mathbf{f}_t = 0.3\mathbf{f}_{t-1} + \sqrt{1 - 0.3^2}\, u_t, \tag{13}$$

$$G_{j,t} = c_j + (0.1j) \cdot G_{j,t-1} + \sqrt{1 - (0.1j)^2}\, w_{j,t}, \tag{14}$$

where $G_{j,t}$ is the j-th element of $\mathbf{G}_t$, which is 1×2 in this section, the intercept c_j equals 1 if $j = 1$ and 0 if $j = 2$, and u_t, $w_{1,t}$ and $w_{2,t}$ are all drawn from $N(0, 1)$. Initial values are set to 0. We make 10,000 replications of panels of size $T \in \{30, 50, 100, 200\}$ and $N \in \{5, 10, 20, 50, 100, 200\}$. Hence, while T is in the same range as in the previous CCE literature (see, for example, Pesaran, [2]), most of the values of N that we consider are intentionally much smaller than in this literature, although we also consider some larger values.

Table 1 reports the average bias and root mean squared error (RMSE) of the TACCE estimator of the first element of $\boldsymbol{\beta}$, β_1 say (both bias and RMSE are $\times 100$). The table also reports the size and power of a nominal 5% level t-test for testing $H_0 : \beta_1 = 1$ versus $H_1 : \beta_1 \neq 1$, where the power results are based on setting $\beta_1 = 0.9$. The first thing to note is that the estimator seems to perform well in small samples, even for N (T) as small as 5 (30). Power can be low in the smallest panels, but it increases very quickly as N and T grow. Unreported results confirm that power is increasing also in the deviation from the null hypothesis, as measured by $|\beta_1 - 1|$. Size accuracy is not perfect, but it is acceptable. RMSE is decreasing not only in T but also in N, as expected given Theorem 1.

The overall picture is that TACCE inherits the good small-sample performance of CCE. In the interest of comparison, Table 2 presents the results for the regular CCE estimator under the same data-generating process as in Table 1. The performance is generally close to that of TACCE. The main exception is when N is small, in which case TACCE performs markedly better, as to be expected given that it does not require N to be large.

Assumption 2 requires that the factors are bounded, which, as pointed out in Section 3.1, rules out factors that are non-stationary. Another assumption is that $\boldsymbol{\beta}$ is time and cross-section invariant. However, these conditions are not necessary and are there mainly to simplify the proofs. This is clear from Tables 3 and 4, which report some results for the case when $\mathbf{f}_t$ is either a linear trend or unit root non-stationary. While obviously not identical, the results are very close to those reported in Table 1, which is partly expected given that the asymptotic distribution of TACCE does not depend on the factors. In Table 5, we consider the case when the slope coefficients vary randomly over time. While the performance is not as good as in Table 1, the difference is never very large, and there are clear improvements as N and especially T increase.

TABLE 1 | Main results: TACCE.

N	T	Bias	RMSE	Size	Power	N	T	Bias	RMSE	Size	Power
5	30	−0.081	13.587	0.084	0.166	50	30	−0.048	2.764	0.072	0.953
	50	0.108	10.223	0.069	0.190		50	0.004	2.099	0.060	0.997
	100	−0.074	7.272	0.063	0.317		100	−0.008	1.468	0.056	1.000
	200	0.039	5.081	0.055	0.526		200	0.006	1.036	0.050	1.000
10	30	0.213	7.041	0.068	0.322	100	30	0.004	2.002	0.075	0.999
	50	0.195	5.405	0.063	0.464		50	0.035	1.458	0.059	1.000
	100	−0.024	3.831	0.059	0.756		100	0.007	1.033	0.056	1.000
	200	−0.056	2.690	0.051	0.962		200	0.011	0.720	0.056	1.000
20	30	0.071	4.615	0.074	0.610	200	30	−0.030	1.435	0.069	1.000
	50	−0.023	3.465	0.061	0.831		50	−0.009	1.031	0.056	1.000
	100	−0.034	2.438	0.054	0.985		100	0.007	0.719	0.058	1.000
	200	−0.058	1.726	0.054	1.000		200	−0.013	0.505	0.052	1.000

Note: The bias and RMSE results are for the TACCE estimator of the first element of $\boldsymbol{\beta}$, β_1 say. Both are multiplied by 100. The size and power results are for a nominal 5% level t-test for testing $H_0 : \beta_1 = 1$ versus $H_1 : \beta_1 = 0.9 \neq 1$.

TABLE 2 | Comparison: Regular CCE.

N	T	Bias	RMSE	Size	Power	N	T	Bias	RMSE	Size	Power
5	30	0.102	13.720	0.246	0.362	50	30	0.024	2.765	0.072	0.960
	50	0.144	10.184	0.226	0.418		50	0.012	2.082	0.056	0.999
	100	0.127	7.261	0.223	0.564		100	0.001	1.462	0.057	1.000
	200	−0.100	5.045	0.222	0.788		200	0.008	1.037	0.058	1.000
10	30	−0.025	7.102	0.113	0.428	100	30	0.017	1.975	0.067	1.000
	50	0.066	5.305	0.103	0.591		50	−0.001	1.446	0.057	1.000
	100	−0.028	3.736	0.097	0.847		100	−0.002	1.027	0.055	1.000
	200	−0.009	2.690	0.102	0.980		200	0.029	0.714	0.052	1.000
20	30	0.026	4.617	0.084	0.674	200	30	−0.019	1.443	0.081	1.000
	50	−0.032	3.467	0.075	0.868		50	0.004	1.049	0.063	1.000
	100	0.016	2.463	0.073	0.990		100	−0.003	0.725	0.056	1.000
	200	0.022	1.745	0.077	1.000		200	−0.001	0.504	0.051	1.000

Note: The bias and RMSE results are for the regular CCE estimator of the first element of $\boldsymbol{\beta}$, β_1 say. Both are multiplied by 100. The size and power results are for a nominal 5% level t-test for testing $H_0 : \beta_1 = 1$ versus $H_1 : \beta_1 = 0.9 \neq 1$.

TABLE 3 | Robustness: Unit root factors.

N	T	Bias	RMSE	Size	Power	N	T	Bias	RMSE	Size	Power
5	30	−0.212	13.738	0.084	0.165	50	30	−0.032	2.766	0.071	0.952
	50	0.089	10.336	0.070	0.198		50	0.009	2.149	0.064	0.997
	100	0.047	7.167	0.060	0.294		100	−0.012	1.456	0.050	1.000
	200	0.038	5.056	0.055	0.518		200	0.011	1.039	0.051	1.000
10	30	−0.148	7.213	0.075	0.336	100	30	0.006	1.964	0.066	0.998
	50	−0.025	5.504	0.070	0.469		50	−0.003	1.485	0.063	1.000
	100	0.071	3.800	0.051	0.741		100	−0.001	1.037	0.056	1.000
	200	0.008	2.712	0.058	0.956		200	0.012	0.733	0.052	1.000
20	30	0.048	4.682	0.075	0.608	200	30	0.043	1.424	0.064	0.999
	50	−0.004	3.537	0.060	0.821		50	0.007	1.064	0.060	1.000
	100	−0.022	2.453	0.058	0.983		100	−0.012	0.730	0.058	1.000
	200	−0.010	1.717	0.054	1.000		200	0.000	0.505	0.050	1.000

Note: The bias and RMSE results are for the TACCE estimator of the first element of $\boldsymbol{\beta}$, β_1 say. Both are multiplied by 100. The size and power results are for a nominal 5% level t-test for testing $H_0 : \beta_1 = 1$ versus $H_1 : \beta_1 = 0.9 \neq 1$. Unlike in Tables 1 and 2, now the factor is generated as $\mathbf{f}_t = \mathbf{f}_{t-1} + u_t$, where u_t is as in the text.

TABLE 4 | Robustness: Linearly trending factors.

N	**T**	**Bias**	**RMSE**	**Size**	**Power**	**N**	**T**	**Bias**	**RMSE**	**Size**	**Power**
5	30	−0.141	13.783	0.088	0.173	50	30	−0.004	2.702	0.069	0.964
	50	−0.281	10.421	0.072	0.210		50	0.041	2.103	0.061	0.998
	100	0.082	7.238	0.063	0.306		100	0.009	1.479	0.052	1.000
	200	−0.006	5.005	0.049	0.511		200	0.028	1.051	0.053	1.000
10	30	0.074	7.038	0.066	0.321	100	30	0.018	1.868	0.068	1.000
	50	−0.022	5.500	0.065	0.478		50	0.005	1.447	0.061	1.000
	100	−0.093	3.798	0.057	0.764		100	−0.018	1.033	0.058	1.000
	200	−0.042	2.704	0.053	0.965		200	0.005	0.727	0.055	1.000
20	30	−0.100	4.514	0.070	0.642	200	30	0.016	1.339	0.072	1.000
	50	0.006	3.433	0.060	0.831		50	−0.004	1.012	0.058	1.000
	100	−0.002	2.468	0.059	0.982		100	−0.002	0.713	0.053	1.000
	200	0.006	1.727	0.055	1.000		200	0.001	0.508	0.055	1.000

Note: The bias and RMSE results are for the TACCE estimator of the first element of $\boldsymbol{\beta}$, β_1 say. Both are multiplied by 100. The size and power results are for a nominal 5% level t-test for testing $H_0 : \beta_1 = 1$ versus $H_1 : \beta_1 = 0.9 \neq 1$. The factor is generated as $\mathbf{f}_t = 1 + 0.5t + u_t$, where u_t is as in the text.

TABLE 5 | Robustness: Time-varying coefficients.

N	**T**	**Bias**	**RMSE**	**Size**	**Power**	**N**	**T**	**Bias**	**RMSE**	**Size**	**Power**
5	30	0.474	14.534	0.084	0.150	50	30	0.024	4.725	0.074	0.598
	50	0.109	11.184	0.073	0.176		50	−0.026	3.588	0.058	0.810
	100	−0.014	7.741	0.059	0.264		100	0.045	2.561	0.058	0.975
	200	0.096	5.521	0.055	0.447		200	0.025	1.778	0.052	1.000
10	30	−0.024	8.274	0.074	0.280	100	30	0.039	4.185	0.073	0.694
	50	0.023	6.274	0.061	0.368		50	−0.050	3.192	0.059	0.883
	100	−0.004	4.438	0.055	0.615		100	0.025	2.311	0.062	0.990
	200	−0.023	3.123	0.050	0.889		200	0.025	1.621	0.059	1.000
20	30	−0.016	6.064	0.076	0.428	200	30	−0.006	3.933	0.070	0.751
	50	−0.018	4.575	0.060	0.606		50	0.044	3.054	0.066	0.910
	100	0.011	3.272	0.054	0.867		100	0.047	2.130	0.057	0.997
	200	0.002	2.324	0.056	0.990		200	−0.004	1.516	0.055	1.000

Note: The bias and RMSE results are for the TACCE estimator of the first element of $\boldsymbol{\beta}$, β_1 say. Both are multiplied by 100. The size and power results are for a nominal 5% level t-test for testing $H_0 : \beta_1 = 1$ versus $H_1 : \beta_1 = 0.9 \neq 1$. The slopes vary over time as follows: $\boldsymbol{\beta}_t = \boldsymbol{\beta} + \mathbf{v}_t$, where $\mathbf{v}_t \sim N(0, 0.04)$. This specification is analogous to the one employed by Pesaran ([2]) for the case of slope heterogeneity over the cross-sectional units.

5 | Empirical Illustration

5.1 | Motivation

We illustrate TACCE by estimating a Solow-type growth equation for the G7 countries (Canada, France, Germany, Italy, Japan, the United Kingdom, and the United States) using annual data from Penn World Table 11.0 (Feenstra et al. [14]).[5] We choose the largest sample period available to us, which is 1951–2023. We therefore have a balanced panel dataset comprised of $N = 7$ countries and $T = 73$ years.

With $N = 7$, CCE cannot be expected to work well, because the cross-sectional averages employed in the defactoring are too noisy to serve as reliable proxies for the common factors. Indeed, the validity of the cross-sectional averaging approach rests on a large-N argument that does not apply here. TACCE, by contrast, deloads using time averages, which are informative about the country-specific loadings even when N is small, provided that T is large enough.

The TACCE model specification in (1) and (3) with loadings as drivers of unobserved heterogeneity is also economically natural here. The logic runs as follows: In the Solow framework, long-run income differences across countries are driven by deep, slow-moving determinants that are largely country-specific. Institutions, the security of property rights, the quality of governance, and accumulated technological capacity all fall into this category. A large body of empirical work in the growth literature has established that these same income determinants also govern a country's saving and investment behaviour, which are standard Solow regressors. Hall and Jones ([15]) show that what they call "social infrastructure", meaning institutions and government policies, is a fundamental driver of differences in both capital accumulation and productivity across countries. Knack and Keefer ([16]) find that the security of property rights and the enforceability of contracts have a significant positive effect on investment rates, and Mauro ([17]) documents that institutional quality also affects economic performance through investment. The implication is that the same country-specific institutional characteristics that determine long-run income levels also shape the regressors in a Solow model. These are precisely the characteristics captured by the loadings in (1) and (3). The present paper is not the first to make this point (see, for example, Eberhardt and Teal, [18]). However, it is to the best of our knowledge the first to allow for multiple loadings that drive the endogeneity of the regressors.

5.2 | Model and Results

In terms of the notation of Section 1, the model that we will be considering in this section is (1) with $y_{i,t} = \log(Y_{i,t}/L_{i,t})$, where $Y_{i,t}$ denotes real GDP at constant national prices for country i and year t, and $L_{i,t}$ is employment measured as the number of persons engaged. The vector of regressors is given by $\mathbf{x}_{i,t} = [\log(s_{i,t}),\ \log(n_{i,t} + g + \delta),\ \Delta\log(hc_{i,t})]'$, where $s_{i,t}$ is the share of output invested in physical capital, $n_{i,t}$ is the rate of population growth, $g + \delta = 0.05$ following Mankiw et al. ([19]), and $hc_{i,t}$ is the Penn World Table's human capital index.[6] Note that $\Delta\log(hc_{i,t})$ should be interpreted as an empirical proxy rather than as an exact human-capital object (see, e.g., Mankiw et al. [19], and Islam, [20]).

The vector of slope coefficients is in this illustration given by $\boldsymbol{\beta} = [\beta_1, \beta_2, \beta_3]'$. The basic Solow model predicts $\beta_1 > 0$ and $\beta_2 < 0$ since higher investment raises the steady-state capital stock per worker while faster population growth dilutes it. The model also implies symmetry in the sense that $\beta_1 + \beta_2 = 0$, which is a direct consequence of the Cobb-Douglas production structure.[7] Mankiw et al. ([19]) use it as their main test of the theory, and we follow them in doing so here.

Table 6 reports the results. Pooled OLS controls for neither loadings nor factors, and we see that the estimated coefficient on $\log(s_{i,t})$ does not even have the expected sign. Fixed effects improve matters, but the estimated coefficient on $\log(n_{i,t} + g + \delta)$ is unreasonably large (in absolute value). Eberhardt and Teal ([18]) point out that fixed effects are generally not enough to control for interactive effects, and that this is likely to lead to bias. Consistent with this, we see that the CCE estimates of the effect of population growth are much smaller in absolute value. However, nothing is significant, and population growth enters with an estimated negative sign, which is partly expected given the smallness of N in this illustration. By contrast, the TACCE estimates have their expected signs, and they are all significant. We also tested the symmetry restriction that $\beta_1 + \beta_2 = 0$ using the TACCE-based Wald test described in Section 3. The resulting p-value is 0.450 (0.992) in the specification with $\Delta\log(hc_{i,t})$ excluded (included), which we interpret as providing strong evidence in favour of symmetry.

The TACCE estimates are broadly in line with the existing literature. Mankiw et al. ([19]) apply the textbook Solow model to a panel of OECD countries. Their results imply a physical-capital share of about one third once human capital is properly accounted for. Under Cobb-Douglas production technology, the physical-capital share is given by $\beta_1/(1 + \beta_1)$. The TACCE estimates reported in Table 6 lead to physical-capital share estimates of between 0.21 and 0.26, which is lower than the one-third benchmark but within the range found in panel studies that account for unobserved heterogeneity (see, for example, Islam, [20]).

TABLE 6 | Empirical results.

Variable	POLS	FE	CCE	TACCE
Basic Solow specification				
$\log(s_{i,t})$	−0.892 ***	0.112	0.030	0.270 ***
	(−5.03)	(0.68)	(0.23)	(7.12)
$\log(n_{i,t} + g + \delta)$	−3.446 ***	−4.324 ***	0.249	−0.215 ***
	(−42.72)	(−15.91)	(1.60)	(−4.34)
Augmented Solow specification				
$\log(s_{i,t})$	−0.835 ***	0.183	0.059	0.358 ***
	(−4.89)	(1.28)	(0.45)	(10.52)
$\log(n_{i,t} + g + \delta)$	−3.540 ***	−4.102 ***	0.124	−0.358 ***
	(−39.04)	(−15.11)	(0.78)	(−6.03)
$\Delta\log(hc_{i,t})$	−30.777 ***	−29.058 ***	6.178	5.308 ***
	(−3.13)	(−3.46)	(1.10)	(5.18)

Note: The dependent variable is the log of real GDP per worker. Entries in parentheses are asymptotic t-statistics based on the variance estimator described in Section 3. "POLS" and "FE" refer to pooled OLS without fixed effects and the within fixed effects OLS estimator, respectively, "CCE" refers to the regular CCE estimator, and "TACCE" refers to the proposed estimator. ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively.

Islam ([20]) finds that the effect of human capital is often insignificant or even of the wrong sign in his panel regressions. He attributes this to the difficulty of measuring the temporal dimension of human capital, and to the fact that within-country changes in schooling proxies may not track true within-country changes in productive human capital. The TACCE results tell a different story. The coefficient on $\Delta \log(hc_{i,t})$ is positive and strongly significant, suggesting that controlling for interactive effects through deloading overcomes the difficulty of estimating the effect of human capital.

All in all, we find that while the TACCE results are consistent with the Solow equation, the results obtained based on pooled and fixed effects OLS, and regular CCE are not. One possible explanation for this difference in the results is that fixed effects are not enough to account for all of the unobserved heterogeneity, and that N is too small and that the cross-sectional heterogeneity captured by the loadings is too substantial for reliable CCE estimation.

6 | Conclusion

The CCE approach of Pesaran ([2]) has attracted considerable interest in recent years, so much so that it has given rise to a separate literature. However, while certainly popular with great empirical appeal, there are situations in which the defactoring performed within CCE is not expected to work well. The present paper proposes a solution based on deloading the data instead.

Funding

This work was supported by the Knut och Alice Wallenberg Foundation through a Wallenberg Academy Fellowship.

Endnotes

[1] While standard in the literature, the common slope condition implicit in (1) can be relaxed by requiring instead that the slope coefficients are randomly distributed around a constant mean (see Pesaran, [2]). In the Monte Carlo study of Section 4, we report some results for the case when the slopes vary randomly over time.

[2] The TACCE estimator considered in the present paper is based on regular pooling, whereby the data are summed over both time periods and cross-sectional units before taking the ratio. Another approach is to use what is often referred to as "group mean" estimation, but that is here more appropriately referred to as "regime mean" estimation, which means that the ratio of cross-sectional sums is taken prior to summing over time. Pooling does, however, have some optimality properties, which is often reflected in superior small sample performance (see Pesaran, [2]). This last observation is confirmed by our unreported Monte Carlo results.

[3] Should the factors be unit root non-stationary, the assumed stationarity of $\epsilon_{i,t}$ is tantamount to requiring that $y_{i,t}$, $\mathbf{x}_{i,t}$ and $\mathbf{f}_t$ are cointegrated with cointegrating vector $(1, -\boldsymbol{\beta}', -\boldsymbol{\gamma}_i')'$ (see Westerlund, [10]).

[4] While not strictly necessary, appropriate specification of r is likely to lead to efficiency gains when compared to overspecification. If efficiency is a concern, information criteria analogous to those considered by Margaritella and Westerlund ([13]) for regular CCE may be used. These will be inconsistent when N is fixed and only T is large, but they will be consistent when both N and T are large.

[5] The Penn World Table 11.0 is available online at https://www.rug.nl/ggdc/productivity/pwt/?lang=en.

[6] Lower bounds of 0.01 and 0.1 are imposed on the investment share and the human capital index, respectively, before taking logs. Population growth is lower-bounded at −0.04 before forming $\log(n_{i,t} + g + \delta)$.

[7] Under the Cobb-Douglas production function $F(K, AL) = K^{\alpha}(AL)^{1-\alpha}$, where $\alpha \in (0, 1)$ is the capital share of production, the coefficients on $\ln(s_{i,t})$ and $\ln(n_{i,t} + g + \delta)$ are $\beta_1 = \alpha/(1-\alpha)$ and $\beta_2 = -\alpha/(1-\alpha)$, respectively. Hence, under this production technology, $\beta_1 + \beta_2 = 0$.

References

1. J. Ditzen and Y. Karavias, "Interactive, Grouped and Non-Separable Fixed Effects: A Practitioner's Guide to the New Panel Data Econometrics," 2025 SSRN 5365941.

2. M. H. Pesaran, "Estimation and Inference in Large Heterogeneous Panels With a Multifactor Error Structure," *Econometrica* 74 (2006): 967–1012.

3. A. Chudik and M. H. Pesaran, "Common Correlated Effects Estimation of Heterogeneous Dynamic Panel Data Models With Weakly Exogenous Regressors," *Journal of Econometrics* 188 (2015): 393–420.

4. J. Westerlund and J.-P. Urbain, "Cross-Sectional Averages Versus Principal Components," *Journal of Econometrics* 185 (2015): 372–377.

5. Y. Kaddoura and J. Westerlund, "Generally Shifting Coefficients in Interactive Effects Panel Data Models and the COVID–19 Regime," (2025). Unpublished manuscript.

6. J. Westerlund, Y. Petrova, and M. Norkute, "CCE in Fixed-*T* Panels," *Journal of Applied Econometrics* 34 (2019): 746–761.

7. P. Chen, *XTTACCE: Stata Module to Compute Time Averaged Common Correlated Effects Estimator (TACCE) for Fixed-N Panel Data and SUR*. Statistical Software Components S459473 (Boston College Department of Economics, 2025).

8. Y. Mundlak, "On the Pooling of Time Series and Cross Section Data," *Econometrica* 46 (1978): 69–85.

9. C. Hansen and Y. Liao, "The Factor-Lasso and *K*-Step Bootstrap Approach for Inference in High-Dimensional Economic Applications," *Econometric Theory* 35 (2019): 465–509.

10. J. Westerlund, "CCE in Panels With General Unknown Factors," *Econometrics Journal* 21 (2018): 264–276.

11. J. Bai, "Panel Data Models With Interactive Fixed Effects," *Econometrica* 77 (2009): 1229–1279.

12. H. R. Moon and M. Weidner, "Linear Regression for Panel With Unknown Number of Factors as Interactive Fixed Effects," *Econometrica* 83 (2015): 1543–1579.

13. L. Margaritella and J. Westerlund, "Using Information Criteria to Select Averages in CCE," *Econometrics Journal* 26 (2023): 405–421.

14. R. C. Feenstra, R. Inklaar, and M. P. Timmer, "The Next Generation of the Penn World Table," *American Economic Review* 105 (2015): 3150–3182.

15. R. E. Hall and C. I. Jones, "Why Do Some Countries Produce So Much More Output Per Worker Than Others?," *Quarterly Journal of Economics* 114 (1999): 83–116.

16. S. Knack and P. Keefer, "Institutions and Economic Performance: Cross-Country Tests Using Alternative Institutional Measures," *Economics and Politics* 7 (1995): 207–227.

17. P. Mauro, "Corruption and Growth," *Quarterly Journal of Economics* 110 (1995): 681–712.

18. M. Eberhardt and F. Teal, "Econometrics for Grumblers: A New Look at the Literature on Cross-Country Growth Empirics," *Journal of Economic Surveys* 25 (2011): 109–155.

19. N. G. Mankiw, D. Romer, and D. N. Weil, "A Contribution to the Empirics of Economic Growth," *Quarterly Journal of Economics* 107 (1992): 407–437.

20. N. Islam, "Growth Empirics: A Panel Data Approach," *Quarterly Journal of Economics* 110 (1995): 1127–1170.

Supporting Information

Additional supporting information can be found online in the Supporting Information section. **Data S1:** Supplementary Material.