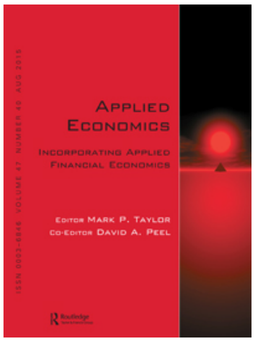

Applied Economics

ISSN: 0003-6846 (Print) 1466-4283 (Online) Journal homepage: https://www.tandfonline.com/loi/raec20

Grouped coefficients to reduce bias in heterogeneous dynamic panel models with small *T*

Nathan P. Hendricks & Aaron Smith

Applied Economics, 2015
Vol. 47, No. 40, 4335–4348, http://dx.doi.org/10.1080/00036846.2015.1029112

Routledge
Taylor & Francis Group

Grouped coefficients to reduce bias in heterogeneous dynamic panel models with small T

Nathan P. Hendricks[a,*] and Aaron Smith[b]

[a]*Department of Agricultural Economics, Kansas State University, Manhattan, KS 66506, USA*
[b]*Department of Agricultural and Resource Economics, University of California Davis, Davis, CA 95616, USA*

We propose the grouped coefficients estimator to reduce bias in dynamic panels with small T that have a multilevel structure to the coefficient and factor loading heterogeneity. If groups are chosen such that the within-group heterogeneity is small, then the grouped coefficients estimator can lead to substantial bias reduction compared to pooled GMM dynamic panel estimators. We also propose using a Wald test that can be used to assess whether pooled estimators suffer from heterogeneity bias. We illustrate the usefulness of grouped coefficients with an application to labour demand in which the coefficients are grouped by sub-sector. Our results suggest that the standard pooled estimates are substantially biased.

Keywords: dynamic panel; varying coefficient; cross-sectional dependence; multilevel; hierarchical

JEL Classification: C23; C52

I. Introduction

We study dynamic panel regression models in which the number of time series observations T is small and all coefficients may vary across individuals. We aim to estimate the average of the coefficients across individuals.[1] Pooling across individuals, and thereby ignoring the heterogeneity, generates large biases in this setting. The fact that ignoring the heterogeneity of the intercept generates large biases has been well known since the work of Nerlove (1967) and Nerlove (1971), and many papers have been written on this topic. However, heterogeneity in the other coefficients can also generate large biases and this issue has received much less attention in the literature.

There is a large literature that proposes alternative estimation methods for dynamic panel models with

*Corresponding author. E-mail: nph@ksu.edu

[1] In many cases, the average effect is of the greatest interest since it corresponds to the aggregate effect. The distribution of coefficients may also be of interest. Our estimator gives group-specific estimates of the coefficients – which is useful – but does not describe the full distribution of individual coefficients, so we focus attention on the average coefficient. Estimating the distribution of individual coefficients requires different methods and may require more stringent assumptions about the model (e.g., Hsiao *et al.*, 1999). Individual coefficients cannot be consistently estimated with small T.

small T that exhibit heterogeneity only in the intercept. Some of the alternative approaches include Generalized Method of Moments (GMM) estimators that use lags as instruments (e.g., Arellano and Bond, 1991; Arellano and Bover, 1995; Blundell and Bond, 1998), deriving bias corrections for fixed effects (FE) (e.g., Kiviet, 1995; Hahn and Kuersteiner, 2002; Bruno, 2005; Bun and Carree, 2005), maximizing a transformed likelihood function (Hsiao *et al.*, 2002) or using alternative differencing procedures (Hahn *et al.*, 2007; Han *et al.*, 2014).

The GMM estimators, in particular, have become extremely popular in empirical applications. For example, Google Scholar indicates nearly 15 000 citations of Arellano and Bond (1991) as of 2014 and the number of citations has been increasing rapidly in recent years (Wansbeek, 2012). The GMM estimators, however, are not consistent when either (i) the coefficient on the lagged dependent variable or an autocorrelated regressor is heterogeneous (Pesaran and Smith, 1995) or (ii) there is cross-sectional dependence (Sarafidis and Robertson, 2009). All of the other estimators listed above were also developed under the assumption of coefficient homogeneity. Given the popularity of the GMM estimators in applied econometrics, it is important to evaluate the magnitude of the bias due to coefficient heterogeneity and cross-sectional dependence and propose a method to reduce the bias.

We argue that, in many empirical settings, individuals can be sorted into relatively homogeneous groups based on observables. Examples include individuals who are more similar to others in their school, county or state than those outside, and firms that are more similar to those in their sector or region than those outside. We use the term *multilevel* to describe this data structure, although the words cluster and hierarchical have also been used in this context (e.g., Gelman, 2006). We make three contributions to the literature. First, we propose the grouped coefficients estimator, which entails pooling within groups to reduce bias in the estimated average effects. Second, we provide Monte Carlo evidence on the large bias of commonly used dynamic panel estimators when coefficients are heterogeneous. Third, we propose a simple test to determine whether grouping significantly reduces bias.

We implement our proposed grouped coefficients by applying dynamic panel estimators separately to groups of relatively homogeneous individuals. We then obtain an estimate of the mean effect by averaging across groups. The success of this estimator in reducing bias depends on the groups being sufficiently large and homogeneous.[2] For example, if all individuals within each group have the same coefficients, then OLS provides consistent group-level estimates and the resulting grouped coefficients estimator is consistent for the average effect. If individuals in each group have heterogeneous intercepts but otherwise homogeneous coefficients, then applying a GMM dynamic panel estimator to each group produces consistent estimates of the average effect. In general, estimating coefficients specific to each group removes the pooling bias from coefficient heterogeneity, whereas pooling within the groups reduces the bias of the coefficient on the lagged dependent variable due to small T.

A multilevel structure is not necessary for bias reduction in panels with large T. Pesaran and Smith (1995) note that, when T is large, a researcher can consistently estimate the coefficients that characterize each individual. Thus, they propose running separate regressions for each individual and then averaging over individuals to estimate the average effect.[3] Related methods for large T are proposed by Maddala *et al.* (1997) and Hsiao and Tahmiscioglu (1997). Maddala *et al.* (1997) consider a random coefficient model of dynamic panels using shrinkage estimators and Hsiao and Tahmiscioglu (1997) estimate separate random coefficient models for groups of homogeneous individuals. These methods do not work well for small T. In essence, one needs a sufficient number of observations in each pool. With small T, we propose increasing the size of the pool beyond the individual to groups of similar individuals.

Few previous articles have dealt with heterogeneous coefficients in dynamic panels with small T. Hsiao *et al.* (1999) suggest the use of a hierarchical Bayes estimator for random coefficients and find that it performs well in Monte Carlo simulations. Hsiao

[2] The relevant asymptotic theory here is one in which group size goes to infinity and T is fixed.
[3] Pesaran and Smith (1995) call their estimator the mean group estimator. This name could also describe our estimator, but we use the term grouped coefficients instead to avoid confusion with Pesaran and Smith's mean group estimator. In addition, for the remainder of this paper, we refer to the mean group estimator as individual coefficients OLS to distinguish from grouped coefficients estimators that estimate separate models for groups of individuals.

et al. (2005) apply this hierarchical Bayes estimator to estimate money demand. Cai and Li (2008) and Tran and Tsionas (2009) take an alternative approach by specifying the coefficients as nonparametric functions of observable covariates. Their specification is flexible because it produces estimates of the heterogeneous coefficients, but it requires the user to specify variables that determine the heterogeneity. In our setting, the heterogeneity is not constrained to be a function of any observables.

Other articles develop approaches to correct bias when the dynamic panel model is misspecified. Lee (2012) considers the effect of misspecifying the lag order and Zhou *et al.* (2014) consider the effect of second-order serial correlation of the error term. Although heterogeneous coefficients also induce serial correlation of the error term in pooled models, Pesaran and Smith (1995) note that the serial correlation due to heterogeneous coefficients has a highly complex form and standard methods of correcting for serial correlation are unlikely to be successful. Therefore, the methods implemented by Zhou *et al.* (2014) are unlikely to resolve the bias due to heterogeneous coefficients.

Another set of previous literature proposes bias reduction methods in dynamic panel estimators with cross-sectional dependence and small T. Sarafidis and Robertson (2009) show that time FE can reduce the bias of GMM estimators with cross-sectional dependence as long as the variance of the factor loadings is small. Phillips and Sul (2007) derive a bias correction for FE with incidental trends and common factors. Nauges and Thomas (2003) propose a double-differencing approach with GMM to remove the influence of common factors. None of these methods exploit a multilevel structure or address heterogeneous coefficients.

The grouped coefficients estimator reduces the bias due to cross-sectional dependence if the dependence is characterized by common factors with loadings that have a multilevel structure. If there is no within-group heterogeneity in the factor loadings, then the common factors can be treated as time FE within each group. The grouped coefficients estimator is especially beneficial when the overall variance of the factor loadings is large, but the variance of the factor loadings within groups is small.

We also provide Monte Carlo evidence on the large impact of heterogeneous coefficients on pooled estimators. Although the bias can be very large, our results indicate that grouped coefficients estimation can substantially reduce the bias. For example, using the Blundell–Bond (BB) GMM estimator, we find that choosing groups with just half of the coefficient variation arising within groups reduces the bias by about half in several realistic settings.

Our work is related to that of Bester and Hansen (2013), who considered a grouped effects estimator (i.e., group-specific intercepts) in nonlinear static panel data models to reduce the bias from the incidental parameters problem. Our setting is different in that we consider a linear dynamic panel model where all coefficients may vary across individuals. But similar to Bester and Hansen (2013), we find that a trade-off exists in the specification of groups. If the groups are large, then pooling reduces the bias of the coefficient on the lagged dependent variable due to small T. But large groups also imply that the groups are likely to have more heterogeneity within groups and the pooling bias from heterogeneous coefficients is larger.

II. Grouped Coefficients Estimator

We consider the following heterogeneous dynamic panel model:

$$y_{igt} = \gamma_{ig} y_{ig,t-1} + \beta_{ig} x_{igt} + \alpha_{ig} + \varepsilon_{igt} \tag{1}$$

for individual i in group g at time t. Each individual is uniquely assigned to a group and group membership does not vary over time. The total number of individuals in the sample is denoted N, the total number of individuals in group g is N_g and the number of time periods is denoted T. We consider the case where T is small and $|\gamma_{ig}| < 1$. Aside from the multilevel structure, this is the same as the model in section 2 of Pesaran and Smith (1995).

We specify the following multilevel structure:

$$\begin{aligned} \gamma_{ig} &= \gamma + u_g^{\gamma} + v_{ig}^{\gamma} \\ \beta_{ig} &= \beta + u_g^{\beta} + v_{ig}^{\beta} \\ \alpha_{ig} &= \alpha + u_g^{\alpha} + v_{ig}^{\alpha} \end{aligned} \tag{2}$$

where the coefficient heterogeneity has a group-level component (denoted by subscript g) as well as an individual-level component (denoted by subscript ig). The individual-level components are independent across individuals. Our objective is to estimate γ and β.

The grouped coefficients estimator exploits knowledge of the multilevel structure of the data. To implement the grouped coefficients estimator, we estimate separate small-T dynamic panel regressions for each group. We consider three estimators for these group-level regressions based on the following moment conditions:

$$\begin{array}{lll} \text{OLS}: & E\left[\mathbf{z}_{igt}\, w_{igt}\right]=0 & t=2,\ldots,T \\ \text{Arellano}-\text{Bond GMM}: & E\left[\mathbf{z}^{*}_{igt}\Delta w_{igt}\right]=0 & t=3,\ldots,T \\ \text{Blundell}-\text{Bond GMM}: & E\left[\mathbf{z}^{*}_{igt}\Delta w_{igt}\right]=0 & t=3,\ldots,T \\ & E\left[\tilde{\mathbf{z}}_{igt}\, w_{igt}\right]=0 & t=3,\ldots,T \end{array}$$

where $w_{igt}=y_{igt}-(\gamma+u_g^{\gamma})y_{ig,t-1}-(\beta+u_g^{\beta})x_{igt}-(\alpha+u_g^{\alpha})$, $\mathbf{z}_{igt}=\left[y_{ig,t-1},x_{igt},1\right]'$, $\mathbf{z}^{*}_{igt}=\left[y_{ig,1},\ldots,y_{ig,t-2},\Delta x_{igt}\right]'$ and $\tilde{\mathbf{z}}_{igt}=\left[\Delta y_{ig,t-1},x_{igt},1\right]'$. We expect the GMM estimators to perform better than OLS when α_{ig} has significant within-group heterogeneity, and we expect OLS to perform better when all parameters are (close to) homogeneous within groups. We use the Arellano–Bond (AB) estimator because it is extensively used, and we use the BB estimator because its additional moment conditions improve performance significantly (Blundell and Bond, 1998).[4] Several other estimators have been proposed to reduce bias in models with heterogeneous intercepts, and these estimators could also be used to form a grouped coefficients estimate.

Let $\hat{\boldsymbol{\theta}}_g$ denote the vector of parameters from the estimator for group g. The grouped coefficients' estimate of the mean coefficient is the weighted average of the coefficients across groups,

$$\hat{\boldsymbol{\theta}}_{GC}=\sum_{g=1}^{N_g} w_g\hat{\boldsymbol{\theta}}_g \tag{3}$$

where w_g is the weight for group g. The weight will typically be the share of observations within each group (N_g/N), but a particular application may call for a different weighting scheme. Assuming independence between individuals, the variance matrix is calculated as $V\left(\hat{\boldsymbol{\theta}}_{GC}\right)=\sum_{g=1}^{N_g} w_g^2\, V\left(\hat{\boldsymbol{\theta}}_g\right)$. Otherwise, the variance matrix could be estimated using a bootstrap procedure that permits cross-sectional dependence of the errors (see Cameron and Trivedi, 2005).

If groups are defined such that there is no within-group coefficient heterogeneity – there may still be within-group heterogeneity of intercepts – and the groups are sufficiently large, then the grouped coefficients estimator consistently estimates the coefficients specific to each group. Of course, it is unlikely in practice to be able to define large groups that have no within-group heterogeneity. But as we show in our Monte Carlo simulations, the bias from coefficient heterogeneity is reduced as groups are defined with less within-group heterogeneity and group size remains large.

The grouped coefficients estimator may also reduce the bias from cross-sectional dependence due to common factors by including time FE in the group-specific regressions.[5] Sarafidis and Robertson (2009) show that GMM estimates with small T can be severely biased in the presence of a common factor even when the common factor is serially uncorrelated. They show that including time FE reduces the bias, but the bias can still be large if the variance of the factor loadings is large. The grouped coefficients estimator further reduces the bias from common factors when the variance of the

[4] The Arellano–Bond estimator is also commonly referred to as difference GMM and the Blundell–Bond estimator as system GMM. The Blundell–Bond estimator is also sometimes called the Arellano-Bover estimator since Arellano and Bover (1995) also proposed using moment conditions based on the level of the composite error.

[5] Alternatively, if all coefficients are expected to be homogeneous except for the factor loadings of common factors, then bias can be reduced by estimating a single model with group-specific coefficients on time fixed effects.

factor loadings is large by estimating coefficients on the time FE separately for groups of relatively homogeneous individuals.

Comparison to other approaches

The individual coefficients OLS estimator proposed by Pesaran and Smith (1995) is a special case of grouped coefficients with $N_g = 1$ and OLS as the within-group estimator.[6] This estimator gives consistent estimates of the short-run and long-run average marginal effects when T is large. Individual coefficients OLS is, however, inconsistent for small T because the coefficients in each individual regression exhibit small-sample bias.

Pooled GMM estimators such as AB or BB are also inconsistent when the coefficients are heterogeneous (Pesaran and Smith, 1995). These estimators use lagged values as instruments for $y_{ig,t-1}$. Intuitively, pooled GMM estimators are biased because the composite error term in the pooled model contains $y_{ig,t-1}$. Thus, lags of the dependent variable are not valid instruments. The grouped coefficients estimator seeks to manage the trade-off between the small-T bias that afflicts individual coefficients OLS and the pooling bias that afflicts pooled GMM. Estimating coefficients specific to each group removes the pooling bias from coefficient heterogeneity, and the small-T dynamic panel estimator removes the bias from heterogeneous intercepts and small T.

An advantage of the grouped coefficients estimator over the Bayesian random coefficients estimator proposed by Hsiao *et al.* (1999) is that the grouped coefficients estimator does not impose a specific distributional form on the coefficient heterogeneity and allows any form of correlation between the coefficients at the group level. The grouped coefficients estimator also allows the unobserved group-specific heterogeneity to be correlated with the regressors. Furthermore, the grouped coefficients estimator is appealing because it is computationally simple compared with Bayesian estimators.

The use of grouped coefficients to reduce bias in dynamic panels is related to the work of Bester and Hansen (2013). They consider a grouped effects estimator (group-specific intercepts) in nonlinear static panel data models with FE to reduce bias from the incidental parameters problem. They find a trade-off from the bias of incidental parameters and the bias from misspecification of the unobserved heterogeneity. Estimating group-specific intercepts reduces the bias from incidental parameters if groups are specified with a large number of individuals, but if groups have a large number of individuals, then there is also likely to be more unobserved heterogeneity within groups. An analogous trade-off exists in the specification of groups in our setting.[7] If the groups are large, then dynamic panel estimators have minimal bias from small T. But large groups also imply that the groups are likely to have more within-group coefficient heterogeneity and the pooling bias is larger.

III. Monte Carlo Simulations

Simulation design

We compare the bias of pooled, individual-coefficients and grouped-coefficients estimators using Monte Carlo simulations. We simulate the heterogeneous dynamic panel model defined in Equation (1) with normally distributed coefficients. The model is

$$\begin{aligned} y_{igt} &= \gamma_{ig} y_{ig,t-1} + \beta_{ig} x_{igt} + \alpha_{ig} + \varepsilon_{igt} \\ x_{igt} &= \mu_{ig}(1-\rho) + \rho\, x_{ig,t-1} + \eta_{igt} \\ &\quad \gamma_{ig} \sim \mathcal{N}\left(\gamma, \sigma_\gamma^2\right),\ \beta_{ig} \sim \mathcal{N}\left(\beta, \sigma_\beta^2\right),\ \alpha_{ig} \sim \mathcal{N}\left(\alpha, \sigma_\alpha^2\right) \\ &\quad \varepsilon_{igt} \sim \mathcal{N}\left(0, \sigma_\varepsilon^2\right),\ \mu_{ig} \sim \mathcal{N}\left(\mu, \sigma_\mu^2\right),\ \eta_{igt} \sim \mathcal{N}\left(0, \sigma_x^2\right) \end{aligned} \tag{4}$$

The regressor x_{igt} is strictly exogenous and is uncorrelated with the unobserved heterogeneity in the intercept α_{ig}. The idiosyncratic errors, ε_{igt}, are independently and identically distributed. The random coefficients are assumed to have a multilevel structure, where the individual and group-level components are independently and normally distributed:

[6] Pesaran and Smith call this the mean group estimator.
[7] The incidental parameters problem underlies the bias of fixed effects in linear dynamic panels as well (see Lancaster, 2000).

$$\begin{aligned}
\gamma_{ig} &= \gamma + u_g^{\gamma} + v_{ig}^{\gamma} \\
&\quad u_g^{\gamma} \sim \mathcal{N}(0,(1-\delta)\sigma_{\gamma}^2) \qquad v_{ig}^{\gamma} \sim \mathcal{N}(0,\delta\sigma_{\gamma}^2), \qquad E\left[u_g^{\gamma} v_{ig}^{\gamma}\right] = 0 \\
\beta_{ig} &= \beta + u_g^{\beta} + v_{ig}^{\beta} \\
&\quad u_g^{\beta} \sim \mathcal{N}(0,(1-\delta)\sigma_{\beta}^2) \qquad v_{ig}^{\beta} \sim \mathcal{N}(0,\delta\sigma_{\beta}^2), \qquad E\left[u_g^{\beta} v_{ig}^{\beta}\right] = 0 \\
\alpha_{ig} &= \alpha + u_g^{\alpha} + v_{ig}^{\alpha} \\
&\quad u_g^{\alpha} \sim \mathcal{N}(0,(1-\delta)\sigma_{\alpha}^2) \qquad v_{ig}^{\alpha} \sim \mathcal{N}(0,\delta\sigma_{\alpha}^2), \qquad E\left[u_g^{\alpha} v_{ig}^{\alpha}\right] = 0
\end{aligned} \tag{5}$$

The parameter δ measures the proportion of the parameter variance that occurs within groups (i.e., δ is the intra-class correlation coefficient).

Pesaran and Smith (1995) derived the asymptotic bias of pooled FE for the large T, large N case. These expressions are complicated, and the analogous expressions for small T are even more complicated. In the supplemental appendix, we present two special cases of the expressions of Pesaran and Smith. These bias expressions allow us to determine the parameters that are likely to affect the bias of pooled estimators when T is small in order to determine which parameters to vary in our simulations. The formulas show that the bias depends on the mean coefficients (γ and β), the degree of parameter heterogeneity (σ_{γ}^2 and σ_{β}^2), the autocorrelation of the independent variable (ρ) and the signal-to-noise ratio ($\sigma_x^2/\sigma_{\varepsilon}^2$). We do not vary β or ρ in our simulations for brevity, but in general we expect pooled estimates of the coefficient on the lagged dependent variable to have less upward bias if the independent variable explains less of the variation in y_{igt} (i.e., β is smaller) or there is less autocorrelation in the independent variable (i.e., ρ is smaller).

Table 1 shows the parameter values used in our simulations. We chose these values to be similar to those in the simulations of Arellano and Bond (1991) and Hsiao *et al.* (1999). The top half of Table 1 shows parameter values that we hold constant throughout all of our simulations. The bottom half of Table 1 shows parameter values that we vary in our simulations. We simulate every potential combination of these parameter values. We hold σ_{ε} constant in our simulations, so that we vary the signal-to-noise ratio by varying σ_x. Our simulated data set has 1600 individuals in 40 groups; thus each group has 40 individuals. For each set of parameters, we perform a Monte Carlo simulation for different estimators with 500 replications.

Table 1. Parameters for Monte Carlo simulations

Parameter	Value(s)
Parameters constant across simulations	
α	0
σ_{α}	1.0
β	1.0
μ	0
σ_{μ}	0
ρ	0.8
σ_{ε}	1.0
N	1600
N_g	40
Number of replications	500
Parameters varying across simulations	
γ	{0.5, 0.8}
σ_x	{0.5, 2.0}
T	{6, 10}
σ_{γ}	{0.25, 0.5}
σ_{β}	{0.5, 1.0}
δ	{0, 1/3, 2/3, 1}

When generating the data, we use the normal distributions in Equation (5), but we censor them to impose stationarity and nonnegative effects for each individual. We replace individual values of γ_{ig} with −0.95 if the random draw is less than −0.95, and replaced γ_{ig} with 0.95 if the random draw is greater than 0.95. For the coefficient on the independent variable (β_{ig}), we ensure that the coefficients for all individuals had the same sign by replacing β_{ig} with zero if the random draw is less than zero. The values of y_{igt} and x_{igt} are set to their mean, 0, in the first period. Then, we simulate $T+10$ periods of data and discard the first ten simulated periods. Different estimators are then used to estimate the mean coefficient on the lagged dependent variable (γ) and the mean coefficient on the independent variable (β).

Monte Carlo results

Figures 1 and 2 show results with $\sigma_x = 0.5$ and $T = 6$. Simulation results with $\sigma_x = 2.0$ and $T = 10$ are presented in the supplemental appendix. These figures plot the mean estimated coefficient as a function of within-group heterogeneity (δ). When $\delta = 0$, none of the parameter variation is within groups. When $\delta = 1$, all of the parameter variation is within groups. The first row of plots shows results with $\sigma_\gamma = 0.25$ and the second row shows results with $\sigma_\gamma = 0.5$. The first column of plots shows results with $\sigma_\beta = 0.5$ and the second column shows results with $\sigma_\beta = 1.0$.

The solid black line shows the true average coefficient (True). Note that the true mean coefficient may differ from the mean of the simulated normal distribution because we censor the simulated coefficients as described in the previous section. The dashed lines show results from pooled estimators. We estimate by OLS, FE, AB and BB. The solid purple line shows results from individual coefficients (IC-OLS), where a separate regression is estimated for each individual. The solid red line is for grouped coefficients OLS (GC-OLS), in which a separate OLS regression is estimated for each group. The solid orange line is for grouped coefficients Arellano–Bond (GC-AB) and the solid blue line is for grouped coefficients Blundell–Bond (GC-BB), where the AB and BB estimators are applied separately for each group. These figures are available in colour in the online version of this article.

The pooled OLS and pooled BB estimates of the coefficient on the lag are biased upwards because the parameter heterogeneity induces positive autocorrelation of the composite error term. The pooled AB estimates of the coefficient on the lag are often biased downwards when $\sigma_x = 0.5$, in agreement with the negative bias found by Bun and Kiviet (2006). But the bias of pooled AB estimates changes direction depending mostly on the magnitude of σ_x (see the supplemental appendix).

Pooled FE have two sources of bias. The bias from small T creates downward bias in the coefficient on the lag and the bias from heterogeneous coefficients creates upward bias. For $T = 6$ and $\sigma_x = 0.5$, the pooled FE estimate of the coefficient on the lag is biased downwards (see Figs 1 and 2). However, when $T = 10$ or $\sigma_x = 2$ the coefficient on the lag can be biased upwards with FE (see the supplemental appendix). Therefore, the standard result from the literature that estimates of the coefficient on the lag from pooled OLS and FE create upper and lower bounds of the true coefficient (e.g., Nerlove, 1971) does not necessarily hold when there is coefficient heterogeneity.

The pooled BB estimator gives severely biased estimates of the coefficient on the lag when there is coefficient heterogeneity. When $\gamma = 0.5$, the pooled BB estimate of the coefficient on the lag is about 0.65 when $\sigma_\gamma = 0.25$ and $\sigma_\beta = 0.5$ (upper left plot of Fig. 1). The upward bias is much larger when there is more coefficient heterogeneity. When $\gamma = 0.8$, we simulate the true average coefficient on the lag as about 0.75 when $\sigma_\gamma = 0.25$ and $\sigma_\beta = 0.5$ because we censor the individual coefficients at 0.95 (upper left plot of Fig. 2). But the pooled BB estimate of the mean coefficient on the lag is roughly 0.97 – larger than any of the individual coefficients.

The source of the bias of the pooled BB estimator is the autocorrelation of the error term. This suggests that the Arellano and Bond (1991) test for autocorrelation of the errors could be used to provide a warning that the BB estimator is biased. However, our Monte Carlo simulations indicate that this test generally lacks power. In Table 2, we report the probability of rejecting the null of no autocorrelation of the error.[8] The probability of rejecting the null hypothesis is relatively low for our parameter values unless the coefficient heterogeneity is very large. For example, when $\gamma = 0.8$, $\sigma_\gamma = 0.25$ and $\sigma_\beta = 0.5$, the pooled BB estimate of the mean coefficient on the lag is roughly 0.97 (Fig. 2) but the AB test only rejects the null hypothesis of no autocorrelation of the error with a probability of 0.07 (Table 2).

Grouped coefficients estimators can substantially reduce the bias from coefficient heterogeneity. When groups are chosen such that there is no within-group parameter heterogeneity ($\delta = 0$), then the grouped coefficients OLS and grouped

[8] That is, we report the probability of rejecting the null of no second-order autocorrelation of the first differences.

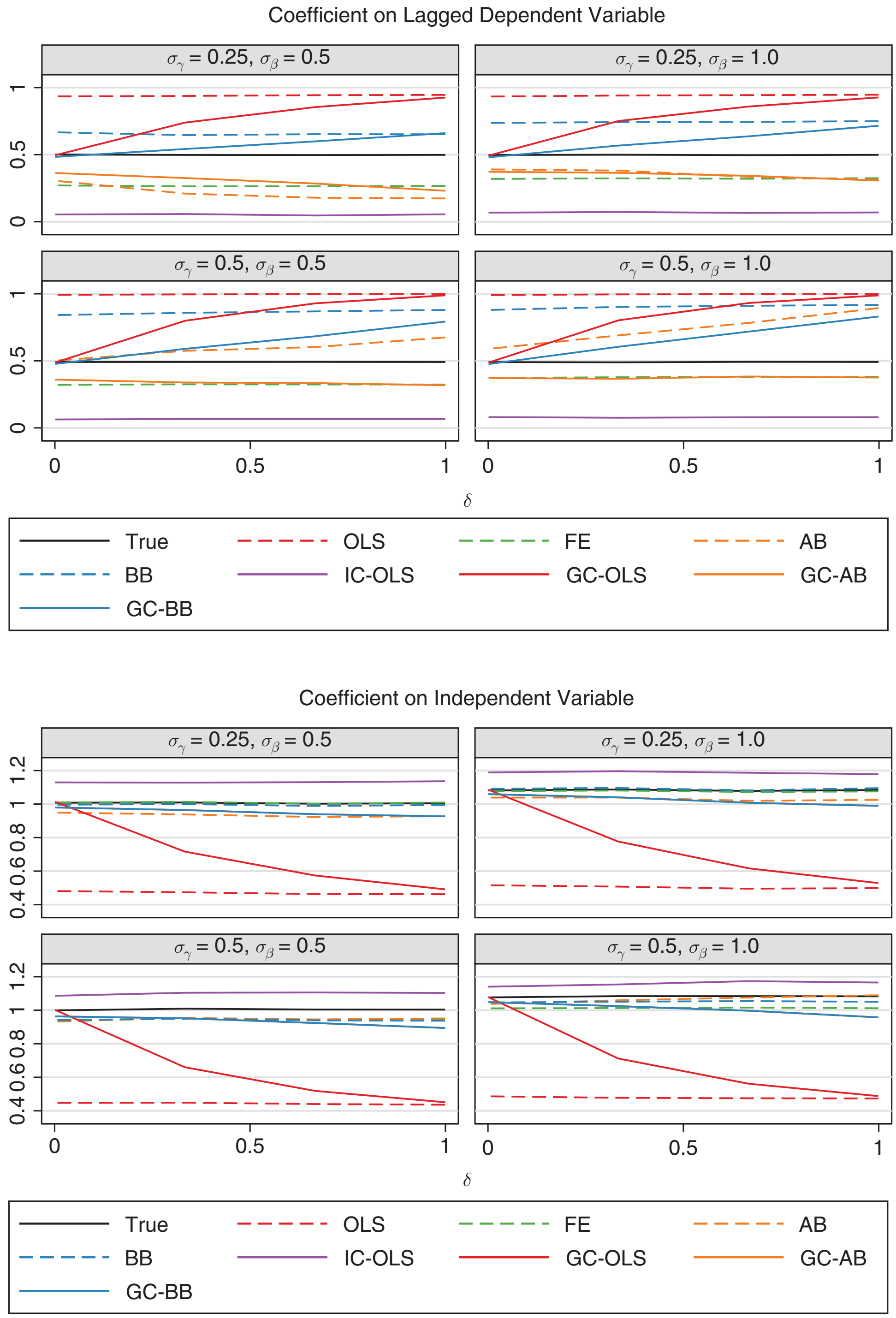

Fig. 1. Monte Carlo results with $\gamma = 0.5$, $\sigma_x = 0.5$ and $T = 6$
Notes: We abbreviate the estimators as follows: ordinary least squares (OLS), Arellano–Bond (AB) and Blundell–Bond (BB). Individual coefficients OLS are denoted IC-OLS. The 'GC-' abbreviation that precedes the estimators indicates grouped coefficients. This figure is available in colour in the online version of the article.

coefficients BB estimators are unbiased.[9] When groups are chosen such that all of the coefficient heterogeneity is within groups ($\delta = 1$), then the bias of grouped coefficients estimate of the

[9] We obtain the result that grouped coefficients OLS is unbiased because we assume that there is no heterogeneity of the intercepts when $\delta = 0$. In practice, there may be substantial heterogeneity of the intercepts within groups in which case the grouped coefficients Blundell–Bond estimator would be preferred.

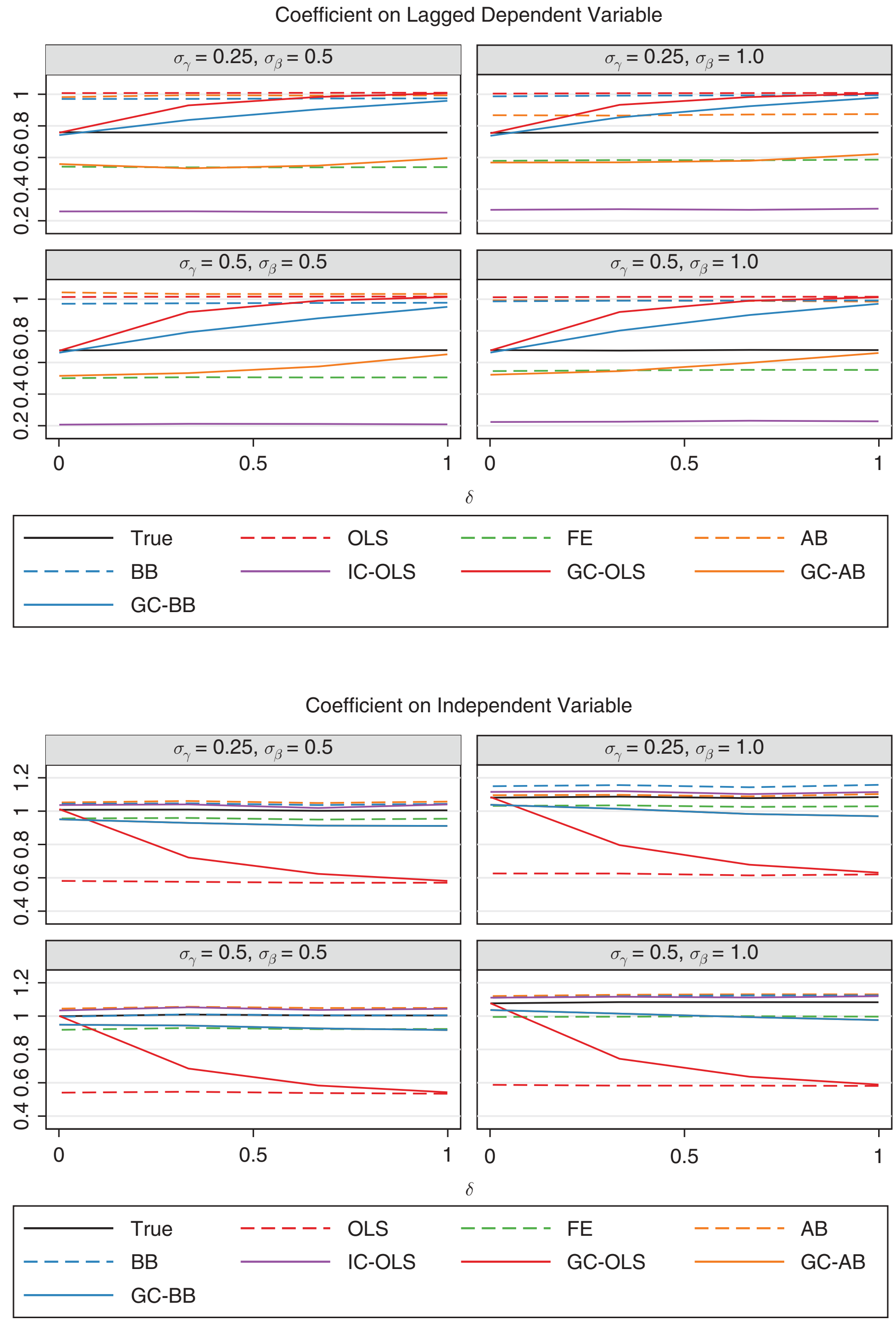

Fig. 2. Monte Carlo results with $\gamma = 0.8$, $\sigma_x = 0.5$ and $T = 6$
Notes: We abbreviate the estimators as follows: ordinary least squares (OLS), Arellano–Bond (AB) and Blundell–Bond (BB). Individual coefficients OLS are denoted IC-OLS. The 'GC-' abbreviation that precedes the estimators indicates grouped coefficients. This figure is available in colour in the online version of the article.

coefficient on the lag is similar to the bias of pooled estimators. For the BB estimator, the bias reduction on the coefficient on the lag is nearly linear in δ. Therefore, even if groups are chosen such that only half of the coefficient heterogeneity is within groups ($\delta = 0.5$), grouped coefficients BB can reduce the bias on the coefficient on the lag by roughly half. The grouped coefficients AB estimator does not perform well when $\sigma_x = 0.5$. This appears, however, to be more a problem of

Table 2. Probability of rejecting null hypothesis of no autocorrelation of error with σ_x= 0.5 and $T = 6$

σ_γ	σ_β	Probability of rejecting null
$\gamma = 0.5$		
0.25	0.5	0.272
0.25	1.0	0.222
0.5	0.5	0.736
0.5	1.0	0.658
$\gamma = 0.8$		
0.25	0.5	0.070
0.25	1.0	0.048
0.5	0.5	0.326
0.5	1.0	0.232

the bias of the AB estimator under these conditions rather than the bias of grouped coefficients.

For the coefficient on the independent variable, the pooled OLS and FE estimators are biased downwards and the individual coefficients estimates are biased upwards. Note that grouped coefficients BB estimates of the coefficient on the independent variable can be biased more than pooled BB when most of the parameter heterogeneity is within groups (large δ). But the bias of grouped coefficient BB is relatively small as long as groups are chosen such that less than half of the coefficient heterogeneity is within groups.

IV. Test for Coefficient Heterogeneity

If the coefficients are homogeneous across individuals, then pooled dynamic panel estimators are consistent. The grouped coefficients estimator is also consistent in this setting, although in general it would be less efficient. If the coefficients are heterogeneous but have no multilevel structure, then the pooled and grouped coefficients estimators are consistent for the same parameter (see the $\delta = 1$ case in our simulations). Both estimators are biased, but because each group is representative of the population, grouping does not enable bias reduction. The two estimators are consistent for different parameters if the coefficients have a multilevel structure.

We test for coefficient heterogeneity using a Hausman-type test with a Wald statistic.[10] To construct the statistic, we combine the data into a pooled model with group-level interaction terms, that is,

$$y_{igt} = \sum_{j=1}^{G} \left(\gamma_{ij} d_{gj} y_{ig,t-1} + \beta_{ig} d_{gj} x_{igt} + \alpha_{ig} d_{gj} \right) + \varepsilon_{igt} \tag{6}$$

where d_{gj} is a dummy variable that equals 1 if $g = j$ and 0 otherwise and G is the total number of groups. Applying the OLS, AB GMM and BB GMM estimators to Equation (6) produces the same coefficient estimates as the group-by-group regressions, but it enables easy tests of cross-group restrictions. We calculate the Wald statistic for a test of the null hypothesis $H_0 : \theta_1 = \theta_2 = \ldots = \theta_G$ with $(G-1)\ k$ degrees of freedom, where k denotes the number of regressors. This test is robust to heteroscedasticity if the variance matrix is estimated with robust SEs.

Under the null hypothesis, there is no coefficient heterogeneity across groups and the pooled estimator and grouped coefficients estimator are consistent for the same parameter. Under the alternative hypothesis, coefficients are significantly different between groups. Rejection of the null hypothesis implies that the pooled estimator is biased unless the parameter heterogeneity is of the sort that does not cause bias because, for example, the y_{igt} and x_{igt} variables are not autocorrelated. If the parameter heterogeneity is of the sort that does not cause bias, then we would expect the average of the grouped coefficients to be similar to the pooled estimates. Rejection of the null does not necessarily imply that the grouped coefficients estimator is consistent for the parameters of interest. Consistency of the grouped coefficients depends on the amount of within-group coefficient heterogeneity. With sufficient data, researchers may be able to investigate this possibility by subdividing

[10] We could instead have implemented the test in the more conventional Hausman form as $H = \left(\hat{\boldsymbol{\theta}}_{GC} - \hat{\boldsymbol{\theta}}_{PL}\right)' \left(N^{-1}\hat{\mathbf{V}}_H\right)^{-1} \left(\hat{\boldsymbol{\theta}}_{GC} - \hat{\boldsymbol{\theta}}_{PL}\right)$, where $\hat{\boldsymbol{\theta}}_{GC}$ denotes the grouped coefficients estimate, $\hat{\boldsymbol{\theta}}_{PL}$ denotes the analogous pooled estimate and $\hat{\mathbf{V}}_H$ denotes an estimate of the asymptotic variance of the limiting distribution of the root-N difference of the two estimates. Hsiao and Pesaran (2008) suggest a similar approach based on individual coefficients OLS when T is large. Standard asymptotic theory implies that the H is asymptotically distributed $\chi^2(k)$ under the null hypothesis.

the groups and assessing whether the estimated parameters change significantly. Overall, the Wald statistic can provide useful information even if a researcher is suspicious that there remains some within-group heterogeneity and therefore that the grouped coefficients estimator is biased.

V. Empirical Example: Labour Demand

We illustrate the grouped coefficients estimator by estimating labour demand using the data from Arellano and Bond (1991). Our estimates are from a sample of 123 manufacturing firms grouped in seven sub-sectors over the 6-year period 1977 to 1982. The original data from Arellano and Bond (1991) are an unbalanced panel of 140 firms grouped in nine sub-sectors over the period 1976 to 1984. Data are missing for roughly half the firms in 1976, 1983 and 1984, so we drop data from these years.[11] We also drop data from two sub-sectors that have little data and produce outlier results.[12]

Several different labour demand equations have been estimated in previous articles. We estimate the same dynamic labour demand equation as in Blundell and Bond (1998):

$$n_{i,t} = \gamma n_{i,t-1} + \beta_1 w_{i,t} + \beta_2 w_{i,t-1} + \beta_3 k_{i,t} + \beta_4 k_{i,t-1} + \alpha_i + \lambda_t + \varepsilon_{it} \qquad (7)$$

where $n_{i,t}$ is the log of the number of employees in firm i in year t, $w_{i,t}$ is the log of the wage rate, $k_{i,t}$ is the log of the capital stock, α_i are firm FE and λ_t are year FE.

Table 3 reports coefficient estimates from seven different estimators. Pooled estimators (columns 1–3) estimate a single coefficient on each variable for all firms. Grouped coefficients (columns 4–6) estimate separate coefficients for each of the seven

Table 3. Estimates of labour demand

	Pooled			Grouped coefficients			
	OLS (1)	AB (2)	BB (3)	OLS (4)	AB (5)	BB (6)	IC-OLS (7)
$n_{i,t-1}$	0.954***	0.900***	0.846***	0.944***	0.535***	0.719***	0.499***
	(0.008)	(0.149)	(0.096)	(0.011)	(0.130)	(0.084)	(0.104)
w_{it}	– 0.380**	– 0.348	– 0.492*	– 0.263***	– 0.546***	– 0.435***	– 0.408**
	(0.169)	(0.293)	(0.277)	(0.075)	(0.197)	(0.138)	(0.200)
$w_{i,t-1}$	0.331**	0.189	0.116	0.232***	– 0.165	0.160	
	(0.162)	(0.190)	(0.198)	(0.074)	(0.171)	(0.392)	
k_{it}	0.334***	0.335*	0.444***	0.307***	0.443***	0.393***	0.487***
	(0.056)	(0.176)	(0.133)	(0.042)	(0.080)	(0.103)	(0.080)
$k_{i,t-1}$	– 0.290***	– 0.424***	– 0.303***	– 0.254***	0.039	– 0.031	
	(0.055)	(0.150)	(0.113)	(0.044)	(0.126)	(0.122)	

Notes: This table reports coefficient estimates of Equation (7). Pooled estimators estimate a single coefficient on each variable for all firms. Grouped coefficients estimate separate coefficients for each sub-sector, then averages coefficients across sub-sectors. We abbreviate the estimators as follows: Ordinary Least Squares (OLS), Arellano–Bond (AB) and Blundell–Bond (BB). Individual coefficients OLS (IC-OLS) estimates a separate OLS regression for each firm and then averages coefficients across firms. Estimates in columns (1)–(6) include year fixed effects. All SEs are robust to heteroscedasticity across firms.
***Significant at the 1% level.
**Significant at the 5% level.
*Significant at the 10% level.

[11] A table with the number of firms in each year is in the supplemental appendix.
[12] We dropped data from sub-sectors labelled ‘3’ and ‘6’ in the data. Sub-sector 6 was dropped since it only contained data from five firms. Sub-sector 3 was dropped because Arellano–Bond and Blundell–Bond estimates of the coefficient on the lagged dependent variable were −1.22 and −0.91, respectively. Including sub-sector 3 in our sample would further strengthen our result that grouped coefficients estimates give a smaller coefficient on the lagged dependent variable.

sub-sectors and calculate a weighted average of the coefficients across sub-sectors, where weights are the share of firms within each sub-sector. Pooled and grouped coefficients OLS omit the firm fixed effect. Individual coefficients OLS (column 7) estimates a separate OLS regression for each firm and calculates the average of the coefficients across firms. We omit lagged wages and lagged capital from the firm-specific regressions since there are only a maximum of six observations per firm. All SEs are robust to heteroscedasticity across firms. The variance of grouped and individual coefficients estimators is calculated assuming independence between firms.

We use two-step GMM for the AB and BB estimators. We do not assume wages and capital are exogenous. Our specification of these estimators is the same as in Blundell and Bond (1998), except that they use one-step GMM and the identity matrix as the initial weighting matrix for GMM. For grouped coefficients, we use the share of firms within each sub-sector as the weights. SEs are robust to heteroscedasticity across firms but assume independence between firms.

The results in Table 3 generally conform to our expectations from the preceding sections. Estimates of the coefficient on the lagged dependent variable do not differ substantially between pooled OLS (column 1) and grouped coefficients OLS (column 4) because there likely remains substantial heterogeneity of the intercepts between firms within sub-sectors. The average coefficient from estimating separate regressions for each firm is only 0.499 (column 7). This estimate is biased downwards because of small-T bias, as seen in our simulations.

Estimates of the coefficient on the lagged dependent variable from pooled AB and pooled BB (columns 2 and 3) are still quite large. For neither of these estimators do we reject the null hypothesis of no serial correlation of the errors at the 10% level (results not reported), giving no indication that there may be misspecification. However, the average coefficient on the lag from grouped coefficients AB and grouped coefficients BB estimators are much smaller. One explanation for the larger coefficient for the pooled estimators is that there is coefficient heterogeneity and the pooled coefficient on the lag is capturing autocorrelation of wages, capital or common factors.

Table 4. Tests for coefficient heterogeneity of labour demand equation (H_0 = coefficient homogeneity)

Variables	χ^2	p-value
Panel A. Arellano–Bond		
Lagged dependent variable	10.15	0.118
Wages and lagged wages	16.63	0.083
Capital and lagged capital	51.28	<0.001
Year fixed effects	72.39	<0.001
All	309.67	<0.001
Panel B. Blundell–Bond		
Lagged dependent variable	12.99	0.023
Wages and lagged wages	40.98	<0.001
Capital and lagged capital	87.26	<0.001
Year fixed effects	42.87	0.007
All	8363.80	<0.001

Table 4 reports test statistics for a test of coefficient heterogeneity. We test for coefficient heterogeneity by estimating the model with interaction terms between each variable and indicators for each group as in Equation (6), and then calculate a Wald statistic testing if the group-specific coefficients are equal for each variable. Panel A of Table 4 reports tests for the AB estimator and panel B for the BB estimator. The last row in each panel reports the statistic for the test with the null hypothesis that coefficients on all variables are homogeneous between groups. We strongly reject coefficient homogeneity for the AB and BB estimators. Other rows in Table 4 test for coefficient heterogeneity for a subset of variables. The Wald statistic is much smaller for homogeneity of the coefficient on the lagged dependent variable than other coefficients. Thus, the primary source of bias in the estimate of the coefficient on the lag is likely due to cross-sectional dependence from heterogeneous coefficients on autocorrelated independent variables (wages and capital) and heterogeneous factor loadings of common factors (year FE).

VI. Conclusion

We propose the use of a grouped coefficients estimator to reduce the bias in small-T dynamic panels with coefficient heterogeneity or cross-sectional dependence from common factors. The grouped coefficients estimator provides substantial bias reduction if groups are defined with relatively little within-group heterogeneity. We assume that the researcher has knowledge of the group membership. A natural extension of our work would be to define methods that assign groups when membership is unknown along the lines of Li *et al.* (2013) and Lin and Ng (2012).

In Monte Carlo simulations, we show that pooled GMM estimates of the coefficient on the lagged dependent variable can be severely biased upwards. Furthermore, the AB test for autocorrelation has little power to reject the null hypothesis of no autocorrelation of the errors. We provide a Hausman-type test that can be used to assess whether this bias is present. Our application to labour demand illustrates the upward bias of the coefficient on the lag from pooled estimators and that grouped coefficients provide a simple method to reduce the bias.

Disclosure Statement

No potential conflict of interest was reported by the authors.

Supplemental Data

Supplemental data for this article can be accessed here.

References

Arellano, M. and Bond, S. (1991) Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations, *The Review of Economic Studies*, **58**, 277–97. doi:10.2307/2297968

Arellano, M. and Bover, O. (1995) Another look at the instrumental variable estimation of error-components models, *Journal of Econometrics*, **68**, 29–51. doi:10.1016/0304-4076(94)01642-D

Bester, C. A. and Hansen, C. B. (2013) Grouped effects estimators in fixed effects models, *Journal of Econometrics*. doi:10.1016/j.jeconom.2012.08.022

Blundell, R. and Bond, S. (1998) Initial conditions and moment restrictions in dynamic panel data models, *Journal of Econometrics*, **87**, 115–43. doi:10.1016/S0304-4076(98)00009-8

Bruno, G. S. (2005) Approximating the bias of the LSDV estimator for dynamic unbalanced panel data models, *Economics Letters*, **87**, 361–6. doi:10.1016/j.econlet.2005.01.005

Bun, M. J. and Kiviet, J. F. (2006) The effects of dynamic feedbacks on LS and MM estimator accuracy in panel data models, *Journal of Econometrics*, **132**, 409–44. doi:10.1016/j.jeconom.2005.02.006

Bun, M. J. G. and Carree, M. A. (2005) Bias-corrected estimation in dynamic panel data models, *Journal of Business & Economic Statistics*, **23**, 200–10. doi:10.1198/073500104000000532

Cai, Z. and Li, Q. (2008) Nonparametric estimation of varying coefficient dynamic panel data models, *Econometric Theory*, **24**, 1321–42. doi:10.1017/S0266466608080523

Cameron, A. C. and Trivedi, P. K. (2005) *Microeconometrics: Methods and Applications*, Cambridge University Press, New York, NY.

Gelman, A. (2006) Multilevel (hierarchical) modeling: what it can and cannot do, *Technometrics*, **48**, 432–5. doi:10.1198/004017005000000661

Hahn, J., Hausman, J. and Kuersteiner, G. (2007) Long difference instrumental variables estimation for dynamic panel models with fixed effects, *Journal of Econometrics*, **140**, 574–617. doi:10.1016/j.jeconom.2006.07.005

Hahn, J. and Kuersteiner, G. (2002) Asymptotically unbiased inference for a dynamic panel model with fixed effects when both n and T are large, *Econometrica*, **70**, 1639–57. doi:10.1111/1468-0262.00344

Han, C., Phillips, P. C. B. and Sul, D. (2014) X-differencing and dynamic panel model estimation, *Econometric Theory*, **30**, 201–51. doi:10.1017/S0266466613000170

Hsiao, C. and Pesaran, M. (2008) Random coefficient models, in *The Econometrics of Panel Data*, Mátyás, L. and Sevestre, P. (Eds), Springer, Berlin, pp. 185–213.

Hsiao, C., Pesaran, M. H. and Tahmiscioglu, A. K. (1999) Bayes estimation of short-run coefficients in dynamic panel data models, in *Analysis of Panels and Limited Dependent Variable Models*, Hsiao, C., Lahiri, K., Lee, L.-F. *et al.* (Eds), Cambridge University Press, Cambridge, UK.

Hsiao, C., Pesaran, M. H. and Tahmiscioglu, A. K. (2002) Maximum likelihood estimation of fixed effects dynamic panel data models covering short time periods, *Journal of Econometrics*, **109**, 107–50.

Hsiao, C., Shen, Y. and Fujiki, H. (2005) Aggregate vs. disaggregate data analysis–a paradox in the estimation of a money demand function of Japan under the low interest rate policy, *Journal of Applied Econometrics*, **20**, 579–601. doi:10.1002/jae.806

Hsiao, C. and Tahmiscioglu, A. K. (1997) A panel analysis of liquidity constraints and firm investment, *Journal of the American Statistical Association*, **92**, 455–65. doi:10.1080/01621459.1997.10473997

Kiviet, J. F. (1995) On bias, inconsistency, and efficiency of various estimators in dynamic panel data models, *Journal of Econometrics*, **68**, 53–78. doi:10.1016/0304-4076(94)01643-E

Lancaster, T. (2000) The incidental parameter problem since 1948, *Journal of Econometrics*, **95**, 391–413. doi:10.1016/S0304-4076(99)00044-5

Lee, Y. (2012) Bias in dynamic panel models under time series misspecification, *Journal of Econometrics*, **169**, 54–60. Recent Advances in Panel Data, Nonlinear and Nonparametric Models: A Festschrift in Honor of Peter C.B. Phillips. doi:10.1016/j.jeconom.2012.01.009

Li, Q., Ouyang, D. and Racine, J. S. (2013) Categorical semiparametric varying-coefficient models, *Journal of Applied Econometrics*, **28**, 551–79. doi:10.1002/jae.1261

Lin, -C.-C. and Ng, S. (2012) Estimation of panel data models with parameter heterogeneity when group membership is unknown, *Journal of Econometric Methods*, **1**, 42–55. doi:10.1515/2156-6674.1000

Maddala, G. S., Trost, R. P., Li, H. *et al.* (1997) Estimation of short-run and long-run elasticities of energy demand from panel data using shrinkage estimators, *Journal of Business & Economic Statistics*, **15**, 90–100.

Nauges, C. and Thomas, A. (2003) Consistent estimation of dynamic panel data models with time-varying individual effects, *Annals of Economics and Statistics*, **70**, 53–75.

Nerlove, M. (1967) Experimental evidence on the estimation of dynamic economic relations from a time-series of cross sections, *Economic Studies Quarterly*, **18**, 42–74.

Nerlove, M. (1971) Further evidence on the estimation of dynamic economic relations from a time series of cross sections, *Econometrica*, **39**, 359–82. doi:10.2307/1913350

Pesaran, M. H. and Smith, R. (1995) Estimating long-run relationships from dynamic heterogeneous panels, *Journal of Econometrics*, **68**, 79–113. doi:10.1016/0304-4076(94)01644-F

Phillips, P. C. and Sul, D. (2007) Bias in dynamic panel estimation with fixed effects, incidental trends and cross section dependence, *Journal of Econometrics*, **137**, 162–88. doi:10.1016/j.jeconom.2006.03.009

Sarafidis, V. and Robertson, D. (2009) On the impact of error cross-sectional dependence in short dynamic panel estimation, *Econometrics Journal*, **12**, 62–81. doi:10.1111/j.1368-423X.2008.00260.x

Tran, K. C. and Tsionas, E. G. (2009) Local GMM estimation of semiparametric panel data with smooth coefficient models, *Econometric Reviews*, **29**, 39–61. doi:10.1080/07474930903327856

Wansbeek, T. (2012) On the remarkable success of the Arellano-Bond estimator, *Aenorm*, **20**, 15–20.

Zhou, Q., Faff, R. and Alpert, K. (2014) Bias correction in the estimation of dynamic panel models in corporate finance, *Journal of Corporate Finance*, **25**, 494–513. doi:10.1016/j.jcorpfin.2014.01.009