1. Linear dynamic panel data models*

Ryo Okui

1. INTRODUCTION

Dynamic panel data models have been widely applied in many empirical studies across various fields of economics. Leading examples include labor economics (for example, Meghir and Pistaferri, 2004), economic growth (for example, Levine et al., 2000; Acemoglu et al., 2008), industrial organization (for example, Nickell 1996; Blundell and Bond, 2000) and international economics (for example, Clerides et al., 1998). This chapter provides a review of the estimation of linear dynamic panel data models, which in the past 30 years has become a very active research area in theoretical econometrics. As such, we now have a good understanding of the mathematical structure of this class of models and the statistical properties of the estimators. It is then sensible to provide an account of current best practice for analyzing these models.

The first half of this chapter reviews the generalized method of moments (GMM) estimation of linear dynamic panel data models. In this, we focus on a first-order autoregressive (AR) model with individual effects, called a panel AR(1) model as the most representative and simplest of the dynamic panel data models. We then consider the GMM estimator proposed by Arellano and Bond (1991) and Holtz-Eakin et al. (1988). While the GMM estimator is the most popular method to estimate dynamic panel data models, in certain circumstances it may not behave well. In particular, GMM may suffer from the problems of weak (Staiger and Stock, 1997; Stock and Wright, 2000) and many (Kunitomo, 1980; Morimune, 1983; Bekker, 1994; Donald and Newey, 2001; Newey and Windmeijer, 2009) instruments. Fortunately, a number of methods are available to alleviate these problems, for which we discuss solutions in situations specific to the GMM estimation of dynamic panel data models.

In the second half of the chapter, we consider the fixed effects estimator. While well known even in introductory econometrics textbooks as the representative estimation method for panel data analysis, it may not be possible to apply the fixed effects estimator to dynamic panel data analysis as often as the GMM estimator. In particular, the fixed effects estimator is not consistent when the length of the time series is fixed (Nickell, 1981), and even with moderately long time series, it is biased and this bias makes statistical inference problematic (Hahn and Kuersteiner, 2002). In this chapter, we discuss the source of this bias and the bias-correction methods provided by Hahn and Kuersteiner (2002) and Dhaene and Jochmans (2015), the outcome being that bias-corrected fixed effects estimators tend to have small variance and are useful when the length of the time series is moderately large.

Ultimately, the chapter identifies the best practices available for the estimation of dynamic panel data analysis based on the existing literature. As useful background, many graduate-level econometrics texts, research monographs and review articles discuss the estimation of dynamic panel data models at length. Wooldridge (2010) is an excellent and popular textbook on graduate-level microeconometrics, while Arellano (2003) is an outstanding monograph on

panel data analysis. Elsewhere, Bun and Sarafidis (2015) provide a recent summary focusing on GMM estimation, with a comprehensive but rather less detailed review available in Sarafidis and Wansbeek (2021). Further, the interested reader may easily implement many, although not all, of the methods discussed in this chapter using readily available econometrics software, including STATA, EViews, R and gretl. These all have built-in commands/functions to analyze dynamic panel data models, and readers can implement the GMM estimator and its variations with these pieces of software very easily. We provide a short empirical illustration at the end of this chapter using a dataset from Acemoglu et al. (2008).

2. SETTINGS AND MODELS

Panel data are data for multiple entities where we observe each at multiple times. A panel has the form

$$y_{it}, \quad i = 1, \dots, N, \quad t = 1, \dots, T,$$

where i denotes an observation unit such as individual, firm, country, and so on and t denotes a time period, such as a year. For example, y_{it} is the income of individual i at time t , and we have a random sample of many individuals over several time periods. Dynamic panel data models describe the dynamics of y_{it} .

In most parts of this chapter, we consider the following panel AR(1) model:

$$y_{it} = \mu_i + \alpha y_{i,t-1} + \varepsilon_{it},$$

where α is the autoregressive coefficient parameter, μ_i is the unobserved individual (fixed) effect for i and ε_{it} is an unobserved idiosyncratic error. The fixed effects, μ_i s, do not depend on t . The error term is assumed to have mean zero, $E(\varepsilon_{it}) = 0$ and to be serially uncorrelated. This panel AR(1) model is the simplest dynamic panel data model. In practice, more complex models could also be estimated. For example, dynamic panel data models often include regressors:

$$y_{it} = \mu_i + \alpha y_{i,t-1} + \beta' X_{it} + \varepsilon_{it},$$

where X_{it} is the vector of regressors and β is a vector of coefficients. Nonetheless, most of the theoretical issues in dynamic panel data models can be discussed under the simple panel AR(1) model. We thus consider the panel AR(1) model for most of the chapter.

3. GMM ESTIMATION

In this section we discuss the GMM estimation of panel AR(1) models. GMM is the most popular estimation method for dynamic panel data models. However, it may not behave well in certain practically relevant situations. In particular, it may face the problems of weak and many instruments. A large number of studies have considered these problems, including proposals to improve the performance of the GMM estimator.

3.1 Endogeneity Issue

We first discuss the motivation for GMM estimation. Why do we use GMM instead of least squares estimators? The answer is that least squares methods face an endogeneity problem and cannot provide a consistent estimator when T is fixed.

Recall that the model is

$$y_{it} = \mu_i + \alpha y_{i,t-1} + \varepsilon_{it}.$$

The parameter of interest is α . This model may not be estimated directly because of the presence of fixed effects. Note that μ_i and $y_{i,t-1}$ are correlated by the structure of the model and OLS does not yield a consistent estimate if we regard μ_i as part of the error term. If μ_i s are regarded as parameters, then the number of parameters tends to infinity as $N \rightarrow \infty$.

A solution is to eliminate the fixed effects by, for example, taking the first difference:

$$y_{it} - y_{i,t-1} = \alpha(y_{i,t-1} - y_{i,t-2}) + (\varepsilon_{it} - \varepsilon_{i,t-1}). \quad (1.1)$$

The OLS estimator applied to this transformed model is known as the first difference (FD) estimator:

$$\hat{\alpha}_{FD} = \frac{\sum_{i=1}^N \sum_{t=3}^T (y_{i,t-1} - y_{i,t-2})(y_{it} - y_{i,t-1})}{\sum_{i=1}^N \sum_{t=3}^T (y_{i,t-1} - y_{i,t-2})^2}.$$

Unfortunately, the FD estimator is not consistent as $N \rightarrow \infty$ when T is fixed. The regressor, $y_{i,t-1} - y_{i,t-2}$, and the error term, $\varepsilon_{it} - \varepsilon_{i,t-1}$, are correlated because $y_{i,t-1}$ and $\varepsilon_{i,t-1}$ are correlated. Note that while there are other ways to eliminate individual effects, this endogeneity problem arises for any transformation.

A solution to the endogeneity problem is to use instruments. Recall that an instrumental variable is correlated with the regression and orthogonal to the error term. Anderson and Hsiao (1981), the seminal paper on dynamic panel data models, observe that lagged dependent variables satisfy the requirement for instrumental variable. See also Anderson and Hsiao (1982). The Anderson–Hsiao estimator, denoted as $\hat{\alpha}_{AH}$, estimates (1) using $y_{i,t-2}$ as an instrument:

$$\hat{\alpha}_{AH} = \frac{\sum_{i=1}^N \sum_{t=3}^T y_{i,t-2}(y_{it} - y_{i,t-1})}{\sum_{i=1}^N \sum_{t=3}^T y_{i,t-2}(y_{i,t-1} - y_{i,t-2})}.$$

Note that $y_{i,t-2}$ and $\varepsilon_{it} - \varepsilon_{i,t-1}$ are uncorrelated under the assumption that ε_{it} is serially uncorrelated. Moreover, we expect that $y_{i,t-2}$ and $(y_{i,t-1} - y_{i,t-2})$ are correlated because both involve $y_{i,t-2}$.

3.2 The Arellano–Bond Estimator

In this section, we introduce the Arellano–Bond estimator, which is arguably the most popular estimator for dynamic panel data models, being an extension of the earlier Anderson–Hsiao estimator. The key observation is that lags other than $y_{i,t-2}$ can be used as instruments. If we use all the available lags as instruments, we obtain the Arellano–Bond estimator.

To define the Arellano–Bond estimator, we introduce some notations. Let Δ be the first difference operator. Define

$$\Delta \varepsilon_i = \begin{pmatrix} \varepsilon_{i3} - \varepsilon_{i2} \\ \dots \\ \varepsilon_{iT} - \varepsilon_{i,T-1} \end{pmatrix}, \quad \Delta y_i = \begin{pmatrix} y_{i3} - y_{i2} \\ \dots \\ y_{iT} - y_{i,T-1} \end{pmatrix}, \quad \text{and } \Delta y_{i,-1} = \begin{pmatrix} y_{i2} - y_{i1} \\ \dots \\ y_{i,T-1} - y_{i,T-2} \end{pmatrix}.$$

Note that $\Delta y_i = \alpha \Delta y_{i,-1} + \Delta \varepsilon_i$. Also, define

$$\mathbf{Z}_i = \begin{pmatrix} [y_{i1}] & 0 & \dots & 0 \\ 0 & [y_{i1}, y_{i2}] & \dots & 0 \\ \dots & \dots & \dots & 0 \\ 0 & \dots & 0 & [y_{i1}, \dots, y_{i,T-2}] \end{pmatrix}.$$

This is a matrix of instruments. The j -th row of $\mathbf{Z}_i$ is a vector of instruments for $y_{i,j+1} - y_{i,j}$. Arellano and Bond (1991) and Holtz-Eakin et al. (1988) propose to estimate α using GMM based on the following moment conditions:

$$E(\mathbf{Z}_i' \Delta \varepsilon_i) = 0.$$

We obtain the Arellano–Bond estimator by minimizing the following objective function:

$$\min_a \left(\frac{1}{N} \sum_{i=1}^n \mathbf{Z}_i' (\Delta y_i - a \Delta y_{i,-1}) \right)' W \left(\frac{1}{N} \sum_{i=1}^n \mathbf{Z}_i' (\Delta y_i - a \Delta y_{i,-1}) \right),$$

where W is a weighting matrix. The following matrix yields an efficient GMM estimator (Hansen, 1982):

$$\hat{W}^{opt} = \left(\frac{1}{n} \sum_{i=1}^n \mathbf{Z}_i \Delta \hat{\varepsilon}_i \Delta \hat{\varepsilon}_i' \mathbf{Z}_i' \right)^{-1},$$

where $\Delta \varepsilon_i = \Delta y_i - \hat{\alpha} \Delta y_{i,-1}$ and $\hat{\alpha}$ is a preliminary consistent estimator of α . The explicit form of the Arellano–Bond estimator is

$$\hat{\alpha}_{AB} = \left\{ \left(\sum_{i=1}^n \Delta y_{i,-1}' \mathbf{Z}_i \right) W \left(\sum_{i=1}^n \mathbf{Z}_i' \Delta y_{i,-1} \right) \right\}^{-1} \left(\sum_{i=1}^n \Delta y_{i,-1}' \mathbf{Z}_i \right) W \left(\sum_{i=1}^n \mathbf{Z}_i' \Delta y_i \right).$$

We may iterate the estimation of α and the computation of $\hat{W}^{opt}$. This iteration converges and the resulting GMM estimator avoids the arbitrariness of the choice of preliminary estimator for $\hat{W}^{opt}$ (Hansen and Lee, 2021).

The Arellano–Bond estimator is arguably the most popular estimator for dynamic panel data models. One reason is that it is consistent and asymptotically normal as $N \rightarrow \infty$ and can be computed easily with commonly available econometric software (for example, STATA, EViews). As a result, many know it simply as the GMM estimator. However, we do not recommend this because GMM is a generic estimation method and other estimators for dynamic panel data models build on GMM principles. Thus, calling it the GMM estimator can be confusing.

Nevertheless, the Arellano–Bond estimator may suffer from bias in certain situations. We now consider two situations in which the Arellano–Bond estimator does not behave well, namely, the weak instruments problem and the many moments problem. These two problems arise in ways specific to dynamic panel data models. We also review the possible solutions in the literature.

3.3 Weak Instruments Problem

The Arellano–Bond estimator suffers from a weak instruments problem when the data are persistent. We first discuss the source of the weak instruments problem and then discuss the System-GMM estimator that may help attenuate the problem.

An instrument is called “weak” if it is only weakly correlated with the regressor. An instrumental variables estimator based on weak instruments suffers from bias and its distribution is not well approximated by a normal distribution (Staiger and Stock, 1997; Stock and Wright, 2000). In our context, $y_{i,t-1}$ is an instrument and the regressor is $y_{it} - y_{i,t-1}$. When the process is close to a unit root (that is, α is close to 1), the correlation between $y_{i,t-1}$ and $y_{it} - y_{i,t-1}$ is small.

Here, we show that the correlation between $y_{i,t-1}$ and $y_{it} - y_{i,t-1}$ tends to zero as $\alpha \uparrow 1$. Note that it is difficult to compute the correlation when $\alpha = 1$ because the variance of $y_{i,t-1}$ may not be well defined under a unit root unless we specify the initial condition. We thus assume stationarity and take α to be one. The discussion here follows that of Blundell and Bond (1998). We assume that μ_i and ε_{it} are independent, the sample is independently and identically distributed (i.i.d.) across i and $\text{Var}(\varepsilon_{it}) = \sigma_\varepsilon^2$ (homoscedasticity). Let $E(\mu_i^2) = \sigma_\mu^2$. Observe that

$$\begin{aligned} E(y_{i,t-1}(y_{it} - y_{i,t-1})) &= E(y_{i,t-1}((\alpha - 1)y_{i,t-1} + \mu_i + \varepsilon_{it})) \\ &= -(1 - \alpha)E(y_{i,t-1}^2) + E(y_{i,t-1}\mu_i). \end{aligned}$$

If y_{it} is stationary, then we have

$$y_{it} = \frac{\mu_i}{1 - \alpha} + \sum_{j=0}^{\infty} \alpha^j \varepsilon_{i,t-j}.$$

It follows that

$$\begin{aligned} &-(1 - \alpha)E(y_{i,t-1}^2) + E(y_{i,t-1}\mu_i) \\ &= -\frac{(1 - \alpha)}{(1 - \alpha)^2}E(\mu^2) + \frac{1}{1 - \alpha}E(\mu^2) - (1 - \alpha)E\left(\left(\sum_{j=0}^{\infty} \alpha^j \varepsilon_{i,t-j}\right)^2\right) \\ &= -\frac{\sigma_\varepsilon^2}{1 + \alpha}. \end{aligned}$$

Thus, the variance of $y_{i,t-1}$ is

$$\frac{\sigma_\mu^2}{(1-\alpha)^2} + \frac{\sigma_\varepsilon^2}{1-\alpha^2},$$

and the variance of $(y_{it} - y_{i,t-1})$ is

$$2\sigma_\varepsilon^2 + \frac{2}{1+\alpha}\sigma_\mu^2.$$

Hence, the correlation coefficient is

$$-\frac{\sqrt{(1-\alpha)^2}}{1+\alpha} \frac{\sigma_\varepsilon^2}{\sqrt{2\sigma_\varepsilon^2 + \frac{2}{1+\alpha}\sigma_\mu^2} \sqrt{\sigma_\varepsilon^2 + \frac{1-\alpha}{1+\alpha}\sigma_\mu^2}},$$

which tends to 0 as $\alpha \uparrow 1$. We note that the weak instruments problem arises not because the covariance becomes zero (note that $-\sigma_\varepsilon^2/2$ is nonzero). It occurs because the variance of the instruments diverges. This is in contrast to standard theoretical analysis of the weak instruments problem, although in the end we encounter the same problem and the Arellano–Bond estimator does not behave well.

We now discuss solutions to the weak instruments problem. The most commonly used solution is to employ an additional set of moment conditions that are relatively robust to persistent data. The System-GMM estimator proposed by Arellano and Bover (1995) and Blundell and Bond (1998) uses this set of moment conditions in addition to those used in the Arellano–Bond estimator. Note that we still assume stationarity and we do not offer a solution to cases with $\alpha = 1$.

The additional moment condition used in the System-GMM estimator is

$$E((\mu_i + \varepsilon_{it})(y_{i,t-1} - y_{i,t-2})) = 0.$$

This moment condition holds for $t = 4, \dots, T$. For $t = 3$, it holds if the initial observations (y_{i0}) are generated from the stationary distribution. Proposed by Arellano and Bover (1995) and Blundell and Bond (1998), these are called “level moment conditions.”

We now show that the level moment conditions hold. First, we observe that $E(\varepsilon_{it}(y_{i,t-1} - y_{i,t-2})) = 0$. Thus, we need to show

$$E(\mu_i(y_{i,t-1} - y_{i,t-2})) = 0.$$

This holds because $E(\mu_i y_{it})$ is constant for any $t \geq 4$. If the initial observations are generated from the stationary distribution, then it holds with $t = 3$.

The System-GMM estimator solves the sample analog of the level moment conditions:

$$E((y_{it} - \alpha y_{i,t-1})(y_{i,t-1} - y_{i,t-2})) = 0,$$

for $t = 3, \dots, T$, in addition to those used for the Arellano–Bond estimator ($E(\mathbf{Z}_i'(\Delta y_i - \alpha \Delta y_{i-1})) = 0$). Blundell and Bond (1998) demonstrate in their simulations that the System-GMM estimator exhibits a small bias.

While the System-GMM estimator is frequently used in empirical research, its good properties may not be guaranteed. Bun and Kiviet (2006) analytically derive an approximation to the bias, which indicates that the results of Blundell and Bond (1998) may be an artifact of their simulation designs. In general, they show that the System-GMM estimator suffers from bias. Hayakawa (2007) shows that the difference moment conditions and the level moment conditions yield biases with opposite directions. This finding implies that we may expect a small bias when we use both types of moment conditions. Alternatively, Bun and Windmeijer (2010) argue that the level moment conditions are weak when the variance of the fixed effects is large. Kruiniger (2009) demonstrates situations in which both types of moment conditions become weak.

The other criticism of the System-GMM estimator is that the System-GMM estimator heavily depends on the assumption of stationarity. If violated, the System-GMM estimator may not behave well. However, note that there are two kinds of possible violations of stationarity relevant in this context. The first is the unit root case ($\alpha = 1$). Bun and Kleibergen (2013) argue that in the Arellano–Bond and System-GMM estimations the identification of the parameter is irregular when α is close to one. The other type of nonstationarity concerns the initial condition. That is, the distribution of the initial observations y_{i0} is not a stationary distribution. Bun and Sarafidis (2015) provide an excellent survey about dynamic panel data models and include a good discussion of the initial condition problem. Hayakawa (2009b) and Hayakawa and Nagata (2016) show that the Arellano–Bond estimator works well when the initial conditions are not generated from the stationary distribution, even when α is close to one.

There is an additional type of moment conditions implied by the model. Ahn and Schmidt (1995) provide the complete list of moment conditions for the panel AR(1) model to achieve efficiency. This additional type of moment conditions is nonlinear and may not frequently apply in applied research. However, Bun and Sarafidis (2015) advocate its use and provide simulation evidence of its good properties.

3.4 Many Instruments Problem

The second problem with the Arellano–Bond estimator is that the number of moment conditions is large when T is large. It is also known that the GMM estimator suffers from bias when the number of instruments is large (Kunitomo, 1980; Morimune, 1983; Bekker, 1994; Donald and Newey, 2001; Newey and Windmeijer, 2009). However, its variance decreases with the number of instruments. The number of instruments (or moment conditions) can be large in the GMM estimation of dynamic panel data models. For example, y_{i1} is an instrument for $y_{i2} - y_{i1}$. But for $y_{i10} - y_{i9}$, there are nine instruments ($y_{i1}, \dots, y_{i9}$). In total, the number of moment conditions is of order $O(T^2)$. Even if T is only moderately large, the number of moment conditions is still large.

Alvarez and Arellano (2003) demonstrate that the Arellano–Bond estimator suffers from bias when both N and T tend to infinity. Bun and Kiviet (2006) and Okui (2009) provide a closer investigation of the problem of many moments bias in the moment-based estimation of dynamic panel data models. In sum, the bias depends on the number of instruments for each regressor, not the total number of moment conditions, such that the variance becomes small as the number of moment conditions increases. Okui (2009) also proposes a method to

choose the number of moment conditions. Relatedly, Carrasco and Nayihouba (2020) propose a regularization method to alleviate the many moments bias.

A simple solution is proposed by Hayakawa (2009a), which is an estimator that is consistent even when T is small. This does not suffer from the many instruments bias yet achieves asymptotic efficiency when T is large and is referred to as the double filter instrumental variable (DFIV) estimator.

The DFIV estimator employs “forward orthogonal deviation” to eliminate the individual effects. Let

$$y_{it}^* = \sqrt{\frac{T-t}{T-t+1}} \left(y_{it} - \frac{y_{i,t+1} + \dots + y_{iT}}{T-t} \right)$$

$$x_{it}^* = \sqrt{\frac{T-t}{T-t+1}} \left(y_{i,t-1} - \frac{y_{i,t} + \dots + y_{i,T-1}}{T-t} \right).$$

This transformation subtracts the average of future values. The term $\sqrt{(T-t)/(T-t+1)}$ equalizes the variances of the variables across t under homoscedasticity. When applied to a time-invariant variable, we obtain zero, thus eliminating the fixed effects.

The model for the transformed variables is

$$y_{it}^* = \alpha x_{it}^* + \varepsilon_{it}^*.$$

The DFIV estimator is an instrumental variable estimator for this transformed model, in which x_{it}^* is instrumented by the lagged dependent variable transformed by “backward orthogonal deviation”:

$$z_{it} = \sqrt{\frac{T-t}{T-t+1}} \left(y_{i,t-1} - \frac{y_{i,t-1} + \dots + y_{i1}}{t-1} \right).$$

Note that this estimator uses only one instrument for each regressor. The closed form of the DFIV estimator is

$$\hat{\alpha}_H = \left(\sum_{i=1}^N \sum_{t=2}^{T-1} z_{it} x_{it}^* \right)^{-1} \sum_{i=1}^N \sum_{t=2}^{T-1} z_{it} y_{it}^*.$$

This estimator is consistent as long as N and/or T tend to infinity. Moreover, when $T \rightarrow \infty$ and ε_{it} is i.i.d. normal, it achieves the efficiency bound derived by Hahn and Kuersteiner (2002) and Iwakura and Okui (2012).

The DFIV estimator is extended to more general dynamic panel data models by Hayakawa, et al. (2019). They consider general panel data models with predetermined regressors with fixed effects including panel AR models and show that the estimator is asymptotically equivalent to a biased corrected version of the fixed effects estimator. Relatedly, Roodman (2009) suggests using the sum of predetermined variables as an instrument to reduce the number of instruments.

3.5 With Regressors

We can also estimate dynamic panel data models with regressors following the same argument. Let X_{it} be a vector of regressors for i at t . Note that all elements of X_{it} need to be time varying. A time-invariant regressor is eliminated from the model in the process of eliminating the individual effects. We consider the following panel AR(1) model with regressors:

$$y_{it} = \mu_i + \alpha y_{i,t-1} + \beta' X_{it} + \varepsilon_{it}.$$

This model can be estimated by GMM with additional moment conditions regarding the relationship between X_{it} and ε_{it} . There are two kinds of exogeneity assumptions regarding X_{it} . The first is strict exogeneity. This means that $E(X_{it}\varepsilon_{is}) = 0$ for any t and s . In this case, we use moment conditions $E(X_{it}(\varepsilon_{is} - \varepsilon_{i,s-1})) = 0$ for any t and s . The second assumption is predeterminedness, where $E(X_{it}\varepsilon_{is}) = 0$ for $t \leq s$. In this case, we use $E(X_{it}(\varepsilon_{is} - \varepsilon_{i,s-1})) = 0$ for $t < s$. Which assumption is more appropriate depends on the application. Let z'_{it} be the row vector of instruments for the equation at time $t + 1$. When X_{it} is strictly exogenous $z_{it} = (y_{i1}, \dots, y_{it}, X'_{i1}, \dots, X'_{iT})$, and when X_{it} is predetermined then $z_{it} = (y_{i1}, \dots, y_{it}, X'_{i1}, \dots, X'_{i,t+1})$. The GMM estimator has the same formula as that of the Arellano–Bond estimator except that $\mathbf{Z}_i$ is replaced by

$$\mathbf{Z}_i = \begin{pmatrix} z_{i1} & 0 & \dots & 0 \\ 0 & z_{i2} & \dots & 0 \\ \dots & \dots & \dots & 0 \\ 0 & \dots & 0 & z_{i,T-2} \end{pmatrix}.$$

Note that we may also add the level moment conditions. The bias of the GMM estimator in models with regressors is examined by Bun and Kiviet (2006).

3.6 Inferences and Specification Tests

The Arellano–Bond estimator is a GMM estimator and asymptotically normal, so statistical inferences can use the usual standard errors for the GMM estimator. However, the empirical size of a test or the coverage probability of a confidence interval based on the usual standard errors may not be close to the nominal size or coverage (see, for example, simulation studies in Okui, 2009). Moreover, if the weighting matrix depends on an initial consistent estimate, the problem becomes more severe. To solve the problem, Windmeijer (2005) proposes an alternative standard error formula called the standard error with the Windmeijer correction. This generates better coverage and is frequently used. Hwang et al. (2021) extend this to provide valid standard errors even in the presence of misspecification.

Another important inference for dynamic panel data models is specification testing. That is, it is important to examine the validity of moment conditions. Because the Arellano–Bond estimator is a GMM estimator, the J -test for overidentifying restrictions by Hansen (1982) can be used. However, the presence of many moment restrictions may cause size distortion. To address this issue, specification tests tailored to dynamic panel data models have been duly proposed.

The key assumption for the estimation of dynamic panel data models is the serial uncorrelatedness of ε_{it} . Given first-order serial correlation is used for the estimation and is not testable, Arellano and Bond (1991) propose a test for the second-order serial correlation in ε_{it} . Yamagata (2008) extends this to construct a joint test for second- and higher-order serial correlations.

4. BIAS-CORRECTED FIXED EFFECTS ESTIMATOR

This section discusses the fixed effects estimator. Although it is the most representative estimation method in the panel data literature, it does not provide a consistent estimator in dynamic panel data models when T is fixed. When T tends to infinity, it is consistent but suffers from bias. Several bias-corrected estimators have been proposed to alleviate the bias when T is moderately large.

4.1 Fixed Effects Estimation

The fixed effects estimator is an ordinary least squares (OLS) estimator applied to panel data using a fixed effects transformation. The fixed effects transformation subtracts the mean of the variables and thus eliminates the fixed effects from the model. However, the fixed effects estimator is biased and is inconsistent when T is fixed.

The fixed effects estimator is computed as follows. We first eliminate the individual effects by subtracting the individual averages:

$$y_{it} - \bar{y}_i = \alpha(y_{i,t-1} - \bar{y}_{-1,i}) + \varepsilon_{it} - \bar{\varepsilon}_i,$$

where $\bar{y}_i = \sum_{t=2}^T y_{it} / T$, $\bar{y}_{-1,i} = \sum_{t=2}^T y_{i,t-1} / T = \sum_{t=1}^{T-1} y_{i,t} / T$ and $\bar{\varepsilon}_i = \sum_{t=2}^T \varepsilon_{it}$. We then apply OLS to this transformed model. The resulting estimator is the fixed effects estimator, denoted as $\hat{\alpha}_{FE}$, namely

$$\hat{\alpha}_{FE} = \left(\sum_{i=1}^N \sum_{t=2}^T (y_{i,t-1} - \bar{y}_{-1,i})^2 \right)^{-1} \left(\sum_{i=1}^N \sum_{t=2}^T (y_{i,t-1} - \bar{y}_{-1,i})(y_{it} - \bar{y}_i) \right).$$

However, the fixed effects estimator is not consistent when $N \rightarrow \infty$ but T is fixed. The source of inconsistency is the correlation between the regressor and the error term. Note that in the transformed model, the regressor is $y_{i,t-1} - \bar{y}_{-1,i}$ and the error term is $\varepsilon_{it} - \bar{\varepsilon}_i$. The problem is that $\bar{y}_{-1,i}$ includes the dependent variables from $t = 1$ to $t = T - 1$ and $\bar{\varepsilon}_i$ includes the error terms from $t = 2$ to T . By the nature of the model, y_{is} and ε_{it} are correlated if $s \geq t$. We thus have

$$E((y_{i,t-1} - \bar{y}_i)(\varepsilon_{it} - \bar{\varepsilon}_i)) \neq 0.$$

This observation implies that

$$\hat{\alpha}_{FE} = \alpha + \left(\sum_{i=1}^N \sum_{t=1}^T (y_{i,t-1} - \bar{y}_{-1,i})^2 \right)^{-1} \left(\sum_{i=1}^N \sum_{t=1}^T (y_{i,t-1} - \bar{y}_{-1,i})(\varepsilon_{it} - \bar{\varepsilon}_i) \right)$$

$$\rightarrow_p \alpha + \left(E \left(\sum_{i=1}^T (y_{i,t-1} - \bar{y}_{-1,i})^2 \right) \right)^{-1} \left(E \left(\sum_{i=1}^T (y_{i,t-1} - \bar{y}_{-1,i})(\varepsilon_{it} - \bar{\varepsilon}_i) \right) \right) \neq \alpha.$$

Thus, $\hat{\alpha}_{FE}$ is not consistent.

Of course, $\hat{\alpha}_{FE}$ is consistent when $T \rightarrow \infty$. The correlation between $y_{i,t-1} - \bar{y}_{-1,i}$ and $\varepsilon_{it} - \bar{\varepsilon}_i$ becomes negligible as T increases. In particular, when N is fixed and $T \rightarrow \infty$, it is a time series problem and the consistency of the OLS estimator has been established in the time series literature.

Nonetheless, the bias may not be negligible in finite samples, even when T is moderately large (Hahn and Kuersteiner, 2002; Alvarez and Arellano, 2003). Moreover, even when T is large, if the ratio between N and T is also large, the asymptotic distribution becomes biased and standard inference procedures may not be valid.

4.2 Bias Correction

We now discuss bias-correction methods for the fixed effects estimator. There have been several attempts to alleviate the bias of the fixed effects estimator. They still require $T \rightarrow \infty$ but the finite sample properties of bias-corrected fixed effects estimators are promising even with moderate T . Moreover, they tend to have a smaller variance than those of the GMM-type estimators.

We first discuss analytical bias correction. Hahn and Kuersteiner (2002) propose a simple bias-corrected estimator:

$$\hat{\alpha}_{HK} = \frac{T+1}{T} \hat{\alpha}_{FE} + \frac{1}{T} = \hat{\alpha}_{FE} + \frac{1}{T} (1 + \hat{\alpha}_{FE}).$$

This estimator is motivated by the following observation. When $N, T \rightarrow \infty$ and $N/T \rightarrow \kappa$, where $0 < \kappa < \infty$, Hahn and Kuersteiner (2002) show that

$$\sqrt{NT}(\hat{\alpha}_{FE} - \alpha) \rightarrow_d N(-\sqrt{\kappa}(1+\alpha), 1-\alpha^2),$$

under homoscedasticity (that is, $Var(\varepsilon_{it}) = \sigma_\varepsilon^2$, where σ_ε^2 does not depend on i or on t). Thus, the asymptotic distribution of the fixed effects estimator is not centered around zero even when $T \rightarrow \infty$. Alternatively, we may write that

$$\sqrt{NT} \left(\hat{\alpha}_{FE} - \alpha + \frac{1}{T}(1+\alpha) \right) \rightarrow_d N(0, 1-\alpha^2).$$

Thus, we have

$$\begin{aligned} \sqrt{NT}(\hat{\alpha}_{HK} - \alpha) &= \sqrt{NT} \left(\hat{\alpha}_{FE} - \alpha + \frac{1}{T}(1+\alpha) \right) + \sqrt{\frac{N}{T}}(1 + \hat{\alpha}_{FE} - (1+\alpha)) \\ &\rightarrow_d N(0, 1-\alpha^2). \end{aligned}$$

Note that, when T is fixed, $\hat{\alpha}_{HK}$ is still inconsistent.

Several studies discuss the bias properties of the fixed effects estimator and bias correction methods for models with regressors. Kiviet (1995) examines the bias problem in panel AR(1) models with regressors and proposes a bias-corrected estimator. The bias is also examined in Bun and Kiviet (2006). Bun and Carree (2005) propose a bias-corrected estimator that is consistent, even when T is fixed. However, the bias-correction methods proposed in these papers may appear more complex than that by Hahn and Kuersteiner (2002). Gouriéroux et al. (2010) propose a bias correction through indirect inference.

The discussion here focuses on panel AR(1) models, but similar arguments hold for the other dynamic panel data models. For example, Lee (2012) considers the fixed effects estimation of panel AR(p) models. Lee et al. (2018) examine the bias of the fixed effects estimator and propose a simple analytical bias correction in panel AR(∞) models. Arellano (2016) considers dynamic panel data models with predetermined regressors. Okui (2010) considers bias-corrected estimations of autocovariances and autocorrelations for panel data with fixed effects.

4.3 Jackknife Bias Correction

An alternative bias correction method draws on jackknife. Here, we consider the half-panel jackknife (HPJ) bias correction proposed by Dhaene and Jochmans (2015) and available for general dynamic panel data models, including nonlinear models.

The HPJ procedure is as follows. Let $\hat{\alpha}$ be an estimator. Let $\hat{\alpha}^{(1)}$ and $\hat{\alpha}^{(2)}$ be the estimators of α computed using $\{\{y_{it}\}_{i=1}^{T/2}\}_{t=1}^N$ and $\{\{y_{it}\}_{i=T/2+1}^T\}_{t=1}^N$, respectively, with even T . Let $\bar{\alpha} = (\hat{\alpha}^{(1)} + \hat{\alpha}^{(2)}) / 2$. The HPJ bias-corrected estimator of H is defined by

$$\hat{\alpha}^H = \hat{\alpha} - (\bar{\alpha} - \hat{\alpha}) = 2\hat{\alpha} - \bar{\alpha}.$$

The main idea for the HPJ bias correction is the following. Consider the fixed effects estimator, $\hat{\alpha}_{FE}$, which exhibits a bias of order T^{-1} . We thus have

$$\hat{\alpha}_{FE} = \alpha + \frac{B}{T} + \text{mean zero term and small terms},$$

where B is a constant. Each $\hat{\alpha}^{(i)}$ for $i = 1, 2$ is based on panel data whose length is $T/2$. Thus, we have

$$\bar{\alpha} = \alpha + \frac{2B}{T} + \text{mean zero term and small terms}.$$

Thus, the bias of order $1/T$ of $\hat{\alpha}_{FE}$ is estimated by $\bar{\alpha} - \hat{\alpha}_{FE}$. Therefore, the HPJ estimator $\hat{\alpha}^H = \hat{\alpha}_{FE} - (\bar{\alpha} - \hat{\alpha}_{FE})$ reduces the bias of order $1/T$ of $\hat{\alpha}_{FE}$.

4.4 Inferences

We suggest bootstrap methods to conduct statistical inferences based on (bias-corrected) fixed effects estimators. Although standard inference procedures, for example t -statistics with usual standard errors, are asymptotically valid, their finite sample properties may not be encouraging because bias correction may not be able to eliminate the bias completely in finite samples, and

this affects the variability of the estimator. Galvao and Kato (2014) propose a cross-sectional bootstrap method. This is a nonparametric bootstrap procedure treating each time series as the unit of observation. Let $y_i = (y_{i1}, \dots, y_{iT})$. The algorithm is as follows:

1. Randomly draw $y_1^*, \dots, y_N^*$ from $\{y_i\}_{i=1}^N$ with replacement.
2. Compute the statistics of interest, say $\hat{S}^*$, from $\{y_i^*\}_{i=1}^N$.
3. Repeat 1 and 2 B times. Let $\hat{S}^*(b)$ be the estimate computed in the b th bootstrap sample.
4. Compute the quantities of interest from the empirical distribution of $\{\hat{S}^*(b)\}_{b=1}^B$.

For example, a 95 per cent confidence interval for α is $[\hat{\alpha} - q_{0.975}^*, \hat{\alpha} - q_{0.025}^*]$ where $\hat{\alpha}$ is an estimate of α , and $q_{0.975}^*$ and $q_{0.025}^*$ are the 97.5 and 2.5 percentiles of the empirical distribution of $\{\hat{\alpha}^*(b) - \hat{\alpha}\}_{b=1}^B$. This bootstrap procedure can be used in any panel data model as long as the panel is cross-sectionally independent. Okui and Yanagi (2019) apply it for statistical inferences for parameters characterizing heterogeneity in dynamics. Gonçalves and Kaffo (2015) propose different bootstrap methods that exploit the features of dynamic panel data models.

5. MOTIVATIONS

In this section, we discuss motivations behind the use of dynamic panel data models. We give the definitions of unobserved heterogeneity and state dependence and discuss their importance in analyzing economic problems. They are important not only in the analysis of dynamics, but also in cases in which the research interest is on the effect of some regressor because of the need to be controlled to obtain an unbiased estimate of the coefficient of the regressor.

5.1 Unobserved Heterogeneity and State Dependence

We consider the importance of distinguishing between unobserved heterogeneity and state dependence. As an example, we consider income dynamics, although analysis of many other economic variables would face similar problems. The income process is typically serially correlated. For example, individuals receiving high incomes in previous years tend to earn a high income this year. This phenomenon can be explained by unobserved heterogeneity and/or state dependence.

Unobserved heterogeneity may explain serial correlation in income dynamics. Those who have high incomes are productive workers (heterogeneity in productivity or ability). Therefore, they tend to receive high incomes throughout their life. For this reason, we may consider it spurious serial correlation because it does not cause serial correlation in the individual dynamics. In other words, we observe serial correlation only in the cross-sectional dimension if caused purely by unobserved heterogeneity.

Conversely, state dependence means that serial correlation in income arises because of a temporal dependence in the income process. That is, those who earned high incomes by chance may subsequently receive high incomes. However, this time dependency in income does not reflect the difference in ability and may result from some market imperfection (such as wage rigidity).

We want to distinguish between unobserved heterogeneity and state dependence for several reasons. First, it helps us to build a better model for income dynamics, which in turn is used as a key part of, for example, large macro models. Second, unobserved heterogeneity and state dependence provide very different policy implications. If serial correlation is caused by unobserved heterogeneity, the observed persistent income inequality is an outcome of the difference in ability. As such, a redistribution policy cannot be justified from any concern with efficiency; instead, it should draw on some normative criteria (for example, we think income inequality is bad). Alternatively, if caused only by state dependence, then persistent income inequality simply reflects that some people are lucky and others are not. As such, a redistribution policy can be justified from an efficiency perspective and thus considered as a risk-sharing mechanism.

Of course, unobserved heterogeneity and state dependence are not mutually exclusive. It would be more reasonable to assume that both play roles in determining the serial correlation structure of income dynamics. Thus, it would be desirable to have statistical methods with which we can evaluate the contribution of each factor while controlling for the other.

The panel AR(1) model is useful to investigate to what extent each interpretation of the persistency in income is relevant. Recall that the panel AR(1) model is

$$y_{it} = \mu_i + \alpha y_{i,t-1} + \varepsilon_{it}.$$

In this model, μ_i (individual (fixed) effect) represents the unobserved heterogeneity, and α (AR(1) parameter) represents the degree of state dependence. This chapter focuses on the estimation of α . That is, we discuss how we can estimate the degree of state dependence while controlling for unobserved heterogeneity.

Distinguishing between state dependence and unobserved heterogeneity becomes important even when we are not interested in the dynamics itself, but rather in the effect of a regressor. Note that controlling for unobserved heterogeneity is standard in panel data analysis. Dynamic panel data models allow us to control both unobserved heterogeneity and state dependence simultaneously.

For example, Asiedu and Lien (2011) consider the effect of democracy on foreign direct investment (FDI) and investigate whether it depends on the availability of natural resources. Their model is

$$\begin{aligned} fdi_{it} = & \mu_i + \alpha fdi_{i,t-1} + \beta dem_{it} + \gamma dem_{it} \times nat_{it} \\ & + \text{other regressors} + \varepsilon_{it}, \end{aligned}$$

where fdi_{it} is the amount of FDI, dem_{it} is an index of the degree of democracy and nat_{it} indicates the availability of natural resources. The parameters of interest are β and γ . In this application, α may not be the parameter of primary interest but $fdi_{i,t-1}$ still needs to be added to estimate β and γ unbiasedly.

It would be natural to think that the degree of democracy correlates with unobserved time-invariant characteristics of countries, such as culture, and that these characteristics would affect the amount of FDI. Moreover, past FDI may affect the current level of democracy through international politics, and we would expect that FDI would possess state dependence. Thus, to estimate the effect of democracy on FDI, we should control for unobserved heterogeneity

across countries and the state dependence of FDI. Otherwise, the regression suffers from omitted variable bias. The dynamic panel data model provides a simple way to achieve this goal. However, the estimation of that model faces several statistical challenges as already discussed in this chapter.

Economic theory may also imply dynamic panel data models. An important example is the estimation of production functions. Suppose that the production is of a Cobb–Douglas type so that the log output y_{it} is determined by

$$y_{it} = a_{it} + \beta k_{it} + \gamma l_{it},$$

where a_{it} is the log productivity level of firm i at period t , k_{it} is log capital and l_{it} is log labor. Suppose that a_{it} follows $a_{it} = \mu_i + \alpha a_{i,t-1} + \zeta_{it}$, where μ_i is the time-invariant part of productivity and ζ_{it} is an idiosyncratic productivity shock. By subtracting αy_{it} from both sides and rearranging terms, we obtain

$$y_{it} = \mu_i + \alpha y_{i,t-1} + \beta k_{it} - \beta \alpha k_{i,t-1} + \gamma l_{it} - \gamma \alpha l_{i,t-1} + \xi_{it}.$$

This transformed model is a dynamic panel data model and the techniques discussed in this chapter may be applied. For details on the estimation of production functions from the perspective of dynamic panel data models, see Blundell and Bond (2000). Akerberg et al. (2015) discuss its relation to other approaches for production function estimation.

6. EMPIRICAL ILLUSTRATION

In this section, we illustrate procedures for the estimation and inference of dynamic panel models through an empirical example. We revisit the study by Acemoglu et al. (2008) on the effect of economic situation on democracy. The statistical computation is conducted with STATA MP15.1 on macOS Big Sur 11.1.

We use the 5 year panel data originally used in Acemoglu et al. (2008).¹ The data consist of observations from 90 countries from 1970 to 2000 (the length of time series is 7) and is balanced.² We investigate the effect of economic situation measured by GDP per capita on a measure of democracy. The Freedom House democracy index is used as the measure of democracy. See Acemoglu et al. (2008) for the details of the data.

We subtract cross-sectional averages from variables to eliminate time effects. Let dem_{it} be the democracy index for country i at year t and gdp_{it} the GDP per capita for county i at year t . Define $\log(\widehat{dem}_{it}) = \log(dem_{it}) - \sum_{j=1}^N \log(dem_{jt}) / N$ and $\log(\widehat{gdp}_{it}) = \log(gdp_{it}) - \sum_{j=1}^N \log(gdp_{jt}) / N$.

We estimate the following dynamic panel data models.

$$\log(\widehat{dem}_{it}) = \beta \log(\widehat{dem}_{i,t-1}) + \theta \log(\widehat{gdp}_{i,t-1}) + \alpha_i + \varepsilon_{it}, \quad (1.2)$$

where α_i is the country effect. Note that time effects are eliminated by cross-sectional demeaning and are not included in the model. We assume that $\log(\widehat{gdp}_{i,t-1})$ is predetermined. The main

parameter of interest is θ . It measures the effect of economic growth on democracy, controlling for the state dependency of democracy and unobserved heterogeneity among countries.

Various packages are available for the estimation of dynamic panel data models with STATA; `xtdpdgm` by Kripfganz (2017) implements all GMM estimators discussed in this chapter; `xtspj` by Sun and Dhaene (2019) implements the half-panel jackknife bias correction for the fixed effects estimator. The estimation results discussed are obtained with these packages. Unfortunately, at the time of writing, it is still burdensome to implement cross-sectional bootstrap for panel data methods using STATA.

Table 1.1 presents the estimation results from various estimation procedures. The estimated values of the coefficients vary across estimation methods. This result indicates the importance of understanding different properties of various estimation methods. The OLS estimate of θ is positive and statistically significant. It suggests an economically important and statistically significant effect of income on democracy. However, this result is overturned once we take into account unobserved heterogeneity. The Anderson–Hsiao estimate is negative and large, but its standard error is large too. The Arellano–Bond estimator also yields negative estimates which are statistically insignificant. The system-GMM estimates of θ are positive. However, the estimate from the iterative GMM is not statistically significant and its magnitude is not large either. The fixed effects estimator of θ is negative and the bias correction inflates the size of the effect. On the other hand, the panel data used in this empirical exercise are short and these fixed effects methods may not be particularly reliable compared with the GMM procedures in the current case.

Our results are consistent with those of Acemoglu et al. (2008). Once we take into account unobserved heterogeneity and state dependence, we tend to find either negative or small and statistically insignificant effect of income on democracy. The system-GMM estimator yields statistically significant positive results, but the results are not robust and the iterative version is small and statistically insignificant. It is thus sensible to conclude that there is no solid statistical evidence that good economic situation promotes democracy.

7. CONCLUSION

This chapter provides a review of the estimation of linear dynamic panel data models. We consider a panel AR(1) model and its extensions. We consider two estimation methods: GMM and the bias-corrected fixed effects estimator. Lastly, we consider the economic motivation to analyze dynamic panel data models.

We suggest the following as best practice depending on the length of time series of the data. When the panel is short (say, $T < 7$), then the Arellano–Bond approach may be used. We would iterate the estimation of the parameter and computation of the weighting matrix as suggested by Hansen and Lee (2020). Statistical inferences are conducted with standard errors based on the Windmeijer (2005) correction. In contrast, when the panel is moderately long (say, $T \geq 7$), bias-corrected fixed estimators such as those suggested by Hahn and Kuersteiner (2002) and Dhaene and Jochmans (2015) are suggested. Statistical inferences may be conducted with cross-sectional bootstraps as in Galvao and Kato (2014). Regardless of the length of time series, the DFIV estimator by Hayakawa (2009a) does not suffer from bias, and inferences based on it have correct coverage and size but may have a large variance in finite samples. Note that the previously-mentioned cutoff at $T = 7$ is suggested

Table 1.1 Estimation results of model (2)

Estimator	weights	OLS	FD	AH	AB one	AB two	AB ite	Sys one	Sys two	Sys ite	FE	HPJ
$\hat{\beta}$		0.675	-0.248	0.348	0.376	0.362	0.346	0.534	0.605	0.912	0.29	0.541
		(0.050)	(0.062)	(0.126)	(0.107)	(0.107)	(0.110)	(0.066)	(0.074)	(0.063)	(0.064)	(0.064)
$\hat{\theta}$		0.070	-0.056	-0.362	-0.228	-0.158	-0.13	0.072	0.080	0.031	-0.064	-0.114
		(0.015)	(0.056)	(0.162)	(0.123)	(0.098)	(0.101)	(0.022)	(0.026)	(0.032)	(0.052)	(0.052)
Country effects		No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Standard errors		CR	CR	CR	CR	WC	WC	CR	WC	WC	CR	CR

Note: FD: first difference estimator. AH: Anderson-Hsiao estimator. AB: Arellano-Bond estimator. Sys: system-GMM estimator. FE: fixed effects estimator. HPJ: fixed effects estimator with half-panel jackknife bias correction. one: weighting matrix optimal under homoskedasticity. two: heteroskedasticity optimal weighting matrix based on the one-step procedure. ite: iterative GMM. CR: cluster robust standard errors. WC: cluster robust standard errors with Windmeijer's correction.

by the simulation results in, for example, Hahn and Kuersteiner (2002) and Okui (2009), but is not definite.

This chapter considers panel AR(1) models with fixed effects, and because of space constraints we do not discuss more complex models also used in the literature. These include models with time effects, individual specific time trends, and interactive effects. Hahn and Moon (2006) show that the bias-correction method by Hahn and Kuersteiner (2002) can be applied even when the model includes time effects. Okui (2014) argues that time effects create bias for the estimation of autocovariances, but because they affect the estimator in a multiplicative way, it is not present for autocorrelation estimators because we take a ratio between autocovariances. Phillips and Sul (2007) examine the role of individual specific time trends in the bias of the fixed effects estimator, and Okui (2011) does so for the estimation of autocovariances and autocorrelations. Models with interactive effects (models with factor structure) are examined by Phillips and Sul (2007) and Moon and Weidner (2015, 2017). The GMM estimation of models with interactive effects is considered by Robertson and Sarafidis (2015). Pesaran (2015) provides an excellent summary of these advanced topics, particularly for panel data with long time series.

Lastly, we suggest some possible future research topics. In particular, we argue that heterogeneity and misspecification are two important topics in the literature on dynamic panel data models. Pesaran and Smith (1995), Hsiao et al. (1999), Phillips and Moon (1999) and Pesaran (2006) consider panel AR(1) models with heterogeneous coefficients, that is, models in which α depends on i , and consider the estimation of the cross-sectional mean of the heterogeneous coefficients. How to estimate the entire distribution of heterogeneous coefficients is examined by Okui and Yanagi (2019, 2020). See also Jochmans and Weidner (2018), Arellano and Bonhomme (2012) and Barras et al. (2018). The other important issue is model misspecification. Okui (2008) shows that popularly used panel AR(1) estimators, such as the Arellano–Bond and fixed effects estimators, converge in probability to first-order autocorrelation when $T \rightarrow \infty$ followed by $N \rightarrow \infty$ even if the panel AR(1) model is misspecified. Okui (2017) demonstrates their probability limits when the model is misspecified and the panel exhibits heterogeneous dynamics. Torgovitsky (2019) uses a potential outcome framework to examine the identification of state dependence and unobserved heterogeneity when y_{it} is binary. It is important to examine what we can learn from the techniques for dynamic panel data models when they exhibit model misspecification and heterogeneity is not totally accounted for in the model. We fully expect research to concentrate on these issues in the future.

NOTES

- * The author gratefully acknowledges the financial support of the School of Social Sciences and a New Faculty Startup Grant at Seoul National University and the Housing and Commercial Bank Economic Research Fund in the Institute of Economic Research at Seoul National University. Junho Choi provided excellent research assistance. The usual disclaimer applies.
- 1. The authors made this dataset available to the research community via openICPSR, a repository of the Inter-university Consortium for Political and Social Research (ICPSR). The URL address is <https://www.openicpsr.org/openicpsr/project/113251/version/V1/view>.
- 2. The dataset is taken from “5 Year Panel” sheet in “Income and Democracy AER adjustment-1” available as a supplemental material of Acemoglu et al. (2008). We use a subset which satisfies *samplebalancegmm* = 1 so that the subset constitutes a balanced panel.

REFERENCES

- Acemoglu, Daron, Simon Johnson, James A. Robinson, and Pierre Yared (2008). "Income and democracy". In: *American Economic Review* 98.3, pp. 808–42.
- Ackerberg, Daniel, Kevin Caves, and Garth Frazer (2015). "Identification properties of recent production function estimators". In: *Econometrica* 83.6, pp. 2411–51.
- Ahn, Seung C. and Peter Schmidt (1995). "Efficient estimation of models for dynamic panel data". In: *Journal of Econometrics* 68, pp. 5–27.
- Alvarez, Javier and Manuel Arellano (2003). "The time series and cross-section asymptotics of dynamic panel data estimators". In: *Econometrica* 71.4, pp. 1121–59.
- Anderson, Theodore Wilbur and Cheng Hsiao (1981). "Estimation of dynamics models with error components". In: *Journal of the American Statistical Association* 76.375, pp. 598–606.
- Anderson, Theodore Wilbur and Cheng Hsiao (1982). "Formulation and estimation of dynamic models using panel data". In: *Journal of econometrics* 18.1, pp. 47–82.
- Arellano, Manuel (2003). *Panel Data Econometrics*. Oxford University Press.
- Arellano, Manuel (2016). "Modelling optimal instrumental variables for dynamic panel data models". In: *Research in Economics* 70, pp. 238–61.
- Arellano, Manuel and Stephen Bond (1991). "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations". In: *Review of Economics Studies* 58, pp. 277–97.
- Arellano, Manuel and Stéphane Bonhomme (2012). "Identifying distributional characteristics in random coefficients panel data models". In: *Review of Economic Studies* 79, pp. 987–1020.
- Arellano, Manuel and Olympia Bover (1995). "Another look at the instrumental variable estimation of error-component models". In: *Journal of Econometrics* 68, pp. 29–51.
- Asiedu, Elizabeth and Donald Lien (2011). "Democracy, foreign direct investment and natural resources". In: *Journal of International Economics* 84, pp. 99–111.
- Barras, Laurent, Patrick Gagliardini, and Olivier Scaillet (2018). "The cross-sectional distribution of fund skill measures", mimeo.
- Bekker, Paul A (1994). "Alternative approximations to the distributions of instrumental variable estimators". In: *Econometrica* 62.3, pp. 657–81.
- Blundell, Richard and Stephen Bond (1998). "Initial conditions and moment restrictions in dynamic panel data models". In: *Journal of Econometrics* 87, pp. 115–43.
- Blundell, Richard and Stephen Bond (2000). "GMM estimation with persistent panel data: an application to production functions". In: *Econometric reviews* 19.3, pp. 321–40.
- Bun, Maurice J. G. and Martin A. Carree (2005). "Bias-corrected estimation in dynamic panel data models". In: *Journal of Business & Economic Statistics* 23.2, pp. 200–210.
- Bun, Maurice J. G. and Jan F. Kiviet (2006). "The effects of dynamic feedbacks on LS and MM Estimator accuracy in panel data models". In: *Journal of Econometrics* 132, pp. 409–44.
- Bun, Maurice J. G. and Frank Kleibergen (2013). "Identification and inference in moments based analysis of linear dynamic panel data models". mimeo.
- Bun, Maurice J. G. and Vasilis Sarafidis (2015). "Dynamic Panel Data Models". In: *The Oxford Handbook of Panel Data*, ed. by Badi H. Baltagi. Oxford University Press, pp. 76–110.
- Bun, Maurice J. G. and Frank Windmeijer (2010). "The weak instrument problem of the system GMM estimator in dynamic panel data models". In: *The Econometrics Journal* 13.1, pp. 95–126.
- Carrasco, Marine and Ada Nayihouba (2020). "A regularization approach to the dynamic panel data model estimation", Mimeo.
- Clerides, Sofronis K., Saul Lach, and James R. Tybout (1998). "Is learning by exporting important? micro-dynamic evidence from Colombia, Mexico, and Morocco". In: *Quarterly Journal of Economics* 113.3, pp. 903–47.
- Dhaene, Geert and Koen Jochmans (2015). "Split-panel jackknife estimation of fixed-effect models". In: *The Review of Economic Studies* 82.3, pp. 991–1030.
- Donald, Stephen G. and Whitney K. Newey (2001). "Choosing the number of instruments". In: *Econometrica* 69.5, pp. 1161–91.
- Galvao, Antonio F. and Kengo Kato (2014). "Estimation and inference for linear panel data models under misspecification when both n and T are large". In: *Journal of Business & Economic Statistics* 32.2, pp. 285–309.
- Gonçalves, Sílvia and Maximilien Kaffo (2015). "Bootstrap inference for linear dynamic panel data models with individual fixed effects". In: *Journal of Econometrics* 186.2, pp. 407–26.
- Gouriéroux, Christian, Peter C. B. Phillips, and Jun Yu (2010). "Indirect inference for dynamic panel models". In: *Journal of Econometrics* 157.1, pp. 68–77.
- Hahn, Jinyong and Guido Kuersteiner (2002). "Asymptotically unbiased inference for a dynamic panel model with fixed effects when both n and T are large". In: *Econometrica* 70.4, pp. 1639–57.
- Hahn, Jinyong and Hyungsik Roger Moon (2006). "Reducing Bias of MLE in a dynamic panel model". In: *Econometric Theory* 22, pp. 499–512.

- Hansen, Bruce E. and Seojeong Lee (2021). "Inference for iterated GMM under misspecification". In: *Econometrica* 89.4, pp. 1419–47.
- Hansen, Lars Peter (1982). "Large sample properties of generalized method of moments estimators". In: *Econometrica* 50.4, pp. 1029–53.
- Hayakawa, Kazuhiko (2007). "Small sample properties of the system GMM estimator in dynamic panel data models". In: *Economics Letters* 95, pp. 32–8.
- Hayakawa, Kazuhiko (2009a). "A simple efficient instrumental variable estimator for panel AR(p) models when both N and T are large". In: *Econometric Theory* 25, pp. 873–90.
- Hayakawa, Kazuhiko (2009b). "On the effect of mean-nonstationarity in dynamic panel data models". In: *Journal of Econometrics* 153, pp. 133–5.
- Hayakawa, Kazuhiko and Shuichi Nagata (2016). "On the behavior of the GMM estimator in persistent dynamic panel data models with unrestricted initial conditions". In: *Computational Statistics and Data Analysis* 100, pp. 265–303.
- Hayakawa, Kazuhiko, Meng Qi, and Jörg Breitung (2019). "Double filter instrumental variable estimation of panel data models with weakly exogenous variables". In: *Econometric Reviews* 38.9, pp. 1055–88.
- Holtz-Eakin, Douglas, Whitney Newey, and Harvey S. Rosen (1988). "Estimating vector autoregressions with panel data". In: *Econometrica* 56.6, pp. 1371–95.
- Hsiao, Cheng, M. Hashem Pesaran, and A. K. Tahmiscioglu (1999). "Bayes estimation of short-run coefficients in dynamic panel data models". In: *Analysis of Panels and Limited Dependent Variables Models*, ed. by K. Lahiri, C. Hsiao, L.F. Lee, and M.H. Pesaran. Cambridge University Press, pp. 268–96.
- Hwang, Jungbin, Byunghoon Kang, and Seojeong Lee (2021). "A doubly corrected robust variance estimator for linear GMM". Forthcoming in the *Journal of Econometrics*.
- Iwakura, Haruo and Ryo Okui (2012). "Asymptotic efficiency in dynamic panel data models with factor structure". mimeo.
- Jochmans, Koen and Martin Weidner (2018). "Inference on a distribution from noisy draws". In: *arXiv preprint arXiv:1803.04991*.
- Kiviet, Jan F. (1995). "On bias, inconsistency, and efficiency of various estimators in dynamic panel data models". In: *Journal of Econometrics* 68, pp. 53–78.
- Kripfganz, Sebastian (Aug. 2017). *XTDPDGMM: Stata module to perform generalized method of moments estimation of linear dynamic panel data models*. Statistical Software Components, Boston College Department of Economics. <https://ideas.repec.org/c/boc/bocode/s458395.html>.
- Kruiniger, Hugo (2009). "GMM estimation and inference in dynamic panel data models with persistent data". In: *Econometric Theory*, pp. 1348–91.
- Kunitomo, Naoto (1980). "Asymptotic expansions of the distributions of estimators in a linear functional relationship and simultaneous equations". In: *Journal of the American Statistical Association* 75, pp. 693–700.
- Lee, Yoon-Jin, Ryo Okui, and Mototsugu Shintani (2018). "Asymptotic inference for dynamic panel estimators of infinite order autoregressive processes". In: *Journal of Econometrics* 204.2, pp. 147–58.
- Lee, Yoonseok (2012). "Bias in dynamic panel models under time series misspecification". In: *Journal of Econometrics* 169, pp. 54–60.
- Levine, Ross, Norman Loayza, and Thorsten Beck (2000). "Financial intermediation and growth: Causality and causes". In: *Journal of Monetary Economics* 46, pp. 31–77.
- Meghir, Costas and Luigi Pistaferri (2004). "Income variance dynamics and heterogeneity". In: *Econometrica* 72.1, pp. 1–32.
- Moon, Hyungsik Roger and Martin Weidner (2015). "Linear regression for panel with unknown number of factors as interactive fixed effects". In: *Econometrica* 83.4, pp. 1543–79.
- Moon, Hyungsik Roger and Martin Weidner (2017). "Dynamic linear panel regression models with interactive fixed effects". In: *Econometric Theory* 33, pp. 158–95.
- Morimune, Kimio (1983). "Approximate distribution of the k-class estimators when the degree of overidentifiability is large compared with the sample size". In: *Econometrica* 51.3, pp. 821–41.
- Newey, Whitney K. and Frank Windmeijer (2009). "Generalized method of moments with many weak moment conditions". In: *Econometrica* 77.3, pp. 687–719.
- Nickell, Stephen (1981). "Biases in dynamic models with fixed effects". In: *Econometrica* 49.6, pp. 1417–26.
- Nickell, Stephen J. (1996). "Competition and corporate performance". In: *Journal of Political Economy* 104.4, pp. 724–46.
- Okui, Ryo (2008). "Panel AR(1) estimators under misspecification". In: *Economics Letters* 101, pp. 210–13.
- Okui, Ryo (2009). "The optimal choice of moments in dynamic panel data models". In: *Journal of Econometrics* 151, pp. 1–16.
- Okui, Ryo (2010). "Asymptotically unbiased estimation of autocovariances and autocorrelations with long panel data". In: *Econometric Theory* 26, pp. 1263–304.
- Okui, Ryo (2011). "Asymptotically unbiased estimation of autocovariances and autocorrelations for panel data with incidental trends". In: *Economics Letters* 112, pp. 49–52.

- Okui, Ryo (2014). "Asymptotically unbiased estimation of autocovariances and autocorrelations with panel data in the presence of individual and time effects". In: *Journal of Time Series Econometrics* 6.2, pp. 129–81.
- Okui, Ryo (2017). "Misspecification in dynamic panel data models and model-free inferences". In: *Japanese Economic Review* 68.3, pp. 283–304.
- Okui, Ryo and Takahide Yanagi (2019). "Panel data analysis with heterogeneous dynamics". In: *Journal of Econometrics* 212.2, pp. 451–75.
- Okui, Ryo and Takahide Yanagi (2020). "Kernel estimation for panel data with heterogeneous dynamics". In: *The Econometrics Journal* 23.1, pp. 156–75.
- Pesaran, M. Hashem (2006). "Estimation and inference in large heterogeneous panels with a multifactor error structure". In: *Econometrica* 74.4, pp. 967–1002.
- Pesaran, M. Hashem (2015). *Time Series and Panel Data Econometrics*. Oxford University Press.
- Pesaran, M. Hashem and Ron Smith (1995). "Estimating long-run relationships from dynamic heterogeneous panels". In: *Journal of Econometrics* 68.1, pp. 79–113.
- Phillips, Peter C. B. and Hyungsik Roger Moon (1999). "Linear regression limit theory for nonstationary panel data". In: *Econometrica* 67.5, pp. 1057–111.
- Phillips, Peter C. B. and Donggyu Sul (2007). "Bias in dynamic panel estimation with fixed effects, incidental trends and cross section dependence". In: *Journal of Econometrics* 137, pp. 162–88.
- Robertson, Donald and Vasilis Sarafidis (2015). "IV estimation of panels with factor residuals". In: *Journal of Econometrics* 185.2, pp. 526–41.
- Roodman, David (2009). "A note on the theme of too many instruments". In: *Oxford Bulletin of Economics and Statistics* 71.1, pp. 135–58.
- Sarafidis, Vasilis and Tom Wansbeek (2021). "Celebrating 40 years of panel data analysis: Past, present and future". In: *Journal of Econometrics* 220, pp. 215–226.
- Staiger, D. and James H. Stock (1997). "Instrumental variables regression with weak instruments". In: *Econometrica* 65.3, pp. 557–86.
- Stock, James H. and Jonathan H. Wright (2000). "GMM with weak identification". In: *Econometrica* 68, pp. 1097–126.
- Sun, Yutao and Geert Dhaene (May 2019). *XTSPJ: Stata module for split-panel jack-knife estimation*. Statistical Software Components, Boston College Department of Economics. <https://ideas.repec.org/c/boc/bocode/s458651.html>.
- Torgovitsky, Alexander (2019). "Nonparametric inference on state dependence in unemployment". In: *Econometrica* 87.5, pp. 1475–505.
- Windmeijer, Frank (2005). "A finite sample correction for the variance of linear efficient two-step GMM estimators". In: *Journal of Econometrics* 126, pp. 25–51.
- Wooldridge, Jeffrey M. (2010). *Econometric Analysis of Cross Section and Panel Data*. Second edition. The MIT Press.
- Yamagata, Takshi (2008). "A joint serial correlation test for linear panel data models". In: *Journal of Econometrics* 146, pp. 135–45.