Journal of Econometrics 180 (2014) 174–197

Contents lists available at ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

Efficient GMM estimation of spatial dynamic panel data models with fixed effects☆

Lung-fei Lee [a], Jihai Yu [b,*]

[a] Department of Economics, Ohio State University, United States
[b] Guanghua School of Management, Peking University, China

ARTICLE INFO

Article history:
Received 31 August 2011
Received in revised form
1 July 2013
Accepted 8 March 2014
Available online 29 March 2014

JEL classification:
C13
C23
R15

Keywords:
Spatial autoregression
Dynamic panels
Fixed effects
Generalized method of moment
Many moments

ABSTRACT

In this paper we derive the asymptotic properties of GMM estimators for the spatial dynamic panel data model with fixed effects when n is large, and T can be large, but small relative to n. The GMM estimation methods are designed with the fixed individual and time effects eliminated from the model, and are computationally tractable even under circumstances where the ML approach would be either infeasible or computationally complicated. The ML approach would be infeasible if the spatial weights matrix is not row-normalized while the time effects are eliminated, and would be computationally intractable if there are multiple spatial weights matrices in the model; also, consistency of the MLE would require T to be large and not small relative to n if the fixed effects are jointly estimated with other parameters of interest. The GMM approach can overcome all these difficulties. We use exogenous and predetermined variables as instruments for linear moments, along with several levels of their neighboring variables and additional quadratic moments. We stack up the data and construct the best linear and quadratic moment conditions. An alternative approach is to use separate moment conditions for each period, which gives rise to many moments estimation. We show that these GMM estimators are $\sqrt{nT}$ consistent, asymptotically normal, and can be relatively efficient. We compare these approaches on their finite sample performance by Monte Carlo.

1. Introduction

Recently, there is a growing literature on spatial panel and dynamic panel models. By including spatial effects into static or dynamic panel models, we can take into account the cross section dependence from contemporaneous or lagged cross section interactions. Kapoor et al. (2007) extend the method of moments estimation to a spatial panel model with error components. Baltagi et al. (2007) consider the testing of spatial and serial dependence in an extended error components model, where serial correlation on each spatial unit over time and spatial dependence across spatial units are in the disturbances. Su and Yang (2007) study a dynamic panel data model with spatial error and random effects. These panel models specify spatial correlations by including spatially correlated disturbances and have emphasized on error components. In the fixed effects setting, Korniotis (2008) estimates a time-space recursive model, where individual time lag and spatial time lag are present, by the least square dummy variable (LSDV) regression approach. Yu et al. (2008, 2012) and Yu and Lee (2010) study the quasi maximum likelihood (QML) estimation for, respectively, the stable, spatial cointegration, and unit root spatial dynamic panel data (SDPD) models, where individual time lag, spatial time lag and contemporaneous spatial lag are all included.

For the stable SDPD model with fixed effects, the asymptotics of the QML estimation in Yu et al. (2008) is developed under $T \to \infty$ where T cannot be too small relative to n. In empirical applications, we might have data sets where n is large while T is relatively

☆ We would like to thank participants of seminars at Fudan University, Shanghai Jiao Tong University, University of Cincinnati, Singapore Management University, 10th World Congress of the Econometric Society (Shanghai, 2010), Shanghai Econometrics Workshop at Shanghai University of Finance and Economics (2011), two anonymous referees and the co-editor, Professor Cheng Hsiao, of this journal for helpful comments. Yu acknowledges funding from the National Science Foundation of China (project no. 71171005, 71322105) and support from the Center for Statistical Science of Peking University.

* Corresponding author. Tel.: +86 10 62760702.
E-mail addresses: lflee@econ.ohio-state.edu (L.-f. Lee), jihai.yu@gmail.com (J. Yu).

http://dx.doi.org/10.1016/j.jeconom.2014.03.003

small. Under this circumstance, in the literature of dynamic panels without spatial interactions, the maximum likelihood estimator (MLE) of the autoregressive coefficient of a linear dynamic panel model, which is also known as the within estimator, is biased and inconsistent when n tends to infinity but T remains finite (Nickell, 1981; Hsiao et al., 2002). This bias is due to the incidental parameter problem in Neyman and Scott (1948). By taking time differences to eliminate individual fixed effects in the dynamic equation, the estimation method of instrumental variables (IV) is popular (see Anderson and Hsiao, 1981; Arellano and Bond, 1991; Arellano and Bover, 1995; Blundell and Bond, 1998; Bun and Kiviet, 2006, etc.). This motivates our study of generalized method of moments (GMM) estimation of the SDPD model in order to cover the scenario that both n and T can be large, but T is small relative to n. The case of a finite T will also be considered.[1] In this paper, we investigate the GMM estimation of an SDPD model with possibly high order spatial lags. The inclusion of high order spatial lags can allow spatial dependence from different interactions characteristics such as geographical contiguity and economic interaction.[2] Compared to QML estimation, the GMM estimation has the following merits for the SDPD model: (1) GMM has a computational advantage over MLE, because GMM does not need to compute the determinant of the Jacobian matrix in the likelihood function for a spatial model, which is especially inconvenient for MLE when n is large or the model has high order spatial lags[3,4]; (2) some GMM methods can be applied to a short SDPD model and is free of asymptotic bias, while ML estimation of the SDPD model requires a large T and a bias correction procedure is needed to eliminate the asymptotic bias. For a finite T case, an initial specification for the first time period observations would also be needed in order to formulate a likelihood function.[5] (3) With carefully designed moment conditions, the GMM estimate can be more efficient than the QML estimate when the true distribution of the disturbances are not normal and has a nonzero degree of excess kurtosis; (4) GMM is also applicable for the SDPD model with time effects and non-row-normalized spatial weights matrices.[6,7]

Compared to dynamic panel data models where serial correlation occursin the time dimension, the SDPD model may have correlation in the time dimension as well as spatial correlation across units. In one approach, we stack up the data and use moment conditions where the IVs have a fixed column dimension for all the periods. In another, we can use separate moment conditions for each time period, which result in many moments. Those many moments not only come from time lags, but are also designed for spatial lags. We focus on the design of estimation methods that can have some asymptotic efficient properties. Normalized asymptotic distributions of IV estimators with a finite number of moments are properly centered at the true parameter vector. In the many moment approach, normalized asymptotic distributions of IV estimates might not be properly centered or an IV estimator might not be consistent due to the many IV moments (but not directly due to the fixed effects). In contrast to the asymptotics in Yu et al. (2008) where there are ratio conditions on how T and n go to infinity in order that ML estimates can be consistent or their normalized asymptotic distributions are properly centered, such ratio conditions may no longer be needed in the proposed GMM estimation with a finite number of moments in the present paper. In the many IVs estimation method, the ratio condition concerns about the number of IVs or moments relative to the total sample size nT, but not directly the ratio of T and n. However, if the total number of IVs is essentially a function of T, then n and T ratio conditions would appear; but in that case, the ratio condition requires that T shall be small relative to n. Thus, the many IVs approach is complementary to the QML approach. In other words, the proposed estimation methods can be applied to some scenarios where the T is small relative to n, while the QML method might not be so, in theory.[8]

The paper is organized as follows. Section 2 introduces the model and discusses moment conditions. Section 3 investigates the consistency and asymptotic distribution of various GMM estimators, and we discuss the asymptotic efficiency of the proposed estimators. Monte Carlo results for various estimators are provided in Section 4. Section 5 concludes the paper and summarizes the contributions. Some lemmas and proofs are collected in the Appendices.[9]

1 The reason for focusing on the asymptotic with $T \to \infty$ instead of a finite T is that, in this framework, we have the best IV or best GMM estimation with proper designs of IVs and moment conditions. This might not be possible for a fixed effects model when T is assumed to be finite.

2 In addition, a high order spatial lag model can be regarded as a general case of the first order spatial lag model with spatial disturbances. To see this, for a cross sectional SAR model $Y_n = \lambda_0 W_n Y_n + X_n \beta_0 + U_n$ where $U_n = \rho_0 M_n U_n + V_n$, with premultiplication of $(I_n - \rho_0 M_n)$, we have $Y_n = (\lambda_0 W_n + \rho_0 M_n - \lambda_0 \rho_0 M_n W_n) Y_{nt} + (I_n - \rho_0 M_n) X_n \beta_0 + V_n$ after re-arrangement. This is a high order spatial lags model with spatial weights matrices W_n, M_n, and $M_n W_n$ and constrained coefficients.

3 For a first order SAR model where the spatial weights matrix is diagonalizable, the determinant of the Jacobian term can be computed by its eigenvalues (see Ord, 1975). If the spatial weights matrix is not diagonalizable or we have some higher order spatial lags, the Ord device might not be applicable.

4 We note that, to construct the best instruments for the GMM in Section 3.1.2, we need to inverse the $n \times n$ matrix $S_n(\lambda) = I_n - \sum_{j=1}^{p} \lambda_j W_{nj}$ in (4) (the matrix inversion is also involved in obtaining the information matrix of ML estimation). This will cause a computation burden if n is large. However, unlike the computation of the determinant in ML estimation that is repeated in the parameter search, the matrix inverse computation needs to be obtained only once given a consistent estimate of parameter vector so that the computation burden is less severe (we can use power series expansion to compute the matrix inverse if necessary).

5 Elhorst (2010) has developed an ML estimation using the initial value approximation in Bhargava and Sargan (1983), which does not have much bias from their Monte Carlo results. Due to the multiple dimension search in the nonlinear variance matrix function, the ML estimation in Elhorst (2010) is computationally complicated; also, it has a larger bias than the GMME. In Yu et al. (2008), the consistency of the ML estimator is derived under large n and large T. The MLE can have satisfactory finite sample results after the bias correction from the Monte Carlo simulation. Both Yu et al. (2008) and Elhorst (2010) work well under a first order SDPD model.

6 It is possible to eliminate the time effects by taking cross sectional difference, but the resulting equation would not have an SAR representation and, therefore, one cannot set up a likelihood function for estimation. The MLE will have an additional incidental parameter problem if time effects need to be estimated in addition to the individual effects.

7 Bell and Bockstael (2000) argue that, based on some underlying economic story, it is not necessary to always row-normalize the spatial weights matrix. In some cases, row-standardizing changes the total impact of neighbors across observations, although it does not change the relative dependence among all neighbors of any given observation. They use the real estate problem to argue that row-standardizing will attach too much weight to the neighbors of remote houses. In social interaction and network literatures, when the social interaction is specified as an SAR model, the measure of centrality in Bonacich (1987) comes out naturally in the reduced form equation. When the indegrees (the sums of each row) of the sociomatrix have a non-zero variation, so does the Bonacich centrality measure, which helps to identify the various interaction effects. Therefore, in empirical applications, sometimes a spatial weights matrix without row-normalization would be appropriate. For estimation procedure in spatial econometrics, Kelejian and Prucha (2010) consider implications on the parameter space of the SAR model when the spatial weights matrix is not row-normalized.

8 However, for the case with multiple spatial weights matrices, when T is not really small, in order to accommodate spatial expansions, the IVs might be too many in order to be practical. This finite sample issue is presented in the Monte Carlo section.

9 Proofs for lemmas and more Monte Carlo results are provided in a supplement file available on request (see Appendix E).