ISSN 1440-771X

Department of Econometrics and Business Statistics

http://business.monash.edu/econometrics-and-business-statistics/research/publications

Celebrating 40 Years of Panel Data Analysis: Past, Present and Future

Vasilis Sarafidis and Tom Wansbeek

February 2020

Working Paper 06/20

Celebrating 40 Years of Panel Data Analysis: Past, Present and Future

Vasilis Sarafidis* Monash University

Tom Wansbeek† University of Groningen

February 3, 2020

Abstract

The present special issue features a collection of papers presented at the 2017 International Panel Data Conference, hosted by the University of Macedonia in Thessaloniki, Greece. The conference marked the 40th anniversary of the inaugural International Panel Data Conference, which was held in 1977 at INSEE in Paris, under the auspices of the French National Centre for Scientific Research. As a collection, the papers appearing in this special issue of the Journal of Econometrics continue to advance the analysis of panel data, and paint a state-of-the-art picture of the field.

Key words: Panel data analysis, unobserved heterogeneity, omitted variables, cross-sectional dependence, dynamic relationships, temporal effects, aggregation bias, nonlinear models, incidental parameter problem, common factor models, multi-dimensional data, multi-level data.

JEL: C23, C33.

*Corresponding author. Department of Econometrics and Business Statistics, Monash University, VIC 3145, Australia. e-mail: vasilis.sarafidis@monash.edu

†Department of Economics, University of Groningen, Groningen 9700, Netherlands. e-mail: t.j.wansbeek@rug.nl

1 Introduction

In the 1960's, it became apparent to policy makers in the U.S. that the enormous economic expansion following World War II, did not cure all major socioeconomic problems, nor prevent new ones from emerging. For instance, despite the fact that over the period 1950-1964, U.S. GDP grew by a staggering 71% in real terms, one in five Americans continued to live below the poverty line. In response, large-scale surveys were set up to collect data on the same families over time, aiming at a better understanding of the dynamics of the distribution of income and employment.

A prominent example of those surveys is the Panel Study of Income Dynamics (PSID), which was created in 1968 at the University of Michigan in order to assess the impact of President Johnson's 'War on Poverty' program. Over the past 50 years, the PSID has collected, and made available, survey data on more than 80,000 individuals, including information on income and poverty, work and employment, housing and commuting to work. An equally important example is the National Longitudinal Survey of Labor Market Experience (NLS), which was initiated in 1966 and originally included more than 5,000 respondents. More recently, the European Community Household Panel (ECHP) was established in 1994 for the purposes of representing the population of the European Union at the household and individual level, containing a wide range of information on living conditions.[1] These data sets constitute a cornerstone of the data infrastructure required for empirically-based research in the social sciences.

As panel data began to emerge, new methods were developed to analyse such data and improve our understanding of economic behaviour. This prompted the creation of new scientific fora, bringing together economists, econometricians, statisticians and social scientists to analyse and study important methodological issues in the field. The inagural International Panel Data Conference was held at INSEE in Paris during 1977, under the auspices of the French National Centre for Scientific Research. A collection of papers presented in that conference was published at the Annales de l'INSEE – No 30/31 1978 (nowadays, the Annals of Economics and Statistics), in a volume titled "The Econometrics of Panel Data", and edited by Marc Nerlove.[2]

Subsequently, the panel data literature has thrived and grown into a major subfield of econometrics. According to an assessment of research impact measures by Chang, McAleer and Oxley (2011), almost a quarter of the top 25 most highly cited papers published in the Journal of Econometrics, lies in the field of panel data econometrics. The paper by Arellano and Bond (1991), which deals with estimation of dynamic panel data models, has recently been listed as the single most cited paper in the field of economics as a whole over the past three decades.[3]

Developments in the panel data literature have accelerated rapidly in recent years, including in areas such as non-linear panels, high-dimensional data, factor models in economics and finance, pseudo-panels, to mention only a few. The present special issue features a collection of papers presented at the 2017 International Panel Data Confer-

[1] See Baltagi (2013), Ch. 1, and Hsiao (2014), Ch. 1 for a detailed description of all three data sets.

[2] See Nerlove (1978).

[3] See Tables 3 and 10 in Linnemer and Visser (2017), who collected citation statistics from the Web of Science database. Five journals are considered in their analysis, namely (in alphabetical order): American Economic Review, Econometrica, Journal of Political Economy, Quarterly Journal of Economics, and The Review of Economic Studies. Table 3 ranks papers according to the number of citations cumulated since 1991 (which favors older papers), while Table 10 normalises these figures according to year of publication.

ence, hosted by the University of Macedonia in Thessaloniki, Greece.[4] The conference marked the 40th anniversary of the inaugural International Panel Data Conference. As a collection, the papers appearing in this special issue of the Journal of Econometrics continue to advance the analysis of panel data and paint a state-of-the-art picture of the field.

2 Basic Motivation

Panel data provide repeated measurements on the same individual agents (such as households, firms, countries) at different points in time. The high popularity of analysing such data over the past four decades can largely be attributed to two main factors. First, the ability to control for certain sources of unobserved heterogeneity and endogeneity, due to (say) omitted variables and measurement error. Second, the ability to estimate dynamic relationships from micro data without suffering aggregation bias, and often using a relatively small number of time series observations.

To illustrate, consider the following linear panel data model:

$$y_{it} = c + \boldsymbol{\beta}'\mathbf{x}_{it} + \boldsymbol{\delta}'\mathbf{z}_{it} + \varepsilon_{it};\ i = 1, ..., N;\ t = 1, ..., T, \tag{1}$$

where y_{it} denotes the observation on the dependent variable for individual i at time t, $\mathbf{x}_{it}$ and $\mathbf{z}_{it}$ denote $[K_x \times 1]$ and $[K_z \times 1]$ vectors of exogenous variables, respectively, and $\boldsymbol{\beta}$, $\boldsymbol{\delta}$ denote the corresponding unknown parameters. Suppose that $\mathbf{z}_{it}$ is unobserved and correlated with $\mathbf{x}_{it}$. In this case, the least-squares estimator of $\boldsymbol{\beta}$ is subject to omitted variable bias. In the absence of repeated observations, consistent estimation of $\boldsymbol{\beta}$ typically requires the use of exogenous instruments. However, if repeated observations on a cross section of individuals are available, then under certain restrictions on $\mathbf{z}_{it}$, it becomes possible to control for omitted variables without instruments. For example, if $\mathbf{z}_{it}$ is time-invariant, i.e. $\mathbf{z}_{it} = \mathbf{z}_i$ for all t, the model can be expressed as follows:

$$\begin{aligned} y_{it} &= c + \boldsymbol{\beta}'\mathbf{x}_{it} + \eta_i + \varepsilon_{it}; & (2)\\ \eta_i &= \boldsymbol{\delta}'\mathbf{z}_i, & (3) \end{aligned}$$

where η_i denotes an individual-specific unobserved effect, which is a linear combination of all time-invariant omitted variables. In this case, a popular identification strategy involves transforming the model in terms of deviations from individual-specific averages to eliminate η_i, a nd applying least-squares. The resulting so-called 'within' or 'fixed effects' (FE) estimator of $\boldsymbol{\beta}$ is unbiased and consistent for T fixed or large, so long as $\mathbf{x}_{it}$ is strictly exogenous with respect to the idiosyncratic error term, i.e. $E(\varepsilon_{it}|\mathbf{x}_{i1}, ..., \mathbf{x}_{iT}) = 0$.

Similarly, suppose that the omitted variables can be decomposed into $\mathbf{z}_{it} = \left(\mathbf{z}_{it}^{(1)\prime},\ \mathbf{z}_{it}^{(2)\prime}\right)'$ such that $\mathbf{z}_{it}^{(1)} = \mathbf{z}_i^{(1)}$ and $\mathbf{z}_{it}^{(2)} = \mathbf{z}_t^{(2)}$ for all i and t respectively, and let $\boldsymbol{\delta}_{(1)}$ and $\boldsymbol{\delta}_{(2)}$ denote the corresponding coefficients of $\mathbf{z}_i^{(1)}$ and $\mathbf{z}_t^{(2)}$. Then, the model can be rewritten as

$$\begin{aligned} y_{it} &= c + \boldsymbol{\beta}'\mathbf{x}_{it} + \eta_i + \tau_t + \varepsilon_{it}; & (4)\\ \eta_i &= \boldsymbol{\delta}_{(1)}'\mathbf{z}_i^{(1)};\ \tau_t = \boldsymbol{\delta}_{(2)}'\mathbf{z}_t^{(2)}. & (5) \end{aligned}$$

[4]The last special issue of the Journal of Econometrics that focused entirely on panel data analysis dates back to 1995 (Vol. 68, no. 1), edited by Badi Baltagi. Many of the papers that appeared in that issue were solicited from the 4th International Panel Data Conference, held in Budapest, Hungary, June 18-19, 1992.

The model in Eq. (4) is commonly referred as the 'two-way effects' model because it controls for two distinct sources of unobserved heterogeneity, η_i and τ_t. Similarly as before, these additive effects can be eliminated by transforming (4) in terms of deviations from both individual- and time-specific averages.

An additional important advantage of panel data analysis is the ability to estimate dynamic or temporal effects from a relatively small number of time series observations, based on large N asymptotics.[5] For instance, consider a simple first-order dynamic panel data model with covariates:

$$y_{it} = c + \alpha y_{it-1} + \boldsymbol{\beta}' \mathbf{x}_{it} + \eta_i + \tau_t + \varepsilon_{it}. \tag{6}$$

The coefficient α has structural significance and captures habit formation, costs of adjustment and 'state dependence'. Thus, it enables a clear distinction between expected short- and long-run partial effects of predictors.

For fixed T, the FE estimator of α is not consistent because the within transformation induces a non-negligible correlation between the lagged dependent variable and the purely idiosyncratic error. This result is known as 'Nickell bias' (Nickell (1981)). The bias is of order $O\left(T^{-1}\right)$ and therefore it vanishes as T grows large.

Following the seminal papers by Anderson and Hsiao (1981) and Arellano and Bond (1991), a popular strategy to deal with 'Nickell bias' involves taking first-differences in Eq. (6), and using lagged values of y_{it-1} as instruments for the endogenous regressor, based on the Generalised Method of Moments (GMM). The properties of this estimator have been studied extensively under a large number of cases, including highly persistent data, weak instruments, and 'too many instruments'; see Bun and Sarafidis (2015) for a recent overview of the dynamic panel data literature.

3 Extensions and Issues

During the past few decades, the methods and models discussed above have been extended in several directions. These include (i) identifying economic relationships using nonlinear models, (ii) controlling for richer structures of unobserved heterogeneity compared to the two-way effects model, (iii) allowing for heterogeneity in the slope coefficients, and (iv) modelling richer data sets, such as panels with multiple dimensions. Almost all articles published in this special issue fit in at least one of these strands of literature. For the purposes of motivating and identifying the various contributions made, in what follows we provide a short (and by no means exhaustive) overview of important issues relevant to each of these strands.

3.1 Nonlinear Models

Many economic problems require fitting a nonlinear relationship between the response and the linear predictor. Prominent examples are 'models for discrete choice analysis'. For this class of models, the method of Maximum Likelihood (hereafter, ML) is the workhorse estimation approach.[6] Unfortunately, for the majority of nonlinear panel data models with fixed effects, it turns out that the ML estimator is not consistent as $N \to \infty$ when T is fixed. This is due to the 'incidental parameter problem', described by Neyman and

[5] See the seminal paper by Balestra and Nerlove (1966).

[6] Recent surveys of this field are provided by Arellano and Bonhomme (2011) and Greene (2015).

Scott (1948). To outline the problem, let $\log\left(f\left(y_{it};\mathbf{x}_{it},\boldsymbol{\beta},\eta_i\right)\right)$ denote the log-likelihood function associated with y_{it} (conditional on $\mathbf{x}_{it}$). Let also $\widehat{\boldsymbol{\beta}}_{ML}$ and $\widehat{\eta}_{i,ML}$ denote the ML estimator of $\boldsymbol{\beta}$ and η_i, respectively. Since there are T observations available to estimate η_i, the ML estimate of η_i remains random as $N \to \infty$ for T fixed. In linear models such randomness averages out, but in nonlinear models it does not. As a result, $\widehat{\boldsymbol{\beta}}_{ML}$ does not approach its true value asymptotically, and has bias of order $O\left(T^{-1}\right)$.

When $T \to \infty$, every $\widehat{\eta}_{i,ML}$ converges to the corresponding true value under certain regularity conditions, enabling point identification of $\boldsymbol{\beta}$. Essentially, the incidental parameter problem becomes an asymptotic bias problem, which is easier to tackle under appropriate assumptions.[7] As we shall shortly see, this result has prompted researchers to seek methods for reducing the small-T bias of the ML estimator, motivated by large T asymptotics.[8]

Several different approaches have been advocated in the literature to deal with the incidental parameter problem when T is small. Instead of maximising the likelihood directly with respect to an increasing number of fixed effects, one approach involves conditioning on a minimal sufficient statistic for the fixed effects, such that the resulting conditional likelihood depends on $\boldsymbol{\beta}$ but not on η_i.[9] Unfortunately, in many models such sufficient statistics do not exist.

An alternative approach involves controlling for the fixed effect using marginalising, differencing, integration or invariance arguments.[10] However, these arguments often rely on strong and restrictive assumptions, and therefore they are not always applicable. Also, removing the unobserved effects from the model precludes the estimation of partial effects.

More recently, approaches aiming at reducing the bias of $\widehat{\boldsymbol{\beta}}_{ML}$ have been advocated in the literature, using either 'model-free' bias correction or analytical bias correction. Within the former approach, prominent methods include panel jackknife (e.g. Dhaene and Jochmans (2015)), integrated likelihood with bias-reducing priors (Arellano and Bohnomme (2009)), and bootstrap (Kim and Sun (2016)). In the case of analytical bias correction, prominent examples include Hahn and Newey (2004) and Hahn and Kuersteiner (2011), both of which derive the approximate bias of $\widehat{\boldsymbol{\beta}}_{ML}$, and Bester and Hansen (2009) and Arellano and Hahn (2016), who derive the approximate bias of the log-likelihood.

Lastly, an emergent strand of the literature has shifted focus on partial identification (see e.g. Honoré and Tamer (2006)). The motivation for this literature is that point identification in nonlinear panels relies on strong assumptions, which in many cases are invoked on computational grounds rather than coherency with the data or economic theory; see Chamberlain (2010). The goal of partial identification analysis is to examine what conclusions can be drawn about the parameters of interest under weaker sets of assumptions, even if point identification fails. As forcefully argued by Manski (1989), identification is not an 'all-or-nothing' concept and there is much to be learned -even if not everything- from credible assumptions about the parameters of interest.[11]

[7] See Fernandez-Val and Weidner (2018) for an up-to-date analysis of the incidental parameter problem in large T panels.

[8] Notwithstanding the usefulness of large N, T asymptotic theory as a tool to guide small T bias correction, in practice one needs to be careful when invoking large T arguments for inference. Therefore, from the empirical point of view, fixed-T panel data theory remains highly relevant.

[9] See Andersen (1970) and Kalbfleisch and Sprott (1970).

[10] See Chamberlain (1985), Manski (1987), Honoré (1992), Lancaster (2002), and Bonhome (2012), among others.

[11] Molinari (2019) provides a useful introduction to this topic.

3.2 Common Factor Models

From at least as far back as Holtz-Eakin et al. (1988), it has been pointed out that the two-way effects model can be potentially too restrictive in practice, since it assumes that the unobserved effects enter in an additive fashion. A prominent framework that generalises the two-way effects model is the common factor approach. This allows multiple effects to enter in a *multiplicative* fashion, as opposed to an additive one, thus giving rise to a 'nonlinear components' model, or 'interactive effects'. Common factor structures offer wider scope for controlling for unobservables, including situations where there is cross-sectional dependence; see Sarafidis and Wansbeek (2012) for a recent overview.[12]

In terms of the motivation provided in Section 2, instead of the omitted variables being restricted to the form $\boldsymbol{\delta}'\mathbf{z}_{it} = \eta_i + \tau_t$, one generalises

$$\boldsymbol{\delta}'\mathbf{z}_{it} = \boldsymbol{\lambda}_i'\mathbf{f}_t, \tag{7}$$

where $\mathbf{f}_t$ and $\boldsymbol{\lambda}_i$ denote $[L \times 1]$ vectors of factors and factor loadings, respectively. As an example, suppose that Eq. (4) represents a model of earnings determination, where y_{it} denotes logged wage, and $\mathbf{x}_{it}$ includes variables such as level of education, experience, and tenure with the same employer. In this case, $\boldsymbol{\lambda}_i$ may absorb different unobserved skills for individual i, and $\mathbf{f}_t$ may capture the market values of such skills, which may vary temporally according to the business cycle of the economy. By contrast, the two-way effects model restricts the business cycle effect on wages (conditional on $\mathbf{x}_{it}$) to be identical across all individuals, regardless of their specific skill set.

It is worth pointing out that the common factor model nests the two way effects model; in particular, the latter is obtained by setting $L = 2$, $\boldsymbol{\lambda}_i = (\eta_i, 1)'$, $\mathbf{f}_t = (1, \tau_t)'$. Notice also that the common factor model can always be decomposed into a linear part (e.g. fixed effects) and a remaining nonlinear part. To see this, let $L = 1$ and define $\tilde{\lambda}_i = \lambda_i - \bar{\lambda}$, and $\tilde{f}_t = f_t - \bar{f}$. Then, one has

$$\lambda_i f_t = \tilde{\lambda}_i \tilde{f}_t + \eta_i + \tau_t + c, \tag{8}$$

where $\eta_i = \bar{f}\lambda_i$, $\tau_t = \bar{\lambda} f_t$, and $c = -\bar{\lambda}\bar{f}$. That is, Eq. (8) consists of two additive effects (with equal mean) plus a zero-mean multiplicative component. Therefore, the single-factor model already contains many features of the two way effects model.

Standard transformations employed for the additive effects model, such as the within transformation or first-differencing, are not capable of eliminating the common factor component. This implies that application of the FE estimator to a factor model may result in a biased estimate of $\boldsymbol{\beta}$, even if $\mathbf{x}_{it}$ is strictly exogenous with respect to ε_{it}.

On the other hand, since the unobserved components enter multiplicatively, estimation of structural parameters becomes more complicated and usually requires nonlinear procedures, unless additional assumptions are imposed in the data generating process (DGP).[13] Moreover, in large panels the incidental parameter problem typically manifests in both dimensions, and therefore bias correction can become cumbersome.

In addition to the use of the common factor approach as a tool to capture rich sources of unobserved heterogeneity, factor models have also been popular for characterising the co-movement of economic variables in high-dimensional data sets. High dimensionality

[12] Chudik and Pesaran (2015a) and Juodis and Sarafidis (2018) provide specialised treatments of this topic in panels with T large and panels with T fixed, respectively.

[13] See Juodis and Sarafidis (2019) for a description of issues arising with nonlinear estimation of common factor models when T is fixed.

brings new challenges, but also provides new insights into the advancement of econometric theory; see Bai and Wang (2016) for a recent overview of this literature.

3.3 Heterogeneous Slopes

Common practice in panel data analysis involves 'pooling' of the data, such that the slope coefficients are restricted to be homogeneous across individuals. There are two main benefits arising from pooling. First, more observations are available for the same set of parameters, which potentially improves the precision of the estimates and increases statistical power. Second, in many cases pooling can simplify derivation of asymptotic theory.

However, the slope parameter homogeneity restriction has often been rejected in empirical analyses and, as such, it has been called into question by some researchers.[14] The basic premise is that variables not included in the specification of the model, could also impact the partial effect of $\mathbf{x}_{it}$ on y_{it}. For instance, in a model of earnings determination (discussed in Section 3.2), the partial effect of an additional year of education on (logged) wage may vary across individuals with different levels of (unobserved) motivation, since the latter can reflect differences in academic performance.

A simple linear model with heterogeneous coefficients can be expressed as follows:

$$y_{it} = c + \boldsymbol{\beta}_i' \mathbf{x}_{it} + \eta_i + \varepsilon_{it}, \tag{9}$$

where $\boldsymbol{\beta}_i$ denotes a $[K_x \times 1]$ vector of heterogeneous partial effects. For large N, $\boldsymbol{\beta}_i$ can be treated as random variables with mean $\boldsymbol{\beta}$ and constant variance across i. When $\boldsymbol{\beta}_i$ is correlated with $\mathbf{x}_{it}$, the pooled FE estimator of the average partial effect, $\boldsymbol{\beta}$, is biased.[15]

Assuming strict exogeneity of $\mathbf{x}_{it}$ with respect to ε_{it}, $\boldsymbol{\beta}$ can be estimated consistently as $N \to \infty$, using the unweighted mean of $\widehat{\boldsymbol{\beta}}_i$, where $\widehat{\boldsymbol{\beta}}_i$ denotes the least squares estimate of $\boldsymbol{\beta}_i$. The resulting estimator is simply defined as $\widehat{\boldsymbol{\beta}}_{MG} = N^{-1} \sum_{i=1}^{N} \widehat{\boldsymbol{\beta}}_i$, and is known as the Mean Group (MG) estimator.[16] Intuitively, the desirable asymptotic properties of the MG estimator hint upon the fact that each estimate of $\boldsymbol{\beta}_i$ is unbiased, and therefore the estimation error associated with $\widehat{\boldsymbol{\beta}}_i$ tends to average out as N grows large.[17]

If some of the regressors are weakly exogenous, i.e. $E(\varepsilon_{it} | \mathbf{x}_{i1}, ..., \mathbf{x}_{it}) = 0$, as it is the case in dynamic panels and models with feedback, $\widehat{\boldsymbol{\beta}}_i$ is biased when T is fixed. Therefore, identification of $\boldsymbol{\beta}$ is not possible in general, unless restrictive assumptions are imposed on the data generating process.[18] However, as T grows large, the bias of $\widehat{\boldsymbol{\beta}}_i$ vanishes and therefore the average partial effect can be identified. In particular, the MG estimator is

[14] See Baltagi, Bresson and Pirotte (2008) for a useful overview on this topic.

[15] Test statistics for the null hypothesis of slope parameter homogeneity have been proposed by Pesaran, Shin and Smith (1996), Phillips and Sul (2003), Pesaran and Yamagata (2008) and Blomquist and Westerlund (2013), among others. Campello, Galvao and Juhl (2019) introduce a method for measuring the magnitude of the slope parameter heterogeneity bias of the FE estimator.

[16] See Chamberlain (1982). Recently, Arellano and Bonhomme (2012) extended the MG approach, studying identification and estimation of higher order moments of the distribution of the heterogeneous partial effects.

[17] Notice that the MG approach is not feasible when $T \leq K_x$.

[18] See Chamberlain (1993). Some limited counter-examples are discussed by Arellano and Honore (2001). An interesting case is a heterogeneous AR(1) panel data model with no individual effects.

consistent and asymptotically normal as $\sqrt{N}/T \to 0$.[19]

Recently, there is increasing interest among researchers in modelling slope heterogeneity using group structures. Under this framework, the slope parameters are restricted to be homogeneous within groups of individuals, but are allowed to vary freely across groups. In this case, the basic linear panel data model can be expressed as in Eq. (9), except that the slopes are restricted to

$$\boldsymbol{\beta}_i = \sum_{\ell=1}^{M} \boldsymbol{\theta}_\ell \mathbf{1}\left\{i \in \mathcal{G}_\ell\right\}, \tag{10}$$

where $\boldsymbol{\theta}_\ell \neq \boldsymbol{\theta}_{\ell'}$ for any $\ell \neq \ell'$, and $\mathcal{G} \equiv \{\mathcal{G}_1, ..., \mathcal{G}_M\}$ denotes a partition of the set $\{1, ..., N\}$. In comparison to the complete homogeneous model in Eq. (2), group structures have the advantage of allowing for some (partial) heterogeneity in the slope coefficients, and hence they are less restrictive. In comparison to the fully heterogeneous model, group structures share the benefit arising from pooling the data, i.e. more observations are available to estimate the slope coefficients.

If the number of groups, M, as well as the true partition/membership of individuals into groups are both known, the problem reduces to a split-sample standard panel data regression, which is straightforward enough to estimate.

More challenging is the problem of determining the optimal partition and the optimal number of groups, jointly with $\boldsymbol{\theta}_\ell$, $\ell = 1, ..., M$. A popular approach for estimating group structures involves minimum within-group sums of squares partitioning. The resulting 'group fixed effects' (GFE) estimator can be expressed as the minimiser of the following objective function:[20]

$$\left(\hat{\boldsymbol{\theta}}_{GFE}, \widehat{\mathcal{G}}\right) = \underset{(\boldsymbol{\theta},\mathcal{G})\in\Theta\times\Theta_{\mathcal{G}}}{\arg\min} \sum_{i=1}^{N}\sum_{t=1}^{T} \left(\tilde{y}_{it} - \boldsymbol{\beta}_i \tilde{\mathbf{x}}_{it}\right)^2, \tag{11}$$

where $\boldsymbol{\theta} = (\boldsymbol{\theta}_1', ..., \boldsymbol{\theta}_M')'$, Θ denotes the full parameter space of $\boldsymbol{\theta}$ (and similarly for $\Theta_{\mathcal{G}}$ in terms of $\mathcal{G}$), while $\tilde{y}_{it}$ and $\tilde{\mathbf{x}}_{it}$ denote observations expressed in terms of deviations from individual-specific averages. For fixed T, the GFE estimator of $\boldsymbol{\theta}$ is consistent and asymptotically normal for a pseudo true value, $\mathring{\boldsymbol{\theta}}$. This pseudo true value, which minimises an expected within-group sum of squared residuals, does not necessarily coincide with the true value of the parameter (Bonhomme and Manresa (2015)). Intuitively, this is because T observations are available upon which to determine membership of individual i to one of the M groups.[21] Notwithstanding, the true value of M can still be estimated consistently for T fixed, using a BIC-type criterion (Sarafidis and Weber (2015)). Alternative methods, based on different objective functions, have also been explored in cases where both N and T tend to infinity. For instance, Lin and Ng (2012) put forward a pseudo threshold approach, which uses the time series estimates of the individual slope coefficients to form threshold variables; Su, Shi and Phillips (2016) develop a group-Lasso approach that serves to shrink individual coefficients to the unknown group-specific coefficients; Liu,

[19] See Pesaran and Smith (1995), and Hsiao, Pesaran and Tahmiscioglu (1999), among others. Pesaran (2006) puts forward MG estimation of large heterogeneous panels with common factors.

[20] For most practical applications in economics, it is infeasible to search over all possible partitions. Therefore, heuristic algorithms are employed in optimisation. The most popular algorithm is known as 'kmeans clustering'. See Lin and Ng (2012) for a useful discussion on the pros and cons of this algorithm.

[21] However, Bonhomme and Manresa (2015) show that in the specific model they consider, the difference between the true value of $\boldsymbol{\theta}$ and $\mathring{\boldsymbol{\theta}}$ vanishes quickly as T increases.

Shang, Zhang and Zhou (2019) study M-estimation of panel data models with group structures under unknown number of groups.[22]

3.4 Panels with Multiple Dimensions or Multiple Levels

The rapid emergence of big datasets has fuelled a burgeoning literature on the analysis of panel data with multiple dimensions or multiple levels.

Simply put, multi-dimensional panel data refer to data containing repeated observations over two or more dimensions. To illustrate, a three-dimensional linear panel data model can be expressed as follows:

$$y_{ijt} = \boldsymbol{\beta}'\mathbf{x}_{ijt} + u_{ijt};\ i = 1, ..., N;\ j = 1, ..., J;\ t = 1, ..., T. \tag{12}$$

Prominent examples of the specification above are models of economic flows, such as a 'gravity model' of international trade, where y_{ijt} typically denotes some measure of volume of trade from country i to country j at time t, and $\mathbf{x}_{ijt}$ contains variables such as the relative size of the two countries, the real exchange rate etc.[23] An important case of a panel with multiple dimensions is a network model, where i and j in (12) are exchangeable. Exchangeability implies that one can swap around indices i and j without changing the distribution of the data.

The extra dimension of the data allows one to extend the two-way effects model in several directions, and therefore to capture additional sources of unobserved heterogeneity. One possibility is to specify a 'three-way' effects model, such that the regression error term in Eq. (12), u_{ijt}, becomes equal to[24]

$$u_{ijt} = \eta_i + \gamma_j + \tau_t + \varepsilon_{ijt}. \tag{13}$$

For example, in gravity models η_i denotes the unobserved effect of the origin country, γ_j denotes the unobserved effect of the destination country, and τ_t is the usual common time effect. Another possibility is to set

$$u_{ijt} = \delta_{ij} + \tau_t + \varepsilon_{ijt}, \tag{14}$$

where δ_{ij} denotes a country-pair effect, i.e. an interaction between unobserved origin-country and destination-country characteristics.[25] It is worth noting that the standard within transformation employed in the two-way fixed effects model is sufficient to eliminate the unobserved effects in both (13) and (14); however this transformation is not optimal, i.e. the resulting FE estimator is not efficient.

A specification that encompasses both (13) and (14) is given by[26]

$$u_{ijt} = \delta_{ij} + \theta_{it} + \psi_{jt} + \varepsilon_{it}, \tag{15}$$

where θ_{it} denotes i-specific time-varying effects, such as the origin country's business cycle, its cultural, political, or institutional characteristics, as well as unobserved factor

[22] See also Ando and Bai (2016), who study group structures in factor models with unknown group membership.

[23] Thus, in this case it is assumed that $i \neq j$ and $N = J$.

[24] See e.g. Mátyas (1997).

[25] See Egger and Pfaffermayr (2003) and Cheng and Wall (2005).

[26] See Baltagi et al (2003), and Aghion et al. (2008), among others.

endowment variables. Likewise, ψ_{jt} accounts for similar influences, except they correspond to the destination country. Optimal transformations for all three specifications above (as well as additional ones) are analysed by Balazsi, Mátyas and Wansbeek (2017).

In a nutshell, data with multiple dimensions offer the ability for practitioners to capture additional sources of unobserved heterogeneity, compared to the usual two-way effects model. A within transformation that is optimal for a particular unobserved effects specification, such as the one in Eq. (14), may not be robust to more general specifications, such as the one in Eq. (15), thus leading to a biased FE estimator. On the other hand, a robust transformation that controls for a more general specification, such as that in Eq. (15), may not be optimal when the true data generating process is given by (13), thus leading to an inefficient FE estimator. One major challenge is to use a FE estimator that is both unbiased and efficient.

In addition to panels with multiple dimensions, there is also a vibrant literature on panel data models with multiple levels, also known as 'hierarchical' or 'nested' models.[27] An important distinction between multi-level and multi-dimensional models is that in the former case the observations are nested; that is, knowledge of the value of i implies knowledge of the value of j. For instance, in a multi-level model of earnings determination, y_{ijt} may denote logged wage of individual i, employed in sector j.[28] By constrast, multi-dimensional models are non-nested in that knowledge of i does not imply knowlege of j. An implication of nesting is that one cannot include fixed effects for both i and j, unlike e.g. Eq. (13). That is, η_i and γ_j cannot be separately identified because they are collinear. This property also carries implications for more sophisticated error structures, such as unobservables with interaction terms.

Last, it is worth pointing out that identification and estimation of nonlinear panel data models with multiple dimensions may be far more complicated compared to the linear model. For instance, even in those rare instances where a sufficient statistic for the (multiple) additive effects exists, optimising the conditional likelihood can be computationally challenging (Charbonneau (2017)).

4 Contributions Made in This Special Issue

The large majority of articles appearing in this special issue deal with the challenges discussed in the previous section. As a collection, these articles paint a state-of-the-art picture of the field. Below we summarise the contributions of each paper. To facilitate exposition, we have grouped papers together according to the main area of contribution, although we note that some papers contribute to multiple areas.

4.1 Nonlinear Models

Second-order corrected likelihood for nonlinear panel models with fixed effects (Dhaene and Sun, 2020)

Dhaene and Sun propose second-order bias correction for static nonlinear panel data models with fixed effects. The correction is made via the log-likelihood function, and

[27] The literature on multi-level models dates back at least to the seminal paper by Fuller and Battese (1973). A good overview of this literature is Raudenbush and Bryk (2002).

[28] This simple definition implies that individual i does not switch between sectors at different points in time. Otherwise, one can view the pair of indices (i, t) as being nested in j.

removes the two leading terms of the bias of the log-likelihood, arising from estimating the fixed effects. Existing methods based on analytical corrections, reduce the bias of the fixed effects estimator from $O(T^{-1})$ to $O(T^{-2})$ (e.g. Arellano and Hahn (2016)). However when T is small, the $O(T^{-2})$ term may still be non-negligible. Indeed, simulation exercises based on logit and probit models show that the second-order correction dominates the first-order correction for all $T \geq 3$ uniformly over all designs examined. This outcome indicates that second-order corrections may already improve on first-order corrections for very small values of T, which can be highly beneficial in empirical applications.

Semiparametric identification in panel data discrete choice models (Aristodemou, 2020)

This paper provides new results on semiparametric identification of dynamic binary response and static ordered response panel data models with fixed effects. It is shown that under mild distributional assumptions on the fixed effect and the time-varying unobservables, informative bounds on the regression coefficients can be derived even if point identification fails. Partial identification is achieved essentially by finding features of the distribution that are independent from the fixed effect. In particular, in the dynamic binary response setting, identification of the regression coefficients relies on individuals who switch in two consecutive time periods, conditional on their initial state. In the static ordered response setting, in addition to the individuals who switch from one period to the next, individuals who choose the 'in-between' category in two consecutive periods also provide a useful source of identification. As a result, tighter bounds can be potentially achieved in this case.

Identifying latent group structures in nonlinear panels (Wang and Su, 2020)

Wang and Su develop estimation and inference procedures for nonlinear panel data models with a group structure, when both $N, T \to \infty$. Specifically, slope parameters are assumed to be homogeneous within groups of individuals but vary freely across groups. The total number of groups and the true membership of individuals into groups are both treated as unknown. To identify the group structure, a variant of the sequential binary segmentation algorithm of Bai (1997) is developed, motivated from the CART-split criterion (Breiman, Friedman, Stone, and Olshen, 1984). This enables classification even if there is no natural ordering of the individual-specific estimates of the slope coefficient vectors across i. Existing extensions of the sequential binary segmentation approach for identification of latent group structures, such as Ke, Li and Zhang (2016), are available for linear models only, and deal with classification of scalar parameters, in which a natural ordering exists. The proposed approach identifies the true latent group structure with probability approaching one as the sample size increases. Moreover, the resulting post-classification QMLE estimator is shown to be asymptotically equivalent to the QMLE estimator that assumes knowledge of group membership and of the total number of groups.

4.2 Common Factor Models

Nonlinear factor models for network and panel data (Chen, Fernández-Val and Weidner, 2020)

This paper studies fixed effects estimation of a class of nonlinear single-index models, such as the logit, probit, ordered probit and Poisson specifications, when both dimensions of the panel grow large. The paper makes a major step from Fernandez-Val and Weidner

(2016), which restricts the unobserved effects to enter in an additive fashion, by allowing for a common factor structure in the residuals. This is particularly appealing in panels with network data, since common factors capture essential features of network formation, such as homophily and clustering. The proposed fixed effects estimator of the slope parameters and average partial effects is consistent and asymptotically normal but might suffer from incidental parameter bias. It is shown that the bias grows proportionally with the number of factors. Both analytical and split-sample corrections are developed for inference purposes.

On the robustness of the pooled CCE estimator (Juodis, Karabıyık and Westerlund, 2020)

Juodis, Karabıyık and Westerlund study the asymptotic properties of the pooled common correlated effects (PCCE) estimator of Pesaran (2006) in a model with weakly exogenous regressors and more cross-sectional averages than unobserved factors. Under proportional asymptotics on N and T, it is shown that the asymptotic distribution of PCCE contains bias terms of order proportional to N and T. Several approaches to bias-correction are examined using simulated data. Specific emphasis is placed on the role of the so-called rank condition. In particular, in a setup where the number of cross-sectional averages employed is larger than the total number of identifiable factors in the covariates, it is shown that the asymptotic distribution of the PCCE estimator is not mixed-normal, in general. The main conclusion is that while asymptotic normality seems fragile, consistency is less of an issue. Furthermore, inclusion of too many cross-sectional averages can be very costly, an insight not previously documented in the literature.

Estimating and testing high dimensional factor models with multiple structural changes (Baltagi, Ka and Wang, 2020)

Motivated by recent literature on the analysis of macroeconomic and financial indicators under severe disruptions, such as the 2007-09 'Great Recession', Baltagi, Kao and Wang study estimation and testing of structural breaks in high-dimensional factor models. The proposed approach allows inference on the presence and number of structural breaks under unknown breakpoint dates. The number of factors may vary across different regimes, which is an important empirical scenario.[29] The method builds upon the fact that a single-factor model with one structural break in the loadings is observationally equivalent to a model with two 'pseudo' factors but no breaks.[30] Moreover, the second moment matrix of the pseudo factors is subject to changes in exactly the same points as the breaks occuring in the loadings. This is crucial because the true factors are unobservable and not estimable without knowledge of the change points in the pseudo factors. Once consistent estimates of the change points are obtained, the number of factors and factor space is estimable in each regime. The paper develops tests for the null of no break vs ℓ breaks, and the null of ℓ breaks vs $\ell + 1$ breaks.

Predicting the VIX and the Volatility Risk Premium: The Role of Short-run Funding Spreads Volatility Factors (Andreou and Ghysels, 2020)

Traditionally, the extraction of risk factors has been confined to a particular asset class each time.[31] Andreou and Ghysels put forward a new approach that allows extracting

[29] See e.g. Stock and Watson (2012).

[30] See also Baltagi et al (2017) and Section 2.1 in Zhu et al (2019) for more details.

[31] For instance, Fama-French factors are extracted from cross-sections of stock returns, which are meant to price equity risk, but not (say) bonds or commodities returns.

volatility factors jointly from several types of economic indicators and different asset classes, such as assets with traded options or high-frequency intraday data. This is appealing because estimated factors from different asset classes may capture different information content, especially during highly volatile periods. The proposed procedure starts by collecting a large panel of asset returns or spreads; for each individual series, a standard ARCH-type volatility model is fitted on the estimated idiosyncratic component of spreads, giving rise to a panel of 'filtered volatilities'. Subsequently, common volatilty factors are extracted using principal components analysis. The combination of volatility filtering and principal components relates to the class of affine diffusions, often used in theoretical asset pricing models. Since filtered volatilities may contain measurement error, the paper employs two alternative IV methods to estimate the factor space consistently in the presence of measurement error. The theoretical properties of such procedure are studied in detail.

4.3 Heterogeneous Slopes

Estimation of heterogeneous panels with systematic slope variations (Breitung and Salish, 2020)

Breitung and Salish study panels with heterogeneous coefficients and additive effects, as in Eq. (9), assuming the regressors are strictly exogenous. The heterogeneous coefficients are decomposed into a systematic part and a remainder (random part), such that the latter is eventually absorbed by the error term. As in Mundlak's (1978) correlated random coefficients (CRC) framework, the systematic part is allowed to be correlated with the regressors. It is shown that the resulting CRC estimator is more efficient than Mean Group, particularly when the variation of the covariates across i is large, and/or the variation of the random part of the heterogeneous coefficients is relatively small. A further advantage of the proposed CRC estimator is that it is relatively robust to the case where the regressors corresponding to the parameters of interest vary little over time.[32] By contrast, the crude MG estimator can perform poorly under these circumstances.[33] The paper also develops two tests statistics for systematic slope parameter heterogeneity using the Lagrange Multiplier and Hausman test principles.

Instrumental variable estimation of dynamic linear panel data models with defactored regressors and a multifactor error structure (Norkute, Sarafidis, Yamagata and Cui, 2020)

Norkute, Sarafidis, Yamagata and Cui develop an instrumental-variables approach for dynamic panels with exogenous covariates and a multifactor error structure, under large N and T asymptotics. The main idea entails (i) using principal components analysis to project out the common factors from the exogenous covariates, and (ii) constructing instruments from defactored covariates. The papers puts forward two IV estimators for models with homogeneous and heterogeneous coefficients. The proposed estimators are linear, and therefore computationally robust and inexpensive. Moreover, they are asymptotically unbiased as both N,T diverge such that $N/T \to c$. By contrast, available estimators extending the so-called CCE and PC approaches of Pesaran (2006) and Bai (2009) to dynamic panels, suffer from incidental parameter bias, depending on the size

[32] This scenario is particularly relevant when the covariates are binary, such as those describing marital status, union membership etc.

[33] Graham and Powell (2012) study identification and estimation of average partial effects when the values of the regressors vary little over time for a subset of the sample.

of T and the true parameter values of the DGP. Simulation evidence shows that this can lead to severe size distortions for these estimators.

Heterogeneous structural breaks in panel data models (Okui and Wang, 2020)

Okui and Wang put forward a new method for testing for structural breaks in models with heterogeneous coefficients. Identification is achieved by imposing a group pattern of slope parameter heterogeneity, as in Eq. (10). In particular within each group, structural breaks are assumed to be common, whereas the number, timing and size of structural breaks can be different across groups. This allows, for example, some structural breaks to affect only a subset of the population. The proposed approach combines shrinkage estimation via an adaptive grouped fused lasso, as proposed by Qian and Su (2016), with minimum within-group sums of squares partitioning, as advocated e.g. in Lin and Ng (2012), Bonhomme and Manresa (2015) and Sarafidis and Weber (2015). The method complements existing state-of-the-art literature, such as Baltagi, Feng and Kao (2016), who consider the case of heterogeneous structural breaks occuring at the same point in time, and Su, Wang and Jin (2019), who study group structural instability that takes the form of continuous time-varying slope coefficients.

Inferential theory for heterogeneity and cointegration in large panels (Trapani, 2020)

Trapani proposes a new estimation and testing framework to assess the presence and the extent of slope heterogeneity and cointegration when the units are a mixture of spurious and/or cointegrating regressions. Method of Moments estimators are developed to estimate the degree of heterogeneity (measured by the dispersion of the slope coefficients around their average), and the fraction of spurious regressions. It is shown that both estimators are consistent across the whole parameter space. Based on this result, two tests for the null hypotheses of slope homogeneity and cointegration are developed. The test for slope homogeneity permits the possibility that some individual time series are not cointegrated due to (say) the presence of neglected nonlinearities in the DGP. By contrast, existing tests require that all individual time series are cointegrated.[34] In addition, the test for cointegration remains valid regardless of the extent of slope heterogeneity, and also allows for cross-sectional dependence via a common factor component.

4.4 Panels with Multiple Dimensions or Multiple Levels

Estimation and inference for multi-dimensional heterogeneous panel datasets with hierarchical multi-factor error structure (Kapetanios, Serlenga and Shin, 2020)

Kapetanios, Serlenga and Shin extend the common correlated effects estimator by Pesaran (2006) to three-dimensional panel data models. The proposed approach generalises existing multi-dimensional panel data literature in that it allows for heterogeneous slope coefficients and strong cross-sectional dependence. This is attractive because multi-dimensional panels, such as those involving network data, are often interdependent by construction. The common factor structure considered in the paper takes a hierarchical form, which distinguishes between 'global' and 'local' factors. The former set of factors hits both i and j units, whereas the latter hits either i or j only. Special cases with homogeneous slope coefficients and homogeneous factors are also examined. The paper develops a pooled CCE estimator and a modified Mean Group estimator, coupled with a new nonparametric estimator for the variance of the modified MG estimator.

[34] See Mark and Sul (2003) and Westerlund and Hess (2011).

An econometric approach to the estimation of multi–level models (Yang and Schmidt, 2020)

Yang and Schmidt establish new theoretical results on nested panel data models with time-invariant regressors, and both fixed and random effects. The paper provides an exhaustive list of the instruments available to this model based on the Hausman-Taylor (1981), Amemiya-MaCurdy (1986) and Breusch-Mizon-Schmidt (1989) IV approaches. Existing applications of Hausman-Taylor methods to the multi-level model, such as Kim and Frees (2007), do not identify all of the relevant instruments and therefore they do not yield asymptotically efficient estimators. In addition, the paper analyses estimation with weakly exogenous and endogenous regressors and discusses the case where conditional homoskedasticity is violated. Furthermore, a Hausman-type test for exogeneity is derived, using a simple variable addition approach.

4.5 Additional contributions

Detecting granular time series in large panels (Brownlees and Mesters, 2020)

Brownlees and Mesters' work builds upon the so-called 'granular hypothesis' (Gabaix (2011)). This postulates that a significant portion of aggregate economic fluctuations are attributable to idiosyncratic shocks on the 'grains' of economic activity, such as a relatively small number of large firms. An important question is how to determine which firms (observed over a period of time) are granular, and how many granular firms exist. The paper formulates the granular detection problem as an *observed* factor model. In particular, it is shown that the column norms of the concentration matrix corresponding to granular series are larger than those for non-granular ones. This implies some ranking of the series, according to the value of their column norm. Moreover, the ratio between ordered column norms is maximized when the column norm of the last granular is divided by the first non-granular series. The resulting statistic selects the true granular series with probability one, as both $N, T \to \infty$. The proposed approach remains valid when the series are hit by additional (unobserved) factors, so long as the signal-to-noise ratio of granular shocks is sufficiently large.

Estimation of a nonparametric model for bond prices from cross-section and time series information (Koo, La Vecchia and Linton, 2020)

Koo, La Vecchia and Linton develop a new methodology for nonparametric estimation of time-varying yield curves using bond prices and their promised cash flows, from panel data with discrete time. The novelty of the proposed approach lies in the combination of two different techniques: cross-sectional nonparametric methods, and kernel estimation for time-varying dynamics in the time series context. Since bond prices and cash flows have a panel data structure, issues such as cross-sectional dependence and temporal dependence naturally arise. The method allows for general forms of cross-sectional and weak temporal dependence in the errors. Moreover, a new variance-covariance estimator for slowly time-varying yield curves is developed, which is consistent under quite general conditions. This paper extends Lee and Robinson (2016), who provide asymptotic theory for series estimation of nonparametric and semiparametric regression models for cross-sectional data under conditions that allow for some form of cross-sectional dependence and heterogeneity in the errors.

Dynamic Panels with MIDAS Covariates: Nonlinearity, Estimation and Fit (Khalaf, Kichian, Saunders and Voia, 2020)

Khalaf, Kichian, Saunders and Voia extend the Mixed Data Sampling (MIDAS) framework, which was first proposed by Ghysels, Santa-Clara and Valkanov (2004) in time series analysis, in the context of panel data analysis. Existing procedures for time series data are not directly applicable due to the dual-indexing of the observations. The proposed approach builds upon the fact that for a fixed value of $\boldsymbol{\theta}$, where $\boldsymbol{\theta}$ denotes the parameter vector associated with the MIDAS aggregation scheme, the corresponding MIDAS regressor becomes an observable aggregation of the high-frequency series. Hence, estimation reverts to a standard context where two statistics are typically available: a criterion to test the significance of the slope coefficient, β, given $\boldsymbol{\theta}$, and a diagnostic test to assess the specification of the model, given $\boldsymbol{\theta}$. The proposed approach constructs a confidence set for $\boldsymbol{\theta}$, by collecting the values that are not rejected by the diagnostic test at the desired level of significance. Subsequently, it puts forth two bound tests for β, based on supremum p-value over the confidence set for $\boldsymbol{\theta}$, or over its entire parameter space. The procedure allows for the possibility of an empty confidence set for $\boldsymbol{\theta}$, which signals model misspecification.

5 Acknowledgements

We are indebted to the authors for making our enterprise successful by delivering excellent papers. We appreciate their cooperation and flexibility in the review process, which often involved nontrivial operations. We have greatly benefited from the assistance of 44 referees, who provided expert advice. We would like to thank the Editor, Oliver Linton, for his support and encouragement to complete this project. Anastasios Panagiotelis, as well as a number of authors in this issue, provided useful feedback on this manuscript. Connie Brown provided excellent secretarial assistance throughout the elaborate editorial process. As stated earlier, this special issue arose out of the 2017 International Panel Data Conference, which was hosted by the University of Macedonia in Thessaloniki, Greece. For the successful organisation of this conference we are eternally grateful to Theologos Pantelidis and Theodore Panagiotidis. This work was supported by the Australian Research Council (ARC) under research grant number DP-170103135.

References

Aghion, P., Burgess, R., Redding, S., Zilibotti, F., 2008. The Unequal effects of liberalization: evidence from dismantling the license Raj in India. Amer. Econ. Rev. 98 (4), 1397–1412.

Amemiya, T., MaCurdy, T.E., 1986. Instrumental variable estimation of an error component model. Econometrica 54, 869–881.

Andersen, E., 1970. Asymptotic properties of conditional maximum-likelihood estimators. J. R. Stat. Soc. Ser. B Stat.. Methodol. 32 (2), 283–301.

Anderson, T.W., Hsiao, C., 1981. Estimation of Dynamic Models with Error Components. J. Amer. Statist. Assoc. 76, 598–606.

Andreou, E., Ghysels, E., 2019. Predicting the VIX and the volatility risk premium: The role of short-run funding spreads volatility factors. J. Econometrics, forthcoming.

Ando, T., Bai, J., 2016. Panel data models with grouped factor structure under unknown group membership. J. Appl. Econometrics 31 (1), 163–191.

Arellano, E., Bond, S.R., 1991. Some specification tests for panel data: Monte Carlo evidence and an application to employment equations. Rev. Econ. Stud. 58, 277–298.

Arellano, M., Bonhomme, S., 2009. Robust priors in nonlinear panel data models. Econometrica 77 (2), 489–536.

Arellano, M., Bonhomme, S., 2011. Nonlinear panel data analysis. Annu. Rev. Econ. 2, 395–424.

Arellano, M., Bonhomme, S., 2012. Identifying distributional characteristics in random coefficients panel data models. Rev. Econom. Stud. 79 (3), 987–1020.

Arellano, M., Hahn, J., 2016. A likelihood-based approximate solution to the incidental parameter problem in dynamic nonlinear models with multiple effects. Global Economic Review 45 (3), 251–274.

Arellano, M., Honoré, B., 2001. Panel data: Some recent developments. In: Heckman, J.J., Leamer, E.E., (Eds.), Handbook of Econometrics, Volume 5, North Holland, pp. 3229–3296.

Aristodemou, E., 2020. Semiparametric identification in panel data discrete response models. J. Econometrics, forthcoming.

Bai, J., 1997. Estimating multiple breaks one at a time. Econometric Theory 13, 315–352.

Bai, J., 2009. Panel data models with interactive fixed effects. Econometrica 77, 1229–1279.

Bai, J., Wang, 2016. Econometric analysis of large factor models. Annu. Rev. Econ. 8, 53–0.

Balazsi, L., Mátyas, L., Wansbeek, T., 2017. Fixed effects models. In: Mátyas, L., (Eds.), The Econometrics of Multi-Dimensional Panels. Springer-Verlag, pp. 1–35.

Balestra, P., Nerlove, M., 1966. Pooling cross section and time series data in the estimation of a dynamic model: The demand for natural gas. Econometrica 34 (3), 585–612.

Baltagi, B.H., 2013. Econometric Analysis of Panel Data, fifth ed. John Wiley & Sons Ltd, Chichester.

Baltagi, B.H., Bresson, G., Pirotte, A., 2008. To pool or not to pool? In: Mátyas L., Sevestre, P., (Eds.), The Econometrics of Panel Data. Springer-Verlag, pp. 517–554.

Baltagi, B.H., Egger, P., Pfaffermayr, M., 2003. A generalized design for bilateral trade flow models. Econom. Lett. 80, 391–397.

Baltagi, B.H., Feng, Q., Kao, C., 2016. Estimation of heterogeneous panels with structural breaks. J. Econometrics (191), 176–195.

Baltagi, B.H., Ka, C., Wang, F., 2020. Estimating and testing high dimensional factor models with multiple structural changes. J. Econometrics, forthcoming.

Baltagi, B.H., Ka, C., Wang, F., 2017. Identification and estimation of a large factor model with structural instability. J. Econometrics 197, 87–100.

Bester, C., Hansen, C., 2009. A penalty function approach to bias reduction in nonlinear panel models with fixed effects. J. Bus. Econom. Statist. 27 (2), 131–148.

Blomquist, J., Westerlund, J., 2013. Testing slope homogeneity in large panels with serial correlation. Econom. Lett. 121, 374–378.

Bonhomme, S., 2012. Functional differencing. Econometrica 80 (4), 1337–1385.

Bonhomme, S., Manresa, E., 2015. Grouped patterns of heterogeneity in panel data. Econometrica 83, 1147–1184.

Breiman, L., Friedman, J., Stone, C.J., Olshen, R.A., 1984. Classification and Regression Trees. CRC Press.

Breitung, J., Salish, N., 2019. Estimation of heterogeneous panels with systematic slope variations. J. Econometrics, forthcoming.

Breusch, T.S., Mizon, G.E., Schmidt, P., 1989. Efficient Estimation Using Panel Data. Econometrica 57, 695–700.

Brownlees, C., Mesters, G., 2019. Detecting granular time series in large panels. J. Econometrics, forthcoming.

Bun, M., Sarafidis, V., 2015. Dynamic panel data models. In: Baltagi, B.H. (Eds.), The Oxford Handbook of Panel Data. Oxford University Press, pp. 76–110.

Campello, M., Galvao, A., Juhl, T., 2019. Testing for slope heterogeneity bias in panel data models. J. Bus. Econom. Statist. 37 (4), 749–760.

Chamberlain, G., 1982. Multivariate regression models for panel data. J. Econometrics 18, 5–46.

Chamberlain, G. 1985. Heterogeneity, omitted variable bias, and duration dependence. In: Heckman J.J. Singer, B. (Eds.), Longitudinal Analysis of Labor Market Data. Cambridge University Press.

Chamberlain, G., 1993. Feedback in panel data models. Unpublished manuscript.

Chamberlain, G., 2010. Binary response models for panel data: Identification and information. Econometrica 78, 159–168.

Chang, C.L., McAleer, M., Oxley, L., 2011. Great Expectatrics: Great Papers, Great Journals, Great Econometrics. Econometric Rev. 30 (6), 583–619.

Charbonneau, K., 2017. Multiple fixed effects in binary response panel data models. Econom. J. 20 (3), 1–13.

Chen, M., Fernández-Val, I., Weidner, M., 2020. Nonlinear factor models for network and panel data. J. Econometrics, forthcoming.

Cheng, I.-H., Wall, H., 2005. Controlling for heterogeneity in gravity models of trade and integration. Federal Reserve Bank of St. Louis Review 87, 49–63.

Chudik, A., and Pesaran, M. H., 2015a. Large panel data models with cross-sectional dependence: A survey. In: Baltagi, B.H. (Eds.), The Oxford Handbook of Panel Data. Oxford University Press, pp. 3–45.

Chudik, A., Pesaran, M.H., 2015b. Common correlated effects estimation of heterogeneous dynamic panel data models with weakly exogenous regressors. J. Econometrics 188, 393–420.

Dhaene, G., Jochmans, K., 2015. Split-panel jackknife estimation of fixed-effect models. Rev. Econ. Stud. 82 (3), 991–1030.

Dhaene, G., Sun, Y., 2020. Second-order corrected likelihood for nonlinear panel models with fixed effects. J. Econometrics, forthcoming.

Egger, P. Pfaffermayr, M., 2003. The Proper econometric specification of the gravity equation: 3-way model with bilateral interaction effects. Empir. Econ. 28, 571–580.

Fernández-Val, I., Weidner, M., 2016. Individual and time effects in nonlinear panel models with large N, T. J. Econometrics 192 (1), 291–312.

Fernández-Val, I., Weidner, M., 2018. Fixed effect estimation of large T panel data models. Working paper.

Fuller, W.A., Battese, G.E. 1973. Transformations for estimation of linear models with nested-error structure. J. Amer. Statist. Assoc. 68, 626–632.

Gabaix, X., 2011. The granular origins of aggregate fluctuations. Econometrica 79, 733–772.
Ghysels, E., Santa-Clara, P., Valkanov, R., 2006. Predicting volatility: getting the most out of return data sampled at different frequencies. J. Econometrics 131 (1), 59–95.
Graham, B.S., Powell, J.L., 2012. Identification and estimation of average partial effects in "irregular" correlated random coefficient panel data models. Econometrica 56, 2105–2152.
Greene, W. 2015. Panel data models for discrete choice. In: Baltagi, B.H. (Eds.), The Oxford Handbook of Panel Data. Oxford University Press, pp. 171–201.
Hahn, J., Kuersteiner, G., 2011. Bias reduction for dynamic nonlinear panel models with fixed effects. Econometric Theory 27 (6), 1152–1191.
Hahn, J., Newey, W., 2004. Jackknife and analytical bias reduction for nonlinear panel models. Econometrica 72 (4), 1295–1319.
Hausman, J.A., Taylor, W.E., 1981. Panel data and unobservable individual effects. Econometrica 49, 1377–1399.
Holtz-Eakin, D., Newey, W., Rosen, H.S., 1988. Estimating vector autoregressions with panel data. Econometrica 56, 1371–1396.
Hsiao, C., 2014. Analysis of Panel Data, third ed. Cambridge University Press, Cambridge.
Hsiao, C., Pesaran, M.H., Tahmiscioglu, A., 1999. Bayes estimation of short-run coefficients in dynamic panel data models. In: Hsiao, C., Lahiri, K., Lee, L.-F., Pesaran, M.H., (Eds.), Analysis of Panels and Limited Dependent Variables, Cambridge University Press, pp. 268–296.
Honoré, B. 1992. Trimmed LAD and least squares estimation of truncated and censored regression models with fixed effects. Econometrica 60 (3), 533–65.
Honoré, B., Tamer, E. 2006. Bounds on parameters in panel dynamic discrete choice models. Econometrica 74 (3), 611–629.
Juodis, A., Karabıyık, H., Westerlund, J., 2020. On the robustness of the pooled CCE estimator. J. Econometrics, forthcoming.
Juodis, A., Sarafidis. V., 2018. Fixed T dynamic panel data estimators with multifactor errors. Econometric Rev. 37 (8), 893–929.
Juodis, A., Sarafidis. V., 2019. A linear estimator for factor-augmented fixed-T panels with endogenous regressors. Working paper.
Kalbfleisch, J.D., Sprott, D.A., 1970. Application of likelihood methods to models involving large numbers of parameters. J. R. Stat. Soc. Ser. B Stat. Methodol. 32 (2), 175–208.

Kapetanios, G., Serlenga, L., Shin, Y., 2020. Estimation and inference for multi-dimensional heterogeneous panel datasets with hierarchical multi-factor error structure. J. Econometrics, forthcoming.
Ke, Y., Li, J., Zhang, W., 2016. Structure identification in panel data analysis. Ann. Statist. 44, 1193–1233.
Kim, J.-S., Frees, E.W., 2007. Multilevel Modeling with Correlated Effects. Psychometrika 72, 505–533.
Kim, M., Sun, Y., 2016. Bootstrap and k-step bootstrap bias corrections for the fixed effects estimator in nonlinear panel data models. Econometric Theory 32 (6), 1523–1568.

Koo, B., La Vecchia, D., Linton, O., 2019. Estimation of a nonparametric model for bond prices from cross-section and time series information. J Econometrics, forthcoming.

Lancaster, T., 2002. Orthogonal parameters and panel data. Rev. Econ. Stud. 69, 647–666.

Lee, J., Robinson, P.M., 2016. Series estimation under cross-sectional dependence. J. Econometrics 190 (1), 1–17.

Lin, C., Ng, S., 2012. Estimation of panel data models with parameter heterogeneity when group membership is unknown. J. Econometric Methods 1, 42–55.

Linnemer, L., Visser, M., 2017. The Most Cited Articles from the Top-5 Journals (1991–2015). Working paper.

Liu, R., Shang, Z., Zhang, Y., Zhou, Q., 2019. Identification and estimation in panel models with overspecified number of groups. J. Econometrics, forthcoming.

Manski, C., 1987. Semiparametric analysis of random effects linear models from binary panel data. Econometrica 55 (2), 357–362.

Manski, C., 1989. Anatomy of the Selection Problem. The Journal of Human Resources 24 (3), 343–360.

Mark, N.C., Sul, D., 2003. Cointegration vector estimation by panel dols and long-run money demand. Oxford B. Econ. Stat. 65 (5), 655–680.

Mátyas, L., 1997. Proper econometric specification of the gravity model. The World Economy 20, 363–369.

Molinari, F., 2019. Econometrics with partial identification. Working paper.

Moon, H. R., Weidner, M., 2017. Dynamic linear panel regression models with interactive fixed effects. Econometric Theory 33, 158–195.

Mundlak, Y., 1978. On the pooling of time series and cross section data. Econometrica 46 (1), 69–85.

Nerlove, M., 1978. Econometric analysis of longitudinal data: approaches, problems and prospects. Annales de l'inséé, No. 30/31, 7–22.

Neyman, J., Scott, E., 1948. Consistent estimates based on partially consistent observations. Econometrica 16 (1), 1–32.

Nickell, S., 1981. Biases in dynamic models with fixed effects. Econometrica 49 (6), 1417–1426.

Norkute, M., Sarafidis, V., Yamagata, T., Cui, G., 2020. Instrumental variable estimation of dynamic linear panel data models with defactored regressors and a multifactor error structure. J. Econometrics, forthcoming.

Okui, R., Wang, W., 2020. Heterogeneous structural breaks in panel data models. J. Econometrics, forthcoming.

Pesaran, M.H., 2006. Estimation and inference in large heterogeneous panels with a multifactor error structure. Econometrica 74, 967–1012.

Pesaran, H., Smith, R., 1995. Estimating long-run relationships from dynamic heterogeneous panels. J. Econometrics 68, 79–113.

Pesaran, H., Smith, R., Im, K.S., 1996. Dynamic linear models for heterogeneous panels. In: Mátyas L., Sevestre, P., (Eds.), The Econometrics of Panel Data: A Handbook of the Theory with Applications. Kluwer, pp. 145–195.

Pesaran, H., Yamagata, T., 2008. Testing slope homogeneity in large panels. J. Econometrics 142 (1), 50–93.

Phillips, P., Sul, D., 2003. Dynamic panel estimation and homogeneity testing under cross section dependence. Econom. J. 6 (1), 217–259.

Qian, J., Su, L., 2016. Shrinkage estimation of common breaks in panel data models via adaptive group fused lasso. J. Econometrics 191 (1), 86–109.
Raudenbush, S.W., Bryk, A.S., 2002. Hierarchical Linear Models: Applications and Data Analysis Methods, second ed. Sage Publications, Newbury Park.
Sarafidis, V., Wansbeek, T., 2012. Cross-sectional dependence in panel data analysis. Econometric Reviews 31 (5), 483–531.
Sarafidis, V., Weber, N., 2015. A partially heterogeneous framework for analyzing panel data. Oxford B. Econ. Stat. 77 (2), 274–296.
Stock, J.H., Watson, M.W., 2012. Disentangling the channels of the 2007-09 recession. Brookings Papers on Economic Activity, pp. 81–156.
Su, L., Shi, Z., Phillips, P., 2016. Identifying latent structures in panel data. Econometrica 84 (6), 2215–2264.
Su, L., Wang, X., Jin S., 2019. Sieve estimation of time-varying panel data models with latent structures. J. Bus. Econom. Statist. 37 (2), 334-349.
Sun, Y., 2005. Estimation and Inference in Panel Structure Models. Unpublished manuscript.
Trapani, L., 2019. Inferential theory for heterogeneity and cointegration in large panels. J. Econometrics, forthcoming.
Wang, W., Su, L., Identifying latent group structures in nonlinear panels. J. Econometrics, forthcoming.
Westerlund, J., Hess, W., 2011. A new poolability test for cointegrated panels. J. Appl. Econometrics 26 (1), 56-88.
Yang, Y., Schmidt, P., 2019. An econometric approach to the estimation of multi-level models. J. Econometrics, forthcoming.
Zhu, H., Sarafidis V., Silvapulle, M., 2019. A New structural Break Test for Panels with Common Factors. Econom. J., forthcoming. https://doi.org/10.1093/ectj/utz018.