DOKUZ EYLÜL UNIVERSITY

GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

A COMPARISON OF GMM ESTIMATORS FOR LINEAR DYNAMIC PANEL MODELS

by

Almila HACIOĞLU

January, 2017

İZMİR

A COMPARISON OF GMM ESTIMATORS FOR LINEAR DYNAMIC PANEL MODELS

A Thesis Submitted to the
Graduate School of Natural and Applied Sciences of Dokuz Eylül University
In Partial Fulfillment of the Requirements for the Master of Science in
Statistics, Statistics Program

by
Almila HACIOĞLU

January, 2017
İZMİR

M.Sc THESIS EXAMINATION RESULT FORM

We have read the thesis entitled **"A COMPARISON OF GMM ESTIMATORS FOR LINEAR DYNAMIC PANEL MODELS"** completed by **ALMİLA HACIOĞLU** under supervision of **PROF. DR. ESİN FİRUZAN** and we certify that in our opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Prof. Dr. Esin FİRUZAN

Supervisor

Doç. Dr. Adil OĞUZHAN

(Jury Member)

Doç. Dr. A. Fırat ÖZDEMİR

(Jury Member)

Prof.Dr. Emine İlknur CÖCEN
Director
Graduate School of Natural and Applied Sciences

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude and thanks to my advisor, Prof. Dr. Esin FİRUZAN for her support, direction, encouragement that she gave to me, positivity and for being cheerful all the time. I am honored to be a student of her.

I wish to thank to Beste BEYAZTAŞ and Ufuk BEYAZTAŞ for their support and helps.

Finally, special thanks to Aylin HACIOĞLU, my mother, Hasan HACIOĞLU, my father and Ferit HACIOĞLU, my brother for their endless support and love during my studies and thanks to Büşra SEVİNÇ and Özge LİMONCU for their motiviation and supporting me in every way when I need.

Almila HACIOĞLU

A COMPARISON OF GMM ESTIMATORS FOR LINEAR DYNAMIC PANEL MODELS

ABSTRACT

Panel data is a data set that consists of cross-sectional units appearing in different time dimensions. Dynamic adjustment processes include specification of lagged values of the covariates and the dependent variable, or both and the processes has become popular application in time series regression models. Generalized method of moment (GMM) is an estimation procedure that have large sample properties. The method allows to be specified while avoiding often unwanted or unnecessary assumptions. GMM estimation method is a standard approach to parameter estimation method that has many optimal properties in estimation: consistency, efficiency and asymptotically normal.

In the thesis, a Monte Carlo simulation study is conducted for comparing the performance of the Arellano-Bond (AB), Arellano-Bover (ABO) and Blundell-Bover (BB) GMM estimators. Two different distributions for residuals are considered for the comparison of three GMM estimators to illustrate the effect of different variance size on the GMM estimators: Standard Normal and t(5) distributions.

AB, ABO and BB are compared bias and standard error (SE) of autoregressive coefficient, coverage probabilities, relative efficiencies, mean squared error (MSE) under two different distributions.

This thesis addresses that BB has great advantages according to AB and ABO when small individual sample size, short time period even high autoregressive parameter in linear dynamic panel models.

Keywords: Panel data models, Linear dynamic panel data models, GMM estimation method, Arellano-Bond GMM estimator, Arellano-Bover GMM estimator, Blundell-Bond GMM estimator.

DOĞRUSAL DİNAMİK PANEL MODELLERİ İÇİN GMM KESTİRİCİLERİNİN KARŞILAŞTIRILMASI

ÖZ

Panel veri, farklı zaman boyutlarındaki yatay kesit gözlemlerden oluşan bir veri setidir. Dinamik süreçler bağımlı değişkenlerin ve bu değişkenlerin gecikmeli değerlerini içermektedir. Bu süreçler zaman serisi regresyon modellerinde popüler bir uygulama haline gelmiştir. GMM büyük örneklem özelliklerine sahip olan bir kestirim yöntemidir. Bu yöntem istenmeyen ve gereksiz varsayımlardan kaçınılmasına izin verir. GMM kestirimi, parametre kestirimi yapmak için birçok optimal özelliğe sahip standart bir yaklaşımdır. GMM kestiricisi tutarlılık, etkinlik ve asimtotik normallik özelliklerine sahiptir.

Bu tezde, AB, ABO ve BB GMM kestiricilerinin performanslarının karşılaştırılması için bir Monte Carlo benzetim çalışması yapılmıştır. GMM kestiricilerinin karşılaştırılmasında GMM kestiricilerinde farklı varyans büyüklüklerini görebilmek için artıkların iki farklı dağılımı kullanılmıştır: Standart Normal dağılım ve t(5) dağılımı. GMM kestiricileri, iki farklı dağılım altında otoregresif katsayının yanlılığı ve standart hatası, kapsama olasılıkları, göreceli etkinlikleri ve hatakareler ortalamasına göre karşılaştırılmıştır.

Bu tez, birim büyüklüğü küçük ve zaman periyodu kısa olduğunda ve otoregressif parametreleri yüksek olsa bile BB GMM kestricisinin AB ve ABO GMM kestiricilerine göre önemli avantajlara sahip olduğunu göstermiştir.

Anahtar kelimeler: Panel veri modelleri, Doğrusal dinamik panel veri modelleri, GMM tahmin yöntemi, Arellano-Bond GMM tahmincisi, Arellano-Bover GMM tahmincisi, Blundell-Bond GMM tahmincisi.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

Page

CHAPTER ONE
INTRODUCTION

This thesis considers linear dynamic panel data models and GMM estimation in the linear dynamic panel data analysis. Time series data is a set of observations collected at usually discrete and equally sapaced time intervals. Cross-sectional data is observations that come from different individuals at a single point in time. Panel data is consist of cross-sectional units appearing in different time dimensions. Panel data know as longitudinal or crosssectional time-series data.

This Chapter reviews the recent literature on dynamic panel data models. Kiviet (1995) has interested with various estimation methods for panel data includes lagged dependent explanatory variables, when the number of observations in the time dimension (T) gets large. Ahn and Schmidt (1995) have proposed the moment conditions. Bond (2002) have studied dynamic panel data analysis have small time dimensions and large individual dimensions. Blundell and Bond (1998) have proposed the hypothesis is required for the consistency of the GMM-system estimator. Mileva (2007) has implemented Arellano-Bond GMM estimation in linear dynamic panel data models. Santos and Barrios have simulated data to investigate both the small and large sample properties of the within-groups estimator and the first difference generalized method of moments estimator of a dynamic panel data model. Calzolari and Magazzini (2013) have proposed propose an LM approach for testing the mean stationarity assumption in dynamic panel data models. Kripfganz and Schwarz have proposed a two-stage estimation procedure to identify the effects of time-invariant regr gressors in a dynamic version of the Hausman-Taylor model providing analytical standard error adjustments for the second-stage.

Because of the various endogeneity problems, least squares based inference methods are biased and inconsistent. Therefore, it has developed Instrumental Variables (IV) methods or the Generalized Method of Moments (GMM), which produce consistent parameter estimates for a finite number of time periods, T, and a large cross-sectional dimension, N. GMM that have large sample properties is an estimation procedure that allows to be specified while avoiding often unwanted or unnecessary assumptions. Researchers has found it useful that GMM estimators can be constructed without specifying the full data generating process Hence, GMM estimation have become widely used, for the following reasons.

The method of moments approach to parameter estimation dates back more than 100 years (Stigler, 1986). Anderson and Hsiao (1981), Holtz-Eakin (1988), Holtz-Eakin, Newey and Rosen (1988), Arellano and Bond (1991) proposed using a GMM where all past outcomes with lags of two or higher are used as instruments for the differenced equations at each time period. is to first difference the equation, and then use instrumental variables (IV). Books with good discussions of GMM estimation with a wide array of applications include: Arellano (2003), Hall (2005. For a theoretical treatment of this method see Hansen (1982) along with the self contained discussions in the books. See also Ogaki (1993) for a general discussion of GMM estimation and applications, and see Hansen (2001) for a complementary entry that, among other things, links GMM estimation to related literatures in statistics. For a collection of recent methodological advances related to GMM estimation see Ghysels and Hall (2002).

CHAPTER TWO

PANEL DATA MODELS

2.1 Panel Data Structures

Cross-sectional data is a type of data that collected by observing many subjects (such as individuals, firms, countries, or regions) at the single point of time in statistics and econometrics. A time series data is a type of well-defined data items obtained through repeated measurements over time.

Panel data consist of cross-sectional units appearing in different time dimensions. Therefore, observations in panel data involve two dimensions: a cross-sectional dimension N, and a time-series dimension T. Panel data is known as longitudinal or crosssectional time-series data. The identification of time series parameters usally based on stationarity, predeterminedness and uncorrelated shocks. Unlike time series, cross-sectional parameters address exogenous instrumental variables and random sampling. By combining the time series and cross-sectional dimensions, panel data sets have developed identification adjustments. Many recent studies have focused on panel data sets.

Panel data is divided into balanced panel and unbalanced panel.

A balanced panel: A balanced panel is panel set that has all its observations. That is, all individuals are presented in all periods in balanced panel data sets. A special case of balance panel is a fixed panel.

An unbalanced panel: An unbalanced panel has some missing data for some time periods or some entities. That is, individuals may beobserved different numbers of times in unbalanced panel data sets.

There are some terms need to be known about panel data analysis:

Individual effect: Variables that reflect the characteristics of the units are called individual effect. Individual effect is a variable that fixed according to time.

Time effect: Variables that reflect the characteristics of the time are called time effect. Time effect is a variable that fixed according to units.

Endogeneity: Endogeneity occurs in a multiple regression if explanatory variables are correlated with error term.

Exogeneity: Exogeneity is one of important assumptions in panel data analysis. It means that explanatory variables are uncorrelated with error term.

Heterogeneity: An heterogeneous panel data model is a model in which all parametersvary across individuals.

Short panel: Small panel is a panel that known as T small and $N \to \infty$.

Long Panel: Long panel is a panel that known as $T \to \infty$ and N small or $N \to \infty$.

Advantages and Disadvantages of Panel Data

Panel data have several advantages and drawbacks over cross-sectional or time-series data.

An important advantage of panel data is increasing attention in estimation. This is the result of an increase in the number of observations due to blending time periods of data for each individual. It is obtained more accurate results for model parameters with panel data analysis. Panel data usually contains more degrees of freedom and less multicollinearity than cross-sectional data which may be viewed as a panel with $T = 1$, or time series data which is a panel with $N = 1$. Panel data has greater capacity than a cross-section or time series data for capturing the complexity.It can be constructed and tested more complicated hypotheses. It controls the impact of omitted variables and contains information on both the intertemporal dynamics and the individuality of the entities that may allow to control the effects of missing or unobserved variables. Panel data uncovers dynamic relationships and generates more accurate predictions. With panel data, we can rely on the inter-individual differences to reduce the collinearity between current and lagged variables to estimate unrestricted time-adjustment patterns (e.g. Pakes & Griliches (1984)). Panel data is suitable for investigating the “homogeneity” versus “heterogeneity”.

Panel data simplifies computation and statistical inference, since when time series data are not stationary, the large sample approximation of the distributions of the least-squares or maximum likelihood estimators are no longer normally distributed, (e.g. Anderson (1959); Dickey & Fuller (1979;1981); Phillips & Durlauf (1986)).

The important drawback of panel methodology is to control the impact of unobserved heterogeneity that represented by the incidental parameters. The appropriate analytical methods have not developed yet, for more sophisticated models that permit more general forms of correlation among the repeated measurements. There is a lack of available computer software for application of these more complex statistical models.

2.2 Static Linear Panel Data Models

The effects of observed explanatory variables, x, are identical across cross-sectional units, i, over time, t, while the effects of variables can be divided into the individual-specific effects, α_i, time-specific effects, λ_t.

The standard static model with *i=1, . . . , N* (individuals), *t=1, . . . ,T* (time), as follows:

$$y_{it} = \alpha_i + x'_{it}\beta + \lambda_t + u_{it} \tag{2.1}$$

where $\alpha_i \sim IID(0, \sigma_\alpha^2)$ and $u_{it} \sim IID(0, \sigma_u^2)$ independent of each other and among themselves. y_{it} is a scalar dependent variable, x_{it} is Kx1 vector of independent variables and the i-th observation on K explanatory variables, β is the parameter vector, $\alpha_i$denotes the unobserved individual-specific effects, λ_t denotes the time-specific effects, $\mathrm{u_{it}}$ is a scalar disturbance term and u_{it} zero mean, constant variance, and is uncorrelated across time andindividuals.

Static linear panel data models consist of two models: Fixed Effects Models and Random Effects Models.

2.2.1 Fixed Effects Models

Fixed-Effects explore the relationship between predictor and outcome variables within an entity. Each entity has its own individual characteristics that may or may not influence the predictor variables.

Fixed effects explicitly deals with that α_i may be correlated with x_{it} .

It is assumed that $\{(y_{i1},\dots,y_{iT},x_{i1},,\dots,x_{iT,}\ \alpha_i),\ i=1,\dots,N\}$ is a random sample and it is shown as in equation (2.2).

$$y_{it}=\alpha_i+x'_{it}\beta+u_{it} \tag{2.2}$$

Assumption 1: For fixed effects models, it is assumed strict exogeneity.

$$E(u_i|x_i,\alpha_i)=0\ \ (t=1,\dots,T)$$

where $u_i=(u_{i1},\dots,u_{iT})'$ and $x_i=(x'_{i1},\dots,x'_{iT})'$. It is observed y_{it} and the K x 1 vector of explanatory variables x_{i1}, but α_i is unobservable time-invariant regressor.

Assumption 2: The errors are conditionally homoskedastic and not serially correlated with explanatory variables.

$$Var(u_i|x_i,\alpha_i)=\sigma^2 I_T$$

2.2.2 Random Effects Models

The rationale behind random effects model is that, unlike the fixed effects model, the variation across entities is assumed to be random and uncorrelated with the predictor or explanatory variables included in the model. Random effects models effectively put α_i in the error term under the assumption that α_i is orthogonal to x_{it} and then accounts for the serial correlation in the composite error.

$$y_{it} = \beta_0 + x_{it}'\beta + \alpha_i + u_{it} \tag{2.3}$$

$$\epsilon_{it} = \alpha_i + u_{it} \tag{2.4}$$

Equaiton (2.3) includes random effects (α_i) and fixed effects (β). The term α_i is assumed to be a random variable, not a fixed, unknown parameter. Thus, the term $\alpha_i$is known as a random effect.

Assumption 1: α_i is unobservable time-invariant regressor and homoscedastic across individuals.

Assumption 2: It is assumed that the regressors x_{it} are uncorrelated with individual effects α_i and u_{it}, if If $E[x_{it}\epsilon_{it}] = 0$. That is, the explanatory variables are exogeneous.

Fixed Effects versus Random Effects:

The fundamental difference is between models with and without fixed effects. In the fixed effects models, α_i is possibly correlated with regressor x_{it} which can be endogenous, but in the random effects models, regressor x_{it} must be exogenous. In the fixed effects models, it can consistently estimated β for time-varying x_{it}, but if short panel, it cannot consistently estimated α_i and the prediction is not possible. Unlike fixed effects models, the prediction is possible in the random effects models.

2.3 Dynamic Linear Panel Data Models

Dynamic panel data models have become common thanks to the increased availability of longitudinal data. Dealing with dynamic adjustment processes including specification of lagged values of the covariates and the dependent variable, or both are popular application in time series regression models. Sufficient characterization of several economic dynamic adjustment processes are provided by inclusion of lagged dependent variable.

The model is dynamic due to the presence of the lagged dependent explanatory variable y_{it}, which has unknown coefficient ρ. It is considered the dynamic linear panel data model on N units observed over $T \geq 2$ time periods for any observation i from a sample of n observations and any time period t from a fixed number T of time periods: The following regression $AR(1)$ model in equation (2.5) presents the dynamic relationships with lagged dependent variable among the explanatory variables:

$$y_{it} = \rho y_{i,t-1} + x_{it}'\beta + \alpha_i + u_{it} \tag{2.5}$$

with i=1, . . . , N (individuals), t=1, . . . ,T (time).

ρ is a scalar, x_{it} is $1 \times K$ and β is $K \times 1$. The initial observations of the dependent variable, y_{i0}, and the regressors, x_{i0}, are assumed to be observed. We will assume that the u_{it} follow a one-way error component model:

$$\varepsilon_{it} = \alpha_i + u_{it} \tag{2.6}$$

where $\alpha_i \sim \mathrm{IID}(0, \sigma_\alpha^2)$ and $u_{it} \sim \mathrm{IID}(0, \sigma_u^2)$ independent of each other and among themselves.

$Y_t = [Y_{1t}, ..., Y_{nt}]$ and $Y_{it} = [y_{i0}, ..., y_{it}]$ are random vectors that values of y_{it} across time and observations and α_i are unobserved effects constant over time and is correlated with the lagged dependent variable by construction, α_i also has unobserved heterogeneity. The dynamic panel data regression model described in (2.5) and (2.6) is characterized by two sources of persistence over time. One of them is the autocorrelation due to the presence of a lagged dependent variable among the regressors and the mother one is individual effects characterizing the heterogeneity among the individuals.

Due to the addition of lagged dependent variable, there are some basic problems. Since y_{it} is a function of α_i, it immediately follows that $y_{i,t-1}$ is also a function of α_i. Therefore, $y_{i,t-1}$, a right-hand regressor in (2.5), is correlated with the error term.

Some assumption should be provided in the dynamic panel models. The dynamic panel data models without covariates have the following assumptions:

1. Stationarity, $|\rho| < 1$.
2. $E[u_{it}|\, y_{i0}, \ldots, y_{i,t-1}, \alpha_i] = 0$, α_i and u_{it} are independent of each other.
3. No serial correlation ($u_{it} \sim IID(0, \sigma_u^2)$)
4. Homoskedasticity assumption ($\alpha_i \sim IID(0, \sigma_\alpha^2)$).

with

$E(u_{it}u_{js})= 0$, $i \neq j$ or $t \neq s$

$E(\alpha_i\alpha_j) = 0$, $i \neq j$,

$E(\alpha_i u_{jt}),= 0$, $\forall$ i,j,t

$E(x_{it}\, u_{js}) = 0$, $\forall$ i,j,t,s

$E(x_{it}\, \alpha_j)$ = unknown, $\forall$ i,j,t .

The first difference of variables removes both the constant term and the individual effect unobserved heterogeneity from the model as shown in equaiton (2.7).

$$y_{it} - \bar{y}_i = \rho(y_{i,t-1} - \bar{y}_{i,-1}) + (x_{it} - \bar{x}_i)\beta + (u_{it-}\bar{u}_i) \qquad (2.7)$$

$$\Delta y_{it} = \rho\Delta y_{i,t-1} + \Delta x_{it}\beta + \quad \Delta u_{it} \tag{2.8}$$

There is still correlation between the differenced lagged dependent variable and the disturbance process, the former contains $y_{i,t-1}$ and the latter contains u_{it-1} introduces a correlation between the transformed lag $(y_{i,t-1} - \bar{y}_{i,-1})$ and the transformed error $(u_{it}{-}\bar{u}_i)$ .Because the average error $(\bar{u}_i = \sum_{\mathrm{i}=1}^{T} u_{it})$ and the estimated ρ remains biased.. u_{it-1} becomes a smaller component of the average error term as T increases. In other words, with higher T the correlation between the lagged dependent variable and the regressionerrors becomes smaller.

CHAPTER THREE

GMM ESTIMATION OF LINEAR DYNAMIC PANEL DATA MODELS

The Generalized Method of Moments (GMM) is an estimation method which can be thought of just as a generalization of the classical Method of Moments. GMM is a set of population moment conditions. GMM is a class of estimators which consists of with exploiting the sample moment conditions in the data generating model. GMM estimators can be constructed without defining the full data generating process, so it is practical in according to researchers. GMM estimators have large sample properties. GMM estimators play an important rolefrom the point ofview of making asymptotic efficiency since characterizing large sample properties of this estimators is easy for comparisons.

The GMM is a statistical method that combines observed economic data with the information in population moment conditions to produce estimates of the unknown parameters of the economic model. GMM is an alternative which depends on minimal set of statistical assumptions. This method can be used in economic models to be specified without having many assumptions, such as specifying a particular distribution for the errors and serial correlations of errors.

GMM extends the classical setup in two important ways: The first is to formally treat the problem of having two or more moment conditions which have information about unknown parameters. GMM allows estimation and inference in systems of Q equations with P unknowns, $P \leq Q$. The second important generalization of GMM is that quantities other than sample moments can be used to estimate the parameters. GMM exploits laws of large numbers and central limit theorems to establish regularity conditions for many different “moment conditions” that may or may not actually be moments. In models for which there are more moment conditions than model parameters, GMM estimation provides a straightforward way to test the specification of the proposed model. This is an important feature that is unique to GMM estimation.

There are two main GMM estimators: GMM-difference estimator and GMM-system estimator. The former is proposed by Arellano and Bond (1991). The latter is seperated to two parts according to instrument matrix. The Arellano-Bover (1995) GMM-system estimator uses differenced observationsΔy_{it-1}, whereas Blundell and Bond (1998) GMM-system estimator uses both original y_{it} and differenced observations Δy_{it-1} in the instrument matrix.

3.1 GMM Estimation of Linear Dynamic Panel Data Models

GMM estimation is used for some dynamic panel models because it allows a flexible specification of the instruments, since this flexibility produces consistent parameter estimates for a finite number of time periods, *T*, and a large cross-sectional dimension, *N* (see e.g. Arellano and Bond, 1991; Arellano and Bover, 1995; Blundell and Bond, 1998).

When there is unobserved unit-specific heterogeneity, it is often hard to disentangle the effects of the observed and the unobserved time-invariant heterogeneity. Fixed and random effects estimators cannot be used due to the various endogeneity problems., when the time dimension is short. That is, fixed and random effects estimators are biased and inconsistent. Therefore, it is common practice in empirical work to apply the generalized method of moments (GMM) framework proposed by Arellano and Bond (1991), Arellano and Bover (1995), and Blundell and Bond (1998). However, as Binder et al. (2005) and Bun and Windmeijer (2010) emphasize, GMM estimators might suffer from a weak instruments problem when the autoregressive parameter approaches unity or when the variance of the unobserved unit-specific effects is large.

In this thesis, AR(1) linear dynamic panel model is investigated for comparing GMM estimators. AR(1) linear dynamic panel model form:

$$y_{it} = \rho y_{i,t-1} + \alpha_i + u_{it} \tag{3.1}$$

The following assumptions on the error components structure:

1. $E(\alpha_i) = 0$; $E(u_{it}) = 0$; $E(u_{it}\alpha_i) = 0$ for $i = 1, \ldots, N$ and $t = 1, \ldots, T$

2. $E(u_{it}u_{is}) = 0$ for $i = 1, \ldots, N$ and $\forall t \neq s$

There are also additional restrictions concerning the initial conditions

3. $E(u_{it}y_{it}) = 0$ for $i = 1, \ldots, N$ and $t = 1, \ldots, T$

For the estimation of such an autoregressive model described by equation (3.1) with assumptions 1-3, Arellano and Bond (1991) proposed transforming the model into first-differences and utilizing the moment conditions:

$$E[y_{it-j}\Delta\varepsilon_{it}] = 0 \quad (3.2)$$

Moment conditions in (3.1) and (3.2) exploit the assumption of lack of serial correlation in ε_{it}, that can be tested using both the Sargan/Hansen test and the test for autocorrelation of first difference residuals proposed by Arellano and Bond (1991).

Blundell and Bond (1998) introduce an additional assumption on the initial condition:

$$E[\varepsilon_{i2}\Delta y_{i1}] = 0 \quad (3.3)$$

labeled 'mean-stationarity' assumption. As ε_{it} is uncorrelated over time, we can also state condition (3.3) in terms of α_i as

$$E[\alpha_i\Delta y_{i1}] = 0 \quad (3.4)$$

that makes more clear that differencing y_{it} eliminates the correlation with α_i, thus excluding correlation between deviations of y_{it} from its long run mean and the individual effect α_i (Roodman,2009).

If condition (3.3) is also satisfied, the following $T - 1$ moment conditions ($t \geq 2$) also hold:

$$E[\Delta y_{it-1}\varepsilon_{it}] = 0 \tag{3.5}$$

As mentioned before, there are two types of GMM estimators. In the following part, assumptions of GMM-system estimator and GMM-differenced estimator are given for AR(1) linear dynamic panel model.

The GMM-difference estimator is used the model which is shown in equations (2.5) and (2.6) under the assumption of lack of serial correlation in ε_{it}. First, in order to wipe out the individual effect α_i, differences of the data are considered. It is assumed that α_i and ε_{it} are uncorrelated for all *i*, and ε_{it} is uncorrelated over time (including $t = 0$).

The following assumptions of the GMM:

1. $E(y_{i0}\varepsilon_{it}) = \sigma_{0\varepsilon}$, constant over time and across units;
2. $E(\varepsilon_{it}\varepsilon_{is}) = \sigma_\alpha^2$ for each $s \neq t$.

Blundell and Bond (1998) make the additional assumption on the initial observation (see also Arellano and Bover, 1995):

3. $E(\varepsilon_{i2}\Delta y_{i1}) = 0$. (excluding cases in which y_0 is not 'mean stationary").

The GMM-system estimator that combines (3.2) and (3.4) also performs better in terms of small sample bias (Hayakawa, 2007). As a result, the GMM-system method is widely used in empirical analysis, thus requiring validity of the 'mean stationarity' assumption. This is customarily tested via the Sargan/Hansen test by looking (i) at the validity of the full set of moment conditions (3.2) and (3.5), or (2) by computing a difference-in-Sargan/Hansen test that takes into account the difference between the value of the minimized GMM criterion function of the GMM-system estimator and the GMM-difference estimator (Bond, Bowsher, and Windmeijer, 2001).

Detail information of Arellano-Bond (AB), Arellano-Bover (ABO) and Blundell-Bond (BB) estimators is given in the following sections:

3.1.1 Arellano and Bond Estimator (AB)

Arellano and Bond (1991) proposed that instruments can be obtained in a dynamicpanel data model in case of the orthogonality conditions that exist between lagged values of y_{it} and the disturbances u_{it} in equation (3.6). The following simple autoregressive model without regressors:

$$y_{it} = \rho y_{i,t-1} + \alpha_i + u_{it} \tag{3.6}$$

where $\varepsilon_{it} = \alpha_i + u_{it}$ with $\alpha_i \sim \mathrm{IID}(0, \sigma_\alpha^2)$ and $u_{it} \sim \mathrm{IID}(0, \sigma_u^2)$, independent of each other and among themselves. To obtain a consistent estimate of ρ as $N \to\infty$ with T fixed, first differenced values are implemented to the analysis; thus individual effects has been removed.

$$y_{it} - y_{i,t-1} = \rho(y_{i,t-1} - y_{i,t-2}) + (u_{it} - u_{i,t-1}) \tag{3.7}$$

$(u_{it} - u_{it-1})$ has MA(1) process with unit root. For t = 3, the first period:

$$y_{i3} - y_{i2} = \rho(y_{i2} - y_{i1}) + (u_{i3} - u_{i2})$$

y_{i1} is a valid instrument, because it is highly correlated with $(y_{i2}-y_{i1})$ and not correlated with $(u_{i3}-u_{i2})$ if u_{it} are not serially correlated. For t = 4, the second period:

$$y_{i4} - y_{i3} = \rho(y_{i3} - y_{i2}) + (u_{i4} - u_{i3})$$

y_{i2} as well as $y_{i1}$are valid instruments for $(y_{i3}-y_{i2})$, becausey_{i2} and $y_{i1}$are not correlated with $(u_{i4}-u_{i3})$. The period can continue for period T and the set of valid instruments becomes$(y_{i1}, y_{i2}, \ldots, y_{i,T-2})$.

This instrumental variable procedure doesn't explain the differenced error term in

$$E(\Delta u_i \Delta u_i') = \sigma_u^2 (I_N \otimes G) \tag{3.8}$$

where $\Delta u_i' = (u_{i3} - u_{i2}, \ldots, u_{it} - u_{i,t-1})$ and toeplitz matrix is

$$G = \begin{pmatrix} 2 & -1 & 0 & \ldots & 0 & 0 & 0 \\ -1 & 2 & -1 & \ldots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & -1 & 2 & -1 \\ 0 & 0 & 0 & \ldots & 0 & -1 & 2 \end{pmatrix} \tag{3.9}$$

where $(T-2)\ x\ (T-2)$, because Δu_i is MA(1) with unit root.

For the estimation of such an autoregressive model described by (3.1) with assumptions 1-3, Arellano and Bond (1991) proposed transforming the model into first-differences and utilizing m_d= 0.5(T -1)(T -2) moment conditions:

$$E(Z_{di}' \Delta \varepsilon_i) = 0 \tag{3.10}$$

where Z_{di} is a $(T-2) \times m_d$ instrument matrix and $\Delta\varepsilon_i$is a (T - 2) column vector of residuals for the first difference equation of individual i.

$$Z_{di} = \begin{bmatrix} [y_{i1}] & & & 0 \\ & [y_{i1}, y_{i2}] & & \\ & & \ddots & \\ 0 & & & [y_{i1}, \ldots, y_{i,T-2}] \end{bmatrix} \tag{3.11}$$

The matrix of instruments is $Z = [Z'_{d1}, \ldots, Z'_{dN}]'$and the moment equations are given by $E(Z'_{di}\Delta u_i) = 0$.

The differenced equation (3.7) in vector form by *Z':*

$$Z'\Delta y = Z'(\Delta y_{-1})\rho + Z\Delta u \tag{3.12}$$

The Arellano and Bond's (1991) one-step consistent estimator is in the following:

$$\hat{\rho}_1 = \left[(\Delta y_{-1})' Z (Z'(I_N \otimes G)Z)^{-1} Z'(\Delta y_{-1})\right]^{-1} \times \left[(\Delta y_{-1})' Z (Z'(I_N \otimes G)Z)^{-1} Z'(\Delta y)\right] \tag{3.13}$$

The optimal GMM estimator of ρ_i for $N \rightarrow \infty$ and T fixed using the moment equations yields the same expression as in (3.13) except that

$$Z'(I_N \otimes G)Z = \sum_{i=1}^{N} Z_{id} G Z_{di} \tag{3.14}$$

is replaced by

$$V_N = \sum_{i=1}^{N} Z(_i\Delta u_i)(_i\Delta u_i)' Z_i \tag{3.15}$$

The GMM estimator requires no knowledge about the initial conditions or the distributions of $u_i$and α_i. Δu_{is} replaced by differenced residuals obtained from the first consistent estimator $\hat{\rho}_1$. The resulting estimator is the two-step Arellano and Bond (1991) GMM estimator:

$$\hat{\rho}_2 = \left[(\Delta y_{-1})' Z (\hat{V}_N)^{-1} Z' (\Delta y_{-1})\right]^{-1} \times \left[(\Delta y_{-1})' Z (\hat{V}_N)^{-1} Z' (\Delta y\right] \quad (3.16)$$

A consistent estimate of the asymptotic $Var(\hat{\rho}_2)$ is given by the first term in the equation (3.6):

$$Var(\hat{\rho}_2) = \left[(\Delta y_{-1})' Z (\hat{V}_N)^{-1} Z' (\Delta y_{-1})\right]^{-1} \quad (3.17)$$

$\hat{\rho}_1$ and $\hat{\rho}_2$ are asymptotically equal if $u_{it} \sim IID(0, \sigma_u^2)$.

When T is relatively large, estimating that many parameters causes the Arellano and Bond estimator to suffer from poor small sample properties in terms of accuracy and inference, which was studied in the context of cross-sectional independence in Alvarez & Arellano (2003); Windmeijer (2005).

Arellano-Bond estimator was found to be ineffcient since it does not make use of all available moment conditions (see Ahn & Schmidt, (1995)); it also has very poor finite sample properties in dynamic panel data models with highly autoregressive series and a small number of time series observations [see Alonso-Borrego & Arellano, (1999); Blundell & Bond, (1998)], since the instruments in those cases become less informative.

3.1.2 Arellano and Bover Estimator (ABO)

Arellano and Bover (1995) develop a GMM framework for looking at efficient instrumental estimators for dynamic panel data models. ABO estimator is one of the System GMM estimators. This estimator uses additional moment conditions, homoscedasticity serial correlation between u_{it} and endogenous variable $y_{i,t-1}$ or stationary initial conditions. Hence, this estimator requires validity of the 'mean stationarity' assumption.

The considered model and moment conditions, which are given in equation (3.6) and (3.12) respectively, are valid for this estimator.

Consider one additional assumption:

$$E[\alpha_i \Delta y_{i2}] = 0 \tag{3.18}$$

a restriction on the initial conditions (see Arellano and Bover (1995)), which holds when the process is mean stationary. This condition, together with assumptions on the error components structure yields the following m_d = 0.5(T-1)(T-2) moment conditions, based on which the level (LEV) GMM estimator is constructed, for the level equation:

$$E(Z'_{di}\Delta\varepsilon_i) = 0 \tag{3.19}$$

The only difference between AB and ABO is the instrument matrix. Z_{di} becomes:

$$Z_{di} = \begin{bmatrix} [\Delta y_{i2}] & 0 & 0 & \dots & 0 & \dots & 0 \\ 0 & [\Delta y_{i2}] & [\Delta y_{i3}] & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & [\Delta y_{i2}] & \dots & [\Delta y_{i,T-1}] \end{bmatrix} \tag{3.20}$$

where Z_{di} is a $(T-2) \times m_d$ instrument matrix and u_i is a (T - 2) column vector of residuals for the level equation of individual i.

To guarantee that $\Delta y_{i,t-1}$ is not correlated with α_i, we require the initial conditions restriction is required as follows:

$$E\left[\left(y_{i1} - \frac{\alpha_i}{1-\rho}\right)\right]_{it} = 0 \tag{3.21}$$

3.1.3 Blundell and Bond GMM Estimator (BB)

Blundell & Bond, (1998) work on the importance of exploiting the initial condition in generating efficient estimators of the dynamic panel data model when T is small. They consider a simple autoregressive panel data model without exogenous regressors,

$$y_{it} = \rho y_{i,t-1} + \alpha_i + u_{it} \tag{3.22}$$

with $E(\alpha_i) = 0$, $E(u_{it}) = 0$ and $E(\alpha_i u_{it}) = 0$ for $i = 1, 2, \ldots, N$; $t = 1, 2, \ldots, T$. Blundell and Bond work on the case where $T = 3$, because there is only one orthogonality condition given by $E(y_{i1} u_{i3}) = 0$, In this case, the first-stage IV regression is obtained by running Δy_{i2} on y_{i1}. This regression can be obtained from (3.22) evaluated at $t = 2$ by subtracting y_{i1} from both sides of this equation.

Blundell & Bond, (1998) and Blundell, Bond & Windmeijer, (2001) point out that the instruments are obtained by the differenced GMM estimator become less informative when the value of the autoregressive coefficient ρ increases towards unity and when the variance of the individual effects σ_α^2 increases relative to the variance of the distinctive (idiosyncratic) errors, σ_u^2.

They proposed using what is called the system GMM estimator, which utilizes all the m_d moment conditions for the difference GMM estimator and a non-redundant subset of the moment conditions for the level equation, a total of $m_s = m_d + (T - 2)$ moment conditions as follows:

$$E(Z'_{si}\tilde{u}_i) = 0 \tag{3.23}$$

where $\tilde{u}_i$ is a vector of error terms in the first-difference equation followed by those in the levelequation of individual i, i.e., $\tilde{u}_i = (\Delta u'_{i,}, u'_i)'$, and Z_{si} is a $2(T-1) \times m_s$ instrument matrix for both equations of individual i

$$Z_{si} = \begin{bmatrix} Z_{di} & 0 & 0 & \dots & 0 \\ 0 & \Delta y_{i2} & 0 & \dots & 0 \\ 0 & 0 & \Delta y_{i3} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \Delta y_{i,T-1} \end{bmatrix} \tag{3.24}$$

:

The one-step system GMM estimator takes the form:

$$\hat{\rho}_1 = \left[(\bar{y}'_{-1})' Z_s W_s Z'_S \bar{y}_{-1}\right]^{-1} \times \left[(\bar{y}'_{-1})' Z_s W_s Z'_S \bar{y}\right] \tag{3.25}$$

where $(\Delta y'_i y'_i)' > 0$ and $W_s = Z'_s (I_N \otimes H_s)$ with

$$H_s = \begin{pmatrix} H_d & 0 \\ 0 & I_{T-2} \end{pmatrix} \tag{3.26}$$

where H_d is a (T -2) square matrix which has twos in the main diagonal, minus ones in the first subdiagonals, and zeros otherwise.

The two-step system GMM estimator is given by replacing the weighting matrix W_s with

$$W_s = \left(\frac{1}{N} \sum_{i=1}^{N} Z'_{si} \hat{\hat{u}}_{1i} \hat{\hat{u}}'_{1i} Z'_{si} \right)^{-1} \tag{3.27}$$

where $\hat{\hat{u}}_{1i}$ are first-step residuals.

$$\tau = \frac{(\sigma_U^2 c)^2}{\sigma_\alpha^2 + \sigma_u^2 c} \to 0 \ \text{ as } \ \rho \to 1 \tag{3.28}$$

where $c = (1-\rho)/(1+\rho)$. The bias term effectively scales the estimated coefficient on the instrumental variable y_{i1}towards zero. They also find that the *F*-

statistic of the first-stage IV regression converges to $\mathcal{X}_1^2$ with noncentrality parameter.

As $\tau \rightarrow 1$, the IV estimator performs poorly. Hence, Blundell and Bond attribute the bias and the poor precision of the first-difference GMM estimator to the problem of weak instruments and characterize this by its concentration parameter τ.

Blundell and Bond (1998) show that an additional mild stationarity restriction onthe initial conditions process allows the use of an extended system GMM estimator that is used lagged differences of y_{it} as instruments for equations in levels (Arellano and Bover (1995)). The system GMM estimator is shown to have dramatic efficiency gains over the basic first-difference GMM as $\rho \rightarrow 1$ and $(\sigma_\alpha^2/\sigma_u^2)$ increases.

The levels restrictions suggested by Arellano and Bover (1995) remain informative in cases where first differenced instruments become weak. Results improve for first-difference GMM as T increases. However, with short T and persistent series, the Blundell and Bond findings support the use of the extra moment conditions.

System GMM estimators presented in Ahn and Schmidt (1995), Arellano and Bover (1995), and Blundell and Bond (1998) are similar to the Arellano and Bond estimator, but additional moment conditions are reauired based on additional assumptions of homoscedasticity serial correlation or stationary initial conditions.

System GMM estimator has been shown to perform much better than the difference GMM estimator in terms of finite sample bias and mean squared error, as well as with regard to coefficient estimator standard errors, since the instruments used for the level equation are still informative as the autoregressive coefficient approaches unity (for detail information, see Blundell and Bond (1998) and Blundell, Bond and Windmeijer (2001)).

CHAPTER FOUR
SIMULATION STUDY

In this chapter, a simulation study has been designed to illustrate how the Arellano-Bond, Arellano-Bover and Blundell-Bond GMM estimators performs. These estimators are performed into two steps, therefore in the tables AB1 and AB2 abbreviations for first step and second step of Arellano-Bond estimator, respectively. Likely, ABO1 and ABO2 are used for Arellano-Bover estimator's step and also BB1 and BB2 represent the Blundell-Bond estimator's two steps.

GMM estimation method is a standard approach to parameter estimation and inference in statistics and has many optimal properties in estimation: consistency (true parameter value that generated the data recovered asymptotically); efficiency (lowest-possible variance of parameter estimates achieved asymptotically); asymptotically normal.

The comparison criterias are chosen bias and standard error (SE) of ρ, coverage probabilities, relative efficiencies, mean squared error (MSE) underlying two different distributions. To show the effect of the distribution of error terms on the estimation, two different distributions are considered for the comparison of three GMM estimators: Standard Normal and t(5) distributions.

Relative efficiency is measured by the ratio R, where R is the variance of one GMM estimator is divided by variance of the another GMM estimator. The relative efficiency of two unbiased estimators $\hat{\theta}_1$, $\hat{\theta}_2$ is the ratio of their variances.

$$R = \frac{Var(\hat{\theta}_1)}{Var(\hat{\theta}_2)}$$

If this ratio is greater than one, it means $\hat{\theta}_1$ estimator has a larger variance and so the $\hat{\theta}_2$estimator will be more appropriate. In order to compare GMM estimators in terms of relative efficiency, estimators that have R ratio less than one are preferred.

For this simulation, the number of individual sample size and time periods are investigated for three levels, respectively: $N = 50, 100, 500$ and $T = 5, 10, 15$. To 0,+see the autoregressive parameter effect, three different autoregressive coefficients are taken into consideration: $\rho = 0.1, 0.5, 0.9$. Nominal significance level is chosen 0.05. Aforementioned criterias are obtained based on 1000 replications.

R programming is used to generate data and to show some illustrations for all above conditions.

Standard Normal Distribution:

The standard normal distribution is a special case of the normal distribution. It is a normal distribution with zero mean and unit variance. The probability density function:

$$f(x \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\sigma^2\pi}} e^{-(x-\mu)^2/2\sigma^2} \qquad -\infty < x < \infty$$

where μ is mean of distribution, σ^2 is variance of distribution.

The normal random variable of a standard normal distribution is called a **z-score**. Every normal random variable *X* can be transformed into a *z* score via the following equation:

$$z = (x - \mu)/\sigma$$

where *X* is a normal random variable, μ is the mean of *X*, and σ is the standard deviation of *X*.

t(5) Distribution:

Student's *t*-distribution is any member of a family of continuous probability distributions that arises when estimating the mean of a normally distributed population in situations where the sample size is small and population standard deviation is unknown. Student's *t*-distribution has the probability density function given by

$$f(x) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi}\,\Gamma(\frac{\nu}{2})}(1+\frac{t^2}{\nu})^{-\frac{\nu+1}{2}} \qquad -\infty < x < \infty$$

where *v* is the number of degrees of freedom and Γ is the gamma function.

To show the similarity and dissimilarity of two distributions, the Figure 4.1. is drawn.

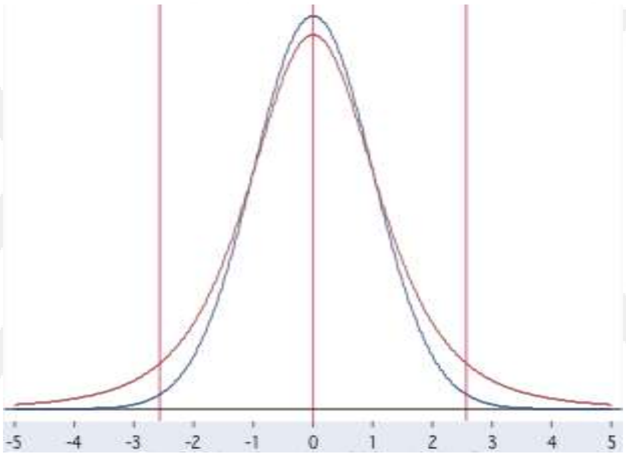

Figure 4.1 The shape of two distributions (SN(0,1) and t(5))

Blue line represents standard normal with 0 mean and 1 variance, red line represents t distribution with 5 degrees of freedom in Figure 4.1.

In Figure 4.1, it is easily seen that the shape of SN(0,1) and t(5) are similar to each other, but the variance of t distribution is slightly greater than standard normal. The tails of t distribution decrease more slowly than the tails of the normal distribution. The t distributions are more spread out than the normal. The spreading effect is huge for 1 degree of freedom. These distributions are chosen to illustrate the effect of different variance size on the GMM estimators.

The algorithm for the simulation study is given below:

Step 1: Definite number of simulation, number of individual (N), time period (T) autoregressive coefficient (ρ) and significant level (α).

Step 2: Generate the fixed effect (α_i) distributed with 0 mean and 1 standard deviation.

Step 3: Generate y_{it} matrix with N column and T row from AR(1) process for the standard normal distributed residuals.

Step4: Generate AB Difference Operator Matrix for AB GMM estimator.

Step 5: Generate AB instrument matrix with $y_{i1}, y_{i2}, . . . y_{i,t-2}$.

Step 6: Generate instrument matrix for ABO GMM estimation with $\Delta y_{i1, . . . ,} \Delta y_{i,t-2}$..

Step 7: Generate instrument matrix for BB GMM estimation with $y_{i1},, . . . y_{i,t-2}$ and $\Delta y_{i1, . . . ,} \Delta y_{i,t-2}$.

Step 8: Definite of the dimension of the simulated data.

Step 9: Get Δy_{it} and $\Delta y_{i,t-1}$ by applying the first differences.

Step 10: Generate the one step AB estimator using AB Difference Operator Matrix, AB instrument matrix, differenced values (Δy_{it}) and lagged of differenced values $\Delta y_{i,t-1}$.

Step 11: Calculate MSE, standard error and bias for the one step AB estimator.

Step 12: Calculate residual $\hat{u}_{it}$ using the first step AB estimator.

Step 13: Calculate $\Delta\hat{u}_{it}$ using $\hat{u}_{it}$.

Step 14: Generate the second step AB estimator using $\Delta\hat{u}_{it}$, $\hat{u}_{it}$, Δy_{it} and $\Delta y_{i,t-1}$ and instrument matrix of AB estimator.

Step 15: Calculate MSE, standard error and bias for the second step AB estimator.

Step 16: Generate the first step ABO estimator using ABO instrument matrix, y_{it} and $y_{i,t-1}$.

Step 17: Count MSE, standard error and bias for the first step ABO estimator.

Step 18: Calculate residual $\hat{u}_{it}$ using the first step ABO estimator.

Step 19: Generate the second step ABO estimator using ABO instrument matrix, y_{it}, $y_{i,t-1}$, and $\hat{u}_{it}$.

Step 20: Calculate MSE, standard error and bias for the second step ABO estimator.

Step 21: Calculate G (Toeplitz matrix) using AB difference operator matrix for BB estimator.

Step 22: Generate the first step BB estimator using G matrix, y_{it}, $y_{i,t-1}$, Δy_{it}, $\Delta y_{i,t-1}$ and BB instrument matrix.

Step 23: Calculate MSE, standard error and bias for the first step BB estimator.

Step 24: Calculate residual $\hat{u}_{it}$ using the first step BB estimator.

Step 25: Generate the second step BB estimator using $\hat{u}_{it}$, G matrix, y_{it}, $y_{i,t-1}$, Δy_{it}, $\Delta y_{i,t-1}$ and BB instrument matrix.

Step 26: Calculate MSE, standard error and bias for the second step BB estimator.

Step 27: Calculate and coverage probabilities and relative efficiencies for AB, ABO and BB estimators.

In terms of comparison of three GMM estimators, the results are given in Table 4.1, 4.3, 4.5, 4.7, 4.9, and 4.11 present parameter values, standard errors and coverage probabilities for different *N* and *T* values under two distributions of error terms: SN(0,1) and t(5). Table 4.2, 4.4, 4.6, 4.8, 4.10, and 4.12 shows estimated ratios of variance of three GMM estimators with each other for individual size N=50, 100 and 500, with *i.i.d*. SN(0,1) and t(5) distributions of residuals. In each case, the smallest three R ratio is marked and the estimators which have less standard error and lower coverage probabilities in comparison with other estimators.

The parameter values, standard errors and coverage probabilities are listed in Table 4.1.

Table 4.1 Parameter values, standard errors and coverage probabilities with *i.i.d.* distributed SN(0,1) with *N*=50. (**AB1-AB2:** Arellano-Bond GMM estimator for first and second step, **ABO1-ABO2:** Arellano-Bover GMM estimator for first and second step, **BB1-BB2**: Blundell-Bover GMM estimator for first and second step).

	ρ=0.1						**ρ=0.5**						**ρ=0.9**					
	AB1	**AB2**	**ABO1**	**ABO2**	**BB1**	**BB2**	**AB1**	**AB2**	**ABO1**	**ABO2**	**BB1**	**BB2**	**AB1**	**AB2**	**ABO1**	**ABO2**	**BB1**	**BB2**
T=5																		
$\hat{\boldsymbol{\rho}}$	0.0746	0.0753	0.1379	0.1279	0.1217	0.1192	0.4318	0.4369	0.5129	0.5155	0.4947	0.5026	0.6705	0.6542	0.8822	0.8851	0.8553	0.8639
$\boldsymbol{SE(\widehat{\rho})}$	0.1368	0.1175	0.1038	0.1125	0.0815	0.0819	0.1774	0.1526	0.1126	0.1112	0.0922	0.0902	0.3055	0.2554	0.1185	0.111	0.1002	0.0876
Coverage Probability	0.8770		0.8450		0.8050		0.8480		0.8420		0.8260		0.7550		0.9130		0.7870	
T=10																		
$\hat{\boldsymbol{\rho}}$	0.0722	0.0726	0.1865	0.1739	0.1429	0.1407	0.4501	0.4488	0.543	0.5379	0.5129	0.5119	0.7673	0.7636	0.8995	0.9012	0.8755	0.8769
$\boldsymbol{SE(\widehat{\rho})}$	0.0660	0.0274	0.0545	0.0271	0.0419	0.0136	0.073	0.0314	0.0543	0.0271	0.0432	0.0141	0.0905	0.0399	0.0444	0.0195	0.0388	0.0118
Coverage Probability	0.5420		0.3490		0.2790		0.4930		0.5030		0.3590		0.2770		0.5790		0.3360	
T=15																		
$\hat{\boldsymbol{\rho}}$	0.0738	0.0631	0.2222	0.2011	0.162	0.1373	0.4553	0.4337	0.5624	0.5573	0.5237	0.4984	0.8048	0.7931	0.9021	0.903	0.88	0.8658
$\boldsymbol{SE(\widehat{\rho})}$	0.0474	0.0416	0.0401	0.0409	0.0306	0.0528	0.0493	0.0449	0.0385	0.0427	0.0303	0.0584	0.0508	0.0353	0.0288	0.0279	0.0247	0.0412
Coverage Probability	0.755		0.383		0.79		0.618		0.626		0.901		0.273		0.817		0.81	

According to Table 4.1, under the standard normal distribution all GMM estimators do no perform very well under N=50. For all autoregressive parameters, when T=5, the standard errors of the parameter have increased. This situation does not change until the parameter value being move to high values. Although having increased standard error, BB GMM estimator has better performance in comparison with other estimators. In general, as T increases, standard error decreases, hence coverage probabilities have become less likely. The interesting point is that when T=15, the standard errors of second step of all estimators have increased, thus coverage probabilities has become more likely in comparison to for T=10. For instance, while for T=10 and N=50, the standard error of BB2 is obtained 0.0136 for ρ=0.1. Besides, as T moves from 10 to 15, the standard error of BB2 reaches to 0.0528. This result is surprising for central limit theorem.

After examining the table, it is concluded that BB GMM estimator has better performance for capturing the actual autoregressive parameter value.

Table 4.2 Estimated ratios of standard errors of three GMM estimators with each other for individual size N=50, with *i.i.d*. SN(0,1) residuals.

	ρ=0.1	**ρ =0.5**	**ρ =0.9**
T=5			
Relative efficiency (ABO/AB)	**1.01878595**	**0.691779020**	**0.54913075**
Relative efficiency (BB/AB)	**0.51786305**	**0.407209182**	**0.20532194**
Relative eficiency (BB/ABO)	**0.56459272**	**0.710141978**	**0.77232868**
T=10			
Relative efficiency (ABO/AB)	**1.104820939**	**0.890624299**	**0.3480136190**
Relative efficiency (BB/AB)	**0.274931469**	**0.229236030**	**0.1193238884**
Relative efficiency (BB/ABO)	**0.278682500**	**0.300187479**	**0.4104300409**
T=15			
Relative efficiency (ABO/AB)	**1.070960833**	**1.021303169**	**0.783739803**
Relative efficiency (BB/AB)	**1.725342120**	**1.820653736**	**1.548718592**
Relative efficiency (BB/ABO)	**1.756683462**	**1.988468346**	**2.472739833**

In terms of relative efficiency, for the small individual size, N=50, and small time period, T=5 and 10, BB GMM estimator has more efficiency compared to AB and ABO. However, when T=15, AB GMM estimator has become more efficient for all autoregressive parameters.

Table 4.3 Parameter values, standard errors and coverage probabilities with *i.i.d*. distributed SN(0,1) with *N*=100. (**AB1-AB2:** Arellano-Bond GMM estimator for first and second step, **ABO1-ABO2:** Arellano-Bover GMM estimator for first and second step, **BB1-BB2**: Blundell-Bover GMM estimator for first and second step)

	ρ=0.1						**ρ=0.5**						**ρ=0.9**					
	AB1	**AB2**	**ABO1**	**ABO2**	**BB1**	**BB2**	**AB1**	**AB2**	**ABO1**	**ABO2**	**BB1**	**BB2**	**AB1**	**AB2**	**ABO1**	**ABO2**	**BB1**	**BB2**
T=5																		
$\hat{\rho}$	0.0768	0.0773	0.116	0.1089	0.1057	0.1031	0.4683	0.4736	0.5065	0.5081	0.4981	0.5035	0.7923	0.7859	0.8889	0.8929	0.8739	0.8792
$SE(\widehat{\rho})$	0.0973	0.0896	0.0741	0.0866	0.0585	0.0646	0.1268	0.1181	0.0802	0.0858	0.0667	0.0733	0.2203	0.2047	0.0832	0.0816	0.0743	0.0729
Coverage Probability	0.891		0.911		0.888		0.89		0.9		0.888		0.856		0.929		0.843	
T=10																		
$\hat{\rho}$	0.0855	0.0854	0.1445	0.1234	0.1216	0.1135	0.4675	0.4669	0.5239	0.5161	0.5058	0.5033	0.8289	0.8246	0.8996	0.9013	0.8861	0.8895
$SE(\widehat{\rho})$	0.0473	0.0338	0.0397	0.0335	0.0303	0.0245	0.0535	0.0386	0.0395	0.0333	0.0316	0.0254	0.0675	0.0494	0.0339	0.0248	0.0297	0.0218
Coverage Probability	0.783		0.7		0.681		0.707		0.742		0.716		0.604		0.802		0.713	
T=15																		
$\hat{\rho}$	0.0865	0.0869	0.1701	0.1668	0.1353	0.1347	0.4757	0.4756	0.5337	0.5322	0.5125	0.5124	0.8419	0.841	0.9012	0.9015	0.8875	0.8878
$SE(\widehat{\rho})$	0.0341	0.0084	0.0295	0.0082	0.0223	0.0047	0.0361	0.009	0.0285	0.0081	0.0223	0.0048	0.0386	0.01	0.0218	0.0057	0.0188	0.0037
Coverage Probability	0.344		0.107		0.141		0.305		0.216		0.206		0.126		0.359		0.233	

Table 4.3 presents that for standard normal distributed residuals, BB estimator has captured the actual autoregressive parameter, ρ=0.1 and 0.5, with nominal size 5% for even the sample size T=5. While the autoregressive parameter value increases, the standard error of the estimator increases as expected. This table shows that as the time period increases, all estimators' coverage probabilities has become low values. When ρ is highly autocorrelated, ABO estimator is better than other estimators.

When ρ=0.1 and 0.5, BB estimator is better than AB and ABO estimators. As ρ parameter increases to 0.9, ABO estimator has better performance. When T=15, the AB estimator has made the better estimation for ρ=0.1, BB estimator is better for ρ=0.5 and ABO estimator is closer than others for ρ=0.9.

When T value increases, the standard errors have decreased for each estimator. Hence, the coverage probabilities have fallen considerably. That is, when T is considerably small for linear dynamic panel data models, parameter estimates appear to be better.

Table 4.4 Estimated ratios of standard errors of three GMM estimators with each other for individual size N=100, with *i.i.d*. SN(0,1) residuals.

	ρ=0.1	**ρ =0.5**	**ρ =0.9**
		T=5	
Relative efficiency (ABO/AB)	**0.990848127**	**0.618318356**	**0.294242021**
Relative efficiency BB/AB)	**0.537531445**	**0.420588063**	**0.194094915**
Relative efficiency (BB/ABO)	**0.575646083**	**0.762142537**	**0.857772413**
		T=10	
Relative efficiency (ABO/AB)	**1.001854277**	**0.776880553**	**0.3111865027**
Relative efficiency BB/AB)	**0.534688614**	**0.447387970**	**0.2264757258**
Relative efficiency (BB/ABO)	**0.543750575**	**0.599105902**	**0.7982158276**
		T=15	
Relative efficiency (ABO/AB)	**1.118819132**	**0.985899702**	**0.4656295800**
Relative efficiency BB/AB)	**0.403221681**	**0.359188478**	**0.1932641055**
Relative efficiency (BB/ABO)	**0.423452014**	**0.440474326**	**0.5419420613**

The results are given in Table 4.4. In terms of relative efficiency, BB estimators generally performs better. In Table 4.4., when both T and ρ have increased under ε_{it}~SN(0,1) distribution, BB generally gives more successful results. If look inside the table in detail, in case of T=15 and ρ=0.9, the ratio of (BB/AB) equals to 0.19326. This means that BB estimator has less standard error in comparison with AB estimator, hence the BB estimator is more efficient that AB and ABO estimators.

Table 4.5 Parameter values, standard errors and coverage probabilities with *i.i.d*. distributed SN(0,1) with *N*=500.

	ρ=0.1						**P=0.5**						**ρ=0.9**					
	AB1	**AB2**	**ABO1**	**ABO2**	**BB1**	**BB2**	**AB1**	**AB2**	**ABO1**	**ABO2**	**BB1**	**BB2**	**AB1**	**AB2**	**ABO1**	**ABO2**	**BB1**	**BB2**
T=5																		
$\widehat{\rho}$	0.0998	0.0999	0.1038	0.1038	0.1032	0.1032	0.4947	0.4958	0.4998	0.5005	0.4991	0.5001	0.8798	0.8804	0.8983	0.8994	0.8945	0.8967
$SE(\widehat{\rho})$	0.0438	0.0431	0.0333	0.0415	0.0265	0.0318	0.0569	0.0561	0.036	0.0415	0.0303	0.0373	0.1	0.0989	0.0367	0.0368	0.0341	0.0359
Coverage Probability	0.948		0.931		0.939		0.944		0.931		0.936		0.943		0.95		0.931	
T=10																		
$\widehat{\rho}$	0.0962	0.0964	0.11	0.1019	0.1045	0.1005	0.4933	0.4934	0.5046	0.5021	0.5009	0.5004	0.8833	0.8836	0.9004	0.9011	0.8971	0.8993
$SE(\widehat{\rho})$	0.0215	0.0201	0.0182	0.0199	0.0138	0.0159	0.0243	0.0229	0.0182	0.0203	0.0146	0.0167	0.032	0.0303	0.0157	0.0152	0.0141	0.0147
Coverage Probability	0.922		0.909		0.915		0.915		0.912		0.914		0.893		0.924		0.908	
T=15																		
$\widehat{\rho}$	0.0976	0.0975	0.1144	0.104	0.1072	0.1019	0.4943	0.4947	0.5081	0.5033	0.5029	0.5014	0.8877	0.8879	0.9005	0.9006	0.8974	0.8991
$SE(\widehat{\rho})$	0.0155	0.0134	0.0136	0.0133	0.0102	0.011	0.0165	0.0144	0.0132	0.0134	0.0103	0.0109	0.0182	0.016	0.0105	0.0094	0.009	0.0089
Coverage Probability	0.883		0.891		0.878		0.876		0.88		0.881		0.819		0.879		0.861	

When the residuals have followed the SN(0,1) distribution, all three estimators give nearly the same satisfactory results especially in the situation of the large sample size N=500. In case of T=5, all GMM estimators have made very close estimation for each ρ values and it seems that coverage probabilities are very close to 95% confidence level. When T=5 and 10, the AB estimator for ρ=0.1, AB and ABO estimators for ρ=0.5 and ABO for ρ=0.9 has the closest coverage probability to 0.95. When T value increases, the standard error has decreased for each estimator. Therefore, the coverage probabilities are little far from 95% confidence level. That is, when the smaller T values, the better parameter estimates. In fact, when N goes to infinity, T value being move to high values does not change anymore.

As shown in Table 4.6, the efficiency alignment is BB, ABO and AB GMM estimators for all ρ values under the SN(0,1) residuals and N=500. For example, when T=5 and ρ=0.9, relative efficiency of (BB/AB) is 0.1516 and relative efficiency of (BB/ABO) is 0.9569. That is, BB estimator is more efficient than ABO and AB estimator.

When T=10 and 15 for all ρ values, it seems that BB estimator is the most efficient estimator. For example; in case of T=10 and ρ=0.9, relative efficiency of (BB/AB) equals to 0.2458, whereas relative efficiency of (BB/ABO) is 0.9364. It means that BB estimator is more efficient than AB and ABO estimators.

Table 4.6 Estimated ratios of standard errors of three GMM estimators with each other for individual size *N*=500, with *i.i.d*. SN(0,1) residuals

	ρ=0.1	**ρ =0.5**	**ρ =0.9**
		T=5	
Relative efficiency (ABO/AB)	**0.9350565**	**0.5646465990**	**0.1612897381**
Relative efficiency (BB/AB)	**0.5479773**	**0.4491786783**	**0.1516965329**
Relative efficiency (BB/ABO)	**0.5933215**	**0.8130770118**	**0.9569864021**
		T=10	
Relative efficiency (ABO/AB)	**0.9801286674**	**0.7904969978**	**0.2653249960**
Relative efficiency (BB/AB)	**0.6238732650**	**0.5351313448**	**0.2458166941**
Relative efficiency (BB/ABO)	**0.6383048106**	**0.6820766818**	**0.9364443760**
		T=15	
Relative efficiency (ABO/AB)	**0.9902813377**	**0.8656730835**	**0.3550718006**
Relative efficiency (BB/AB)	**0.6714768210**	**0.5716106262**	**0.3182370428**
Relative efficiency (BB/ABO)	**0.6792069133**	**0.6637664997**	**0.9046768563**

Table 4.7 Parameter values, standard errors and coverage probabilities with *i.i.d.* distributed t(5) with *N*=50.

	ρ=0.1						**ρ=0.5**						**ρ=0.9**					
	AB1	**AB2**	**ABO1**	**ABO2**	**BB1**	**BB2**	**AB1**	**AB2**	**ABO1**	**ABO2**	**BB1**	**BB2**	**AB1**	**AB2**	**ABO1**	**ABO2**	**BB1**	**BB2**
T=5																		
$\hat{\rho}$	0.0733	0.0766	0.1287	0.1207	0.1112	0.1122	0.4319	0.4447	0.5001	0.5052	0.4854	0.4931	0.6237	0.6065	0.8938	0.8972	0.8724	0.8823
$SE(\widehat{\rho})$	0.1008	0.1027	0.081	0.0977	0.0622	0.072	0.1328	0.1363	0.8639	0.9705	0.6972	0.7851	0.2691	0.2724	0.0739	0.0771	0.068	0.0682
Coverage Probability	0.853		0.851		0.805		0.831		0.823		0.783		0.706		0.846		0.77	
T=10																		
$\hat{\rho}$	0.0754	0.0756	0.1547	0.1486	0.1229	0.1219	0.4534	0.4547	0.5275	0.5256	0.5024	0.5027	0.7528	0.7499	0.8963	0.897	0.8776	0.8785
$SE(\widehat{\rho})$	0.0501	0.0227	0.0432	0.0224	0.0326	0.0111	0.0548	0.0254	0.0421	0.0212	0.0332	0.0112	0.0759	0.0372	0.0305	0.0141	0.0279	0.0086
Coverage Probability	0.488		0.379		0.267		0.435		0.426		0.291		0.24		0.491		0.304	
T=15																		
$\hat{\rho}$	0.0722	0.0666	0.1778	0.1683	0.1328	0.1231	0.4572	0.4455	0.5361	0.5347	0.5062	0.4946	0.7953	0.7879	0.9026	0.904	0.8832	0.8775
$SE(\widehat{\rho})$	0.0362	0.0343	0.0319	0.0336	0.0239	0.0443	0.0373	0.0368	0.0303	0.0342	0.0234	0.0479	0.0417	0.033	0.0202	0.0202	0.018	0.0333
Coverage Probability	0.682		0.458		0.82		0.604		0..6510		0.888		0.231		0.734		0.834	

According to Table 4.7, all GMM estimators do no perform very well under N=50 under the t(5) distribution same as standard normal distribution. In general, it is concluded that AB first step and second step estimators have made underestimate for autoregressive values. As ρ parameter values moves to unit root, the difference between actual value and estimated value have dramatically increased.

For all autoregressive parameters, when T=5, the standard errors of the parameter have increased. This situation does not change until the parameter value being move to high values. Although having increased standard error, BB GMM estimator has better performance in comparison with other estimators. In general, as T increases, standard error decreases, hence coverage probabilities have become less likely. The interesting point is that when T=15, the standard errors of second step of all estimators have increased, thus coverage probabilities has become more likely in comparison to for T=10. For instance, while for T=10 and N=50, the standard error of BB2 is obtained 0.0111 for ρ=0.1. Besides, as T moves from 10 to 15, the standard error of BB2 reaches to 0.0443. This result is same as standard normal distribution case.

After examining the table, it is concluded that BB GMM estimator has better performance for capturing the actual autoregressive parameter value. In general, the BB estimator has become overestimate statues for the lower autoregressive values, whereas the BB estimator has become underestimate statues for the higher autoregressive values. As a result, there is no significant statistical difference for all estimators under two distributions. The performance of three estimators is free of distribution of residuals, or at least there is no big difference in performance of GMM estimation underlying residual distributions.

Table 4.8 Estimated ratios of standard errors of three GMM estimators with each other for individual size *N*=50, with *i.i.d.* t(5) residuals.

	ρ=0.1	ρ =0.5	ρ =0.9
		T=5	
Relative efficiency (ABO/AB)	**1.05922636**	**0.8031565**	**0.202853205**
Relative efficiency (BB/AB)	**0.53399841**	**0.3981728**	**0.121369033**
Relative efficiency (BB/ABO)	**0.58491640**	**0.7241677**	**0.846442887**
		T=10	
Relative efficiency (ABO/AB)	**1.113860370**	**0.828687549**	**0.213911748**
Relative efficiency (BB/AB)	**0.273056306**	**0.222086963**	**0.076397224**
Relative efficiency (BB/ABO)	**0.275651922**	**0.313997265**	**0.406487388**
		T=15	
Relative efficiency (ABO/AB)	**1.091805591**	**0.986151870**	**0.474271503**
Relative efficiency (BB/AB)	**1.797455895**	**1.831334697**	**1.207715352**
Relative efficiency (BB/ABO)	**1.856195959**	**2.142815914**	**3.030710306**

When Table 4.8 is examined, and compared with results of SN(0,1) under the same conditions in Table 4.2., the all three estimators of autoregressive parameters are slightly bigger than the ones which are obtained in Table 4.2. This results, however, do not change the decision of the BB estimator has more efficient and less standard error in comparison with the others.

Table 4.9 Parameter values, standard errors and coverage probabilities with *i.i.d.* distributed t(5) with *N*=100.

	ρ=0.1						**ρ =0.5**						**ρ=0.9**					
	AB1	**AB2**	**ABO1**	**ABO2**	**BB1**	**BB2**	**AB1**	**AB2**	**ABO1**	**ABO2**	**BB1**	**BB2**	**AB1**	**AB2**	**ABO1**	**ABO2**	**BB1**	**BB2**
T=5																		
$\hat{\rho}$	0.0905	0.0919	0.1143	0.113	0.1071	0.1081	0.475	0.4775	0.5041	0.5063	0.4975	0.5017	0.7478	0.7459	0.8954	0.8989	0.8834	0.8899
$SE(\widehat{\rho})$	0.0714	0.0798	0.0575	0.0767	0.0445	0.0585	0.0954	0.1081	0.0604	0.0755	0.0501	0.0657	0.2021	0.2307	0.0508	0.0579	0.0482	0.0551
Coverage Probability	0.895		0.874		0.86		0.89		0.898		0.871		0.824		0.907		0.86	
T=10																		
$\hat{\rho}$	0.0855	0.0855	0.1264	0.1155	0.1097	0.1062	0.4735	0.4763	0.5141	0.5119	0.5005	0.5021	0.8129	0.8131	0.8986	0.8995	0.888	0.8916
$SE(\widehat{\rho})$	0.0358	0.0292	0.0312	0.0291	0.0235	0.0215	0.0399	0.0329	0.0303	0.0279	0.0241	0.0221	0.0579	0.0498	0.0222	0.0182	0.0205	0.0167
Coverage Probability	0.742		0.685		0.676		0.724		0.702		0.663		0.553		0.744		0.674	
T=15																		
$\hat{\rho}$	0.0867	0.0868	0.1439	0.1418	0.1192	0.1188	0.4778	0.4777	0.5199	0.5191	0.5039	0.5038	0.8388	0.8389	0.8989	0.8992	0.8886	0.8888
$SE(\widehat{\rho})$	0.0259	0.0068	0.0231	0.0068	0.0172	0.004	0.0271	0.0073	0.0221	0.0065	0.0171	0.004	0.0319	0.0093	0.015	0.0041	0.0135	0.0028
Coverage Probability	0.28		0.163		0.168		0.263		0.243		0.197		0.125		0.3		0.183	

In the Table 4.9, the performances of GMM estimators are shown when the residuals are distributed t with degrees of freedom 5 and N=100. BB estimator is better than AB and ABO estimators where T=5 and 10, when ρ=0.1 and 0.5. ABO estimator is better than other estimators when ρ=0.9. It seems that coverage probability is very close to 95% confidence level. When T=5 and 10, the AB estimator has the closest coverage probability to 0.95 for ρ=0.1, AB estimator and ABO estimator for ρ=0.5 and ABO estimator for ρ=0.9.

When T=15, the BB estimator has made the best estimation ρ=0.1 and 0.5, while ABO is the best estimator for ρ=0.9.

As T increase, it seems that the standard error of estimators decrease considerably. Therefore, when T=15, the coverage probabilities of estimators are quite far from 95% confidence level. In contrast to in case of S(0,1) distributed residuals, as T increase, parameter estimates and coverage probabilities appear to be better.

Table 4.10 Estimated ratios of standard errors of three GMM estimators with each other for individual size N=100, with *i.i.d.* t(5) residuals.

	ρ=0.1	**ρ =0.5**	**ρ =0.9**
		T=5	
Relative efficiency (ABO/AB)	0.998075108	0.665204328	0.112319076
Relative efficiency (BB/AB)	0.560790834	0.411464491	0.095667668
Relative efficiency (BB/ABO)	0.606599310	0.792596423	0.923622389
		T=10	
Relative efficiency (ABO/AB)	1.019891345	0.7631787615	0.1602382246
Relative efficiency (BB/AB)	0.554438987	0.4657422780	0.1338505025
Relative efficiency (BB/ABO)	0.559363026	0.6399709411	0.8540372329
		T=15	
Relative efficiency (ABO/AB)	1.163300483	0.9573753269	0.2746509559
Relative efficiency (BB/AB)	0.437779145	0.3874657894	0.1283926437
Relative efficiency (BB/ABO)	0.433777277	0.4849777340	0.5843618558

In the Table 4.10, *N* moves from 50 to 100. In the case of *T*=5, ABO estimator is more efficient than with a little difference for ρ=0.1. As ρ increases, ABO estimator becomes much more efficient than AB estimator. For each ρ value, it seems that BB estimator is more efficient than other estimators. For example; when ρ=0.9, relative efficiency (BB/AB)=0.0957, so BB estimator is much more efficient AB estimator.

When T=10 and 15, the efficiency of ABO estimator has deteriorated in comparison with AB estimator for ρ=0.1. For all ρ values, it seems that BB estimator is the most efficient estimator. For example; in case of *T*=10 and ρ=0.9, relative efficiency (BB/AB)=0.1338 is less than relative efficiency(BB /ABO)=0.8540. BB estimator is much more efficient than AB when compared with ABO estimator. As ρ value increases, it seems that BB estimator is much more efficient than other estimators.

Table 4.11 Parameter values, standard errors and coverage probabilities with *i.i.d.* distributed t(5) with *N*=500.

	ρ =0.1						**ρ =0.5**						**ρ =0.9**					
	AB1	**AB2**	**ABO1**	**ABO2**	**BB1**	**BB2**	**AB1**	**AB2**	**ABO1**	**ABO2**	**BB1**	**BB2**	**AB1**	**AB2**	**ABO1**	**ABO2**	**BB1**	**BB2**
T=5																		
$\hat{\rho}$	0.0966	0.097	0.1006	0.0988	0.0996	0.0998	0.4936	0.4952	0.5008	0.5008	0.4984	0.5002	0.875	0.8763	0.9006	0.9019	0.8982	0.9012
$SE(\widehat{\rho})$	0.0318	0.0397	0.0257	0.0382	0.0199	0.0305	0.0422	0.0529	0.0271	0.0374	0.0227	0.0345	0.0911	0.1147	0.0223	0.0278	0.0216	0.0276
Coverage Probability	0.93		0.913		0.92		0.924		0.928		0.936		0.927		0.933		0.927	
T=10																		
$\hat{\rho}$	0.0977	0.0976	0.1071	0.102	0.1033	0.1011	0.4949	0.4958	0.5032	0.5017	0.5004	0.5008	0.8794	0.8811	0.9001	0.9008	0.8976	0.8999
$SE(\widehat{\rho})$	0.0162	0.0188	0.0141	0.0185	0.0106	0.0151	0.0181	0.0212	0.0139	0.0182	0.011	0.0157	0.0277	0.0327	0.0101	0.0119	0.0095	0.0117
Coverage Probability	0.907		0.904		0.893		0.905		0.909		0.909		0.884		0.9		0.892	
T=15																		
$\hat{\rho}$	0.0974	0.0976	0.1089	0.1027	0.1039	0.1015	0.4956	0.4958	0.5039	0.5016	0.5007	0.5004	0.8852	0.8859	0.8996	0.9	0.8972	0.8989
$SE(\widehat{\rho})$	0.0118	0.0124	0.0106	0.0123	0.0079	0.0103	0.0124	0.0131	0.0102	0.0121	0.0079	0.0102	0.0153	0.0165	0.0069	0.0075	0.0063	0.0074
Coverage Probability	0.857		0.849		0.827		0.863		0.854		0.857		0.788		0.862		0.845	

The standard errors of parameter, coverage probabilities of all estimators are very close the results which obtained with SN(0,1). In case of T=5, all GMM estimators have made very close estimation for each ρ value and it seems that coverage probability is very close to 95% confidence level.

When T=5 and 10, the AB estimator has the closest coverage probability to 95% confidence level for ρ=0.1, BB and ABO estimators are better for ρ=0.5 and AB estimator is better for ρ=0.9.

When T increases, it seems that the standard error of estimators has decreased. As T increases, parameter estimates and coverage probabilities appear to be worse.

In the Table 4.12, the comparisons have been made by relative efficiency of the GMM estimators under t(5) distributed residuals and N=500. In the case of T=5, ABO estimator is more efficient than AB estimator for all ρ values. As ρ increase, ABO estimator has become much more efficient than AB estimator. BB estimator is more efficient than other estimators. For example; when ρ=0.9, relative efficiency of (BB/AB)=0.0668 is much more less than relative efficiency of (BB/ABO)=0.9818. Therefore, BB estimator is more efficient than ABO estimator, but BB estimator is much more efficient than AB estimator in comparison with ABO estimator.

When T=10 and 15, for all ρ values, it is concluded that BB estimator is the most efficient estimator. For example; in case of T=10 and ρ=0.9, relative efficiency of (BB/AB)=0.1336 is less than relative efficiency(BB/ABO)=0.9728. Hence, BB estimator is much more efficient than AB estimator in comparison to ABO estimator. It seems that BB estimator is the most efficient estimator for all ρ and T values.

Table 4.12.Estimated ratios of standard errors of three GMM estimators with each other for individual size *N*=500, with *i.i.d*. t(5) residuals.

	ρ=0.1	**ρ =0.5**	**ρ =0.9**
		T=5	
Relative efficiency (ABO/AB)	0.9462349585	0.5196351	0.0681175548
Relative efficiency (BB/AB)	0.5954088739	0.4366594	0.0668885861
Relative efficiency (BB/ABO	0.6437207653	0.8595173	0.9818306421
		T=10	
Relative efficiency (ABO/AB)	0.9827377657	0.7497152330	0.1376443
Relative efficiency (BB/AB)	0.6520081760	0.5567914666	0.1335922
Relative efficiency (BB/ABO	0.6664108141	0.7488649028	0.9728980
		T=15	
Relative efficiency (ABO/AB)	0.9917941910	0.8468937060	0.2143688
Relative efficiency (BB/AB)	0.6904657922	0.6023541678	0.2033951
Relative efficiency (BB/ABO	0.6980466375	0.7152526155	0.9521950

Figure 4.2 shows the bias of three GMM estimators with individual size N=50, 100, 200, 300, 400, 500 in case of *i.i.d*. SN(0,1) distribution of residuals when T=5, 10 and ρ =0.1, 0.5, 0.9. Figure 4.3 and Figure 4.4 present the MSE of three GMM estimators individual size N=50, 100, 200, 300, 400, 500 in case of *i.i.d*. SN(0,1) and t(5) distribution of residuals when T=5, 10 and ρ =0.1, 0.5, 0.9.

When the residuals follow SN(0,1) distribution, the bias of the estimators are examined for the case of T=5, 10 and ρ=0.1, 0.5, 0.9.

The bias of the three GMM estimators are decreasing when N increases for all ρ and T values. In the case of ρ=0.1,0.5, AB estimator is negatively biased while both ABO and BB estimators are positively biased for all T values. Moreover, it can be seen that BB estimator has the smallest bias for all N values.

When the case of ρ=0.9, AB estimator has the smallest bias and all estimators are negatively biased. It means that when the autoregressive coefficient converge to the unit root, the bias is becoming negative. For the case of T=10, the bias values are more when N is smaller.

While N value is increasing, lines getting closer to each other and the biases of estimators are approach to 0.

The results when residuals distributed t(5) did not shown graphically because of giving similar results with when the residuals follow SN(0,1) distribution.

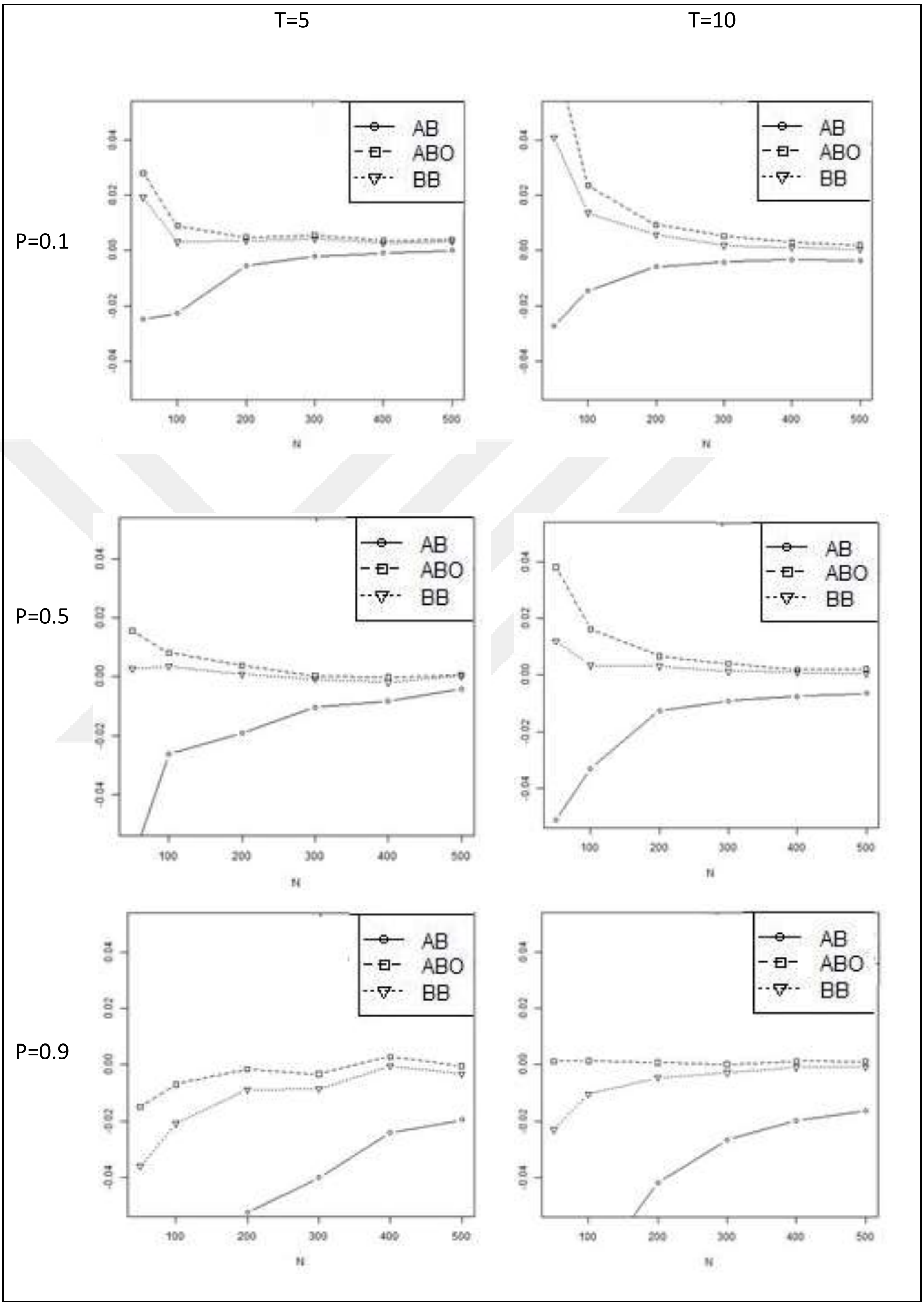

Figure 4.2 Bias of three GMM estimators for i.i.d SN(0,1) distributed residuals (T=5, 10)

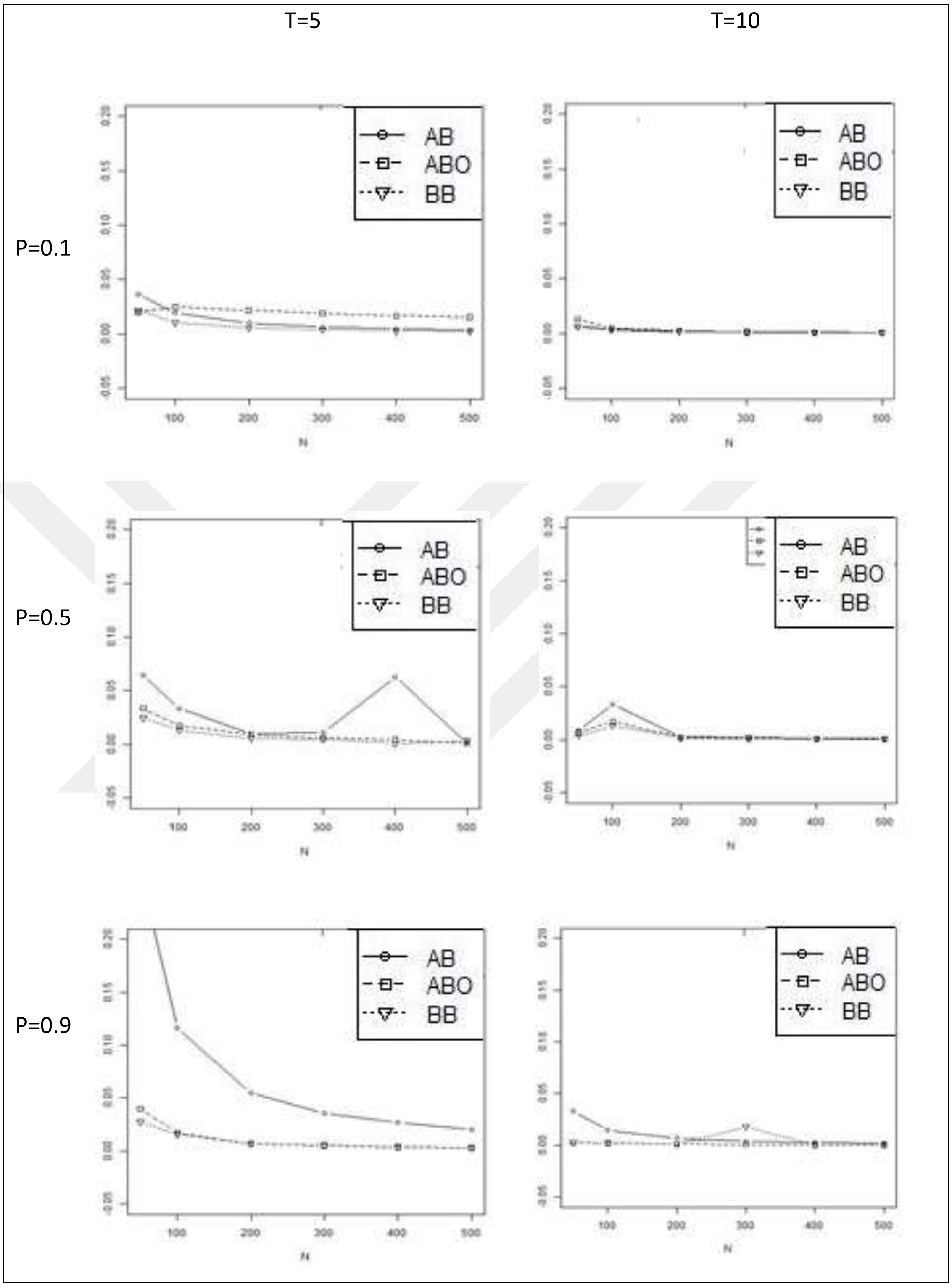

Figure 4.3 MSE of three GMM estimators for i.i.d SN(0,1) distributed residuals (T=5, 10)

In Figure 4.3, MSE values are examined for all GMM estimators when the residuals follow SN(0,1) distribution in the case of T=5,10 and ρ=0.1, 0.5, 0.9. It can be seen that, MSE is very close to 0 for all N values in the case of T=5 and ρ=0.1 MSE is more closer to 0 for all ρ values when T=10. When ρ converges to unit root, the MSE of AB estimator is bigger in the case of N is smaller and T=5. The MSEs of all estimators are smaller when T is increasing. The MSE of BB estimator is smaller than the MSE of other estimators for all ρ and T values. When N is increasing, the MSEs of all estimators are getting closer to 0.

In Figure 4.4, MSE values are examined for all GMM estimators when the residuals distributed t(5) in the case of T=5,10 and ρ=0.1, 0.5, 0.9.. It can be seen that, MSE is very close to 0 for all N values in the case of T=10 and ρ=0.1, 0.5. The MSEs of all GMM estimators are smaller when T is increasing. When ρ converges to unit root, the MSE of AB estimator is bigger in the case of N is smaller and T=5, 10. The MSE of BB estimator is smaller than the MSEs of other GMM estimators for all ρ and T values. When N is increasing, the MSEs of all estimators are getting closer to 0.

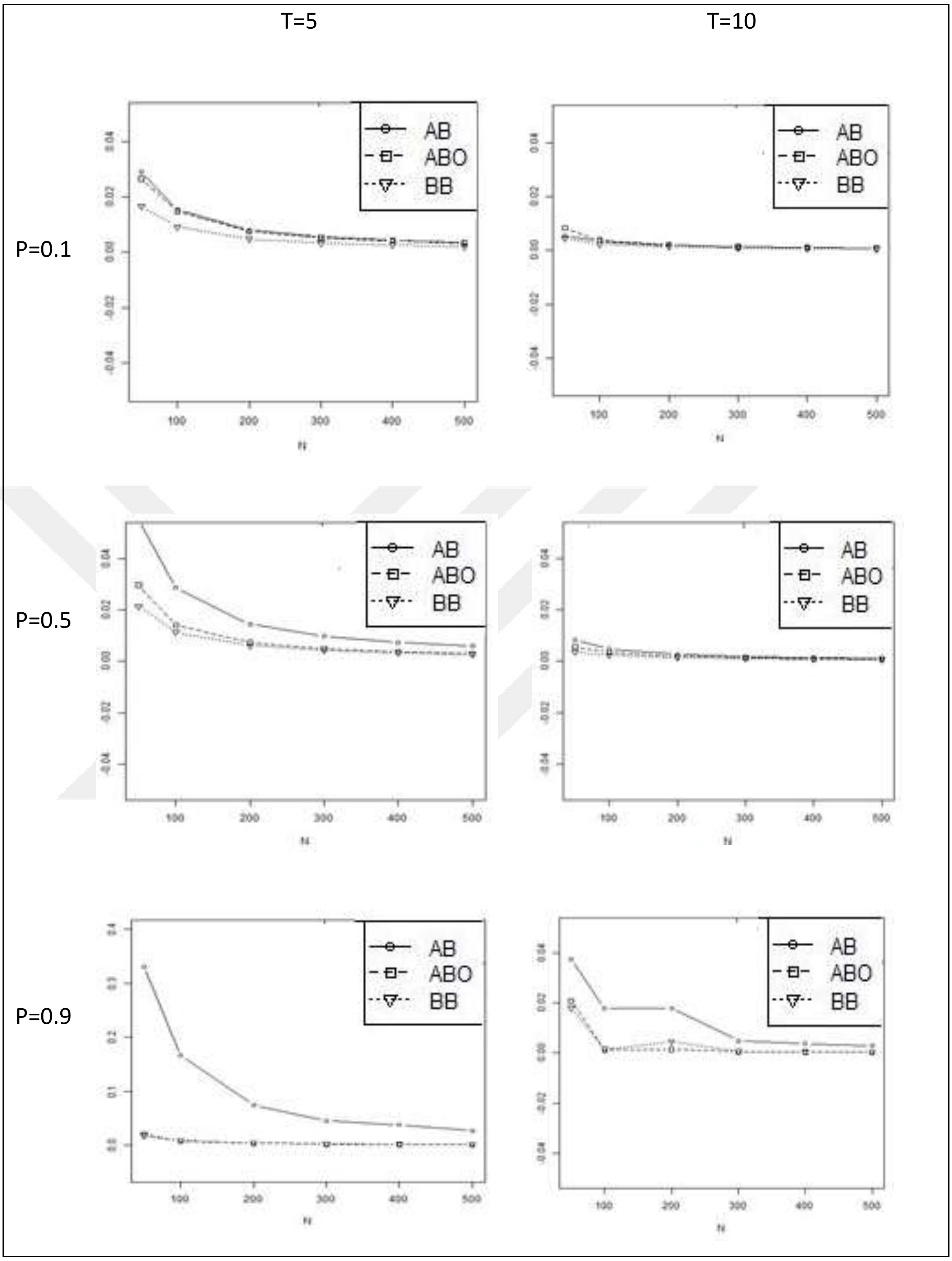

Figure 4.4 MSE of three GMM estimators for t(5) distributed residuals (T=5, 10)

CHAPTER FIVE
CONCLUSION

.In the thesis, we examined the comparison of the AB, ABO and BB GMM estimators. To illustrate the effect of different variance size on the GMM estimators two different distributions are considered for the comparison of three GMM estimators: Standard Normal and t(5) distributions. The comparison criterias are chosen bias, sampling distributions and standard error (SE) of ρ, coverage probabilities, relative efficiencies, mean squared error (MSE) underlying two different distributions.

For this simulation, the number of individual sample size and time periods and autoregressive coefficients are investigated for three levels, respectively: $N = 50, 100, 500$, $T = 5, 10, 15$ and $\rho = 0.1, 0.5, 0.9$. Nominal significance level is chosen 0.05.

The AB, ABO and BB GMM estimators are performed into two steps, therefore in the tables AB1 and AB2 abbreviations for first step and second step of AB, respectively. Likely, ABO1 and ABO2 are used for ABO estimator's step and also BB1 and BB2 represent the BB estimator's two steps.

GMM estimation method has many optimal properties in estimation: consistency, efficiency, asymptotically normal.

In general, it is concluded that first step and second step AB estimators have made underestimate for autoregressive values. As ρ parameter values moves to unit root, the difference between actual value and estimated value have dramatically increased.

As T increases, standard error decreases, therefore coverage probabilities has become less likely. The interesting point is that when N=50, the standard errors of second step of all estimators have increased in case of T=15, thus coverage probabilities has become more likely in comparison to for T=10.

In general, the BB estimator has made overestimate for the lower autoregressive values, whereas the BB estimator has made underestimate for the higher autoregressive values.

As a result, there is no significant statistical difference for all estimators under two distributions. The performance of three estimators is free of distribution of residuals, or at least there is no big difference in performance of GMM estimation.

The all GMM estimators of autoregressive parameters are slightly bigger than the ones which are obtained. These results do not change the decision of the BB estimator has more efficient and less standard error in comparison with the other GMM estimators.

It has examined the bias of the three GMM estimators under two distributions for the residuals. The bias of the three GMM estimators are decreasing when N increases for all ρ and T values. The results when residuals distributed t(5) are similar with when the residuals follow SN(0,1) distribution.

MSE values are examined for all GMM estimators when the residuals that distributed SN(0,1) snd t(5) in the case of T=5,10 and p=0.1, 0.5, 0.9. The MSE of BB estimator is smaller than the MSE of other estimators for all ρ and T values. When N is increasing, the MSEs of all estimators are getting closer to 0.

REFERENCES

Abonazel, Mohamed R. (2015). R-codes to calculate GMM estimations for dynamic panel data models, 1-6.

Ahn, S. C., & Schmidt, P. (1995). Efficient estimation of models for dynamic panel data. *Journal of econometrics*, *68*(1), 5-27.

Alfaro, R (2008). Estimation of a dynamic panel data: The case of corporat investment in Chile, 1-6.

Alonso-Borrego, C., & Arellano, M. (1999). Symmetrically normalized instrumental-variable estimation using panel data. *Journal of Business & Economic Statistics*, *17*(1), 36-49.

Alvarez, J., & Arellano, M. (2003). The Time Series and Cross-Section Asymptotics of Dynamic Panel Data Estimators. *Econometrica*, *71*(4), 1121-1159.

Anderson, T. W. (1959). On asymptotic distributions of estimates of parameters of stochastic difference equations. *The Annals of Mathematical Statistics*, 676-687.

Anderson, T. W., & Hsiao, C. (1981). Estimation of dynamic models with error components. *Journal of the American statistical Association*, *76*(375), 598-606.

Arellano, M. and Bond, S. (1991). Tests of specification for panel data: Monte Carlo evidence and an application to employment equations, *Review of Economic Studies*, *58*, 277-97.

Arellano, M., & Bover, O. (1995). Another look at the instrumental variable estimation of error-components models. *Journal of econometrics*, *68*(1), 29-51.

Arellano, M., & Honoré, B. (2001). Panel data models: some recent developments. *Handbook of Econometrics*, *5*, 3229-3296.

.

Baltagi, B. H., & Li, D. (2002). Series estimation of partially linear panel data models with fixed effects. *Annals of Economics and Finance*, *3*(1), 103-116.

Baltagi, B. H. (2005). Econometric Analysis of Panel Data 3rd Edition England JW &Sons.

Binder, M., Hsiao, C., & Pesaran, M. H. (2005). Estimation and inference in short panel vector autoregressions with unit roots and cointegration. *Econometric Theory*, *21*(04), 795-837.

Blundell, R., & Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models, *Journal of Econometrics, 87* (1), 115-143.

Blundell, R., Bond, S., & Windmeijer, F. (2001). Estimation in dynamic panel data models: improving on the performance of the standard GMM estimator. *Nonstationary Panels, Panel Cointegration, and Dynamic Panels*, *15*, 53-91.

Bond, S., Bowsher, C., & Windmeijer, F. (2001). Criterion-based inference for GMM in autoregressive panel data models. *Economics Letters*, *73*(3), 379-388.

Bond, S. R. (2002). Dynamic panel data models: a guide to micro data methods and practice. *Portuguese Economic Journal*, *1* (2), 141-162.

Bun, M. J., & Windmeijer, F. (2010). The weak instrument problem of the system GMM estimator in dynamic panel data models. *The Econometrics Journal*, *13*(1), 95-126.

Calzolari, G., & Magazzini, L. (2013). A powerful test of mean stationarity in dynamic models for panel data: Monte Carlo evidence *Working Paper Series Department of Economics University of Verona*. 1-4.

Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American statistical association*, *74*(366a), 427-431.

Dickey, D. A., & Fuller, W. A. (1981). Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica: Journal of the Econometric Society*, 1057-1072.

Flannery, M. J., & Hankins, K. W. (2013). Estimating dynamic panel models in corporate finance. *Journal of Corporate Finance*, *19*, 1-19.

Ghysels, E., & Hall, A. (2002). Editors' introduction to twentieth anniversary commemorative issue of the Journal of Business and Economic Statistics. *Journal of Business & Economic Statistics*, *20*(1), 1-4.

Hall, R. E. (2005). Employment fluctuations with equilibrium wage stickiness. *The American Economic Review*, *95*(1), 50-65.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica: Journal of the Econometric Society*, 1029-1054.

Hansen, G. (2001). A bias-corrected least squares estimator of dynamic panel models. *AStA Advances in Statistical Analysis*, *2*(85).

Hayakawa, K. (2007). Small sample bias properties of the system GMM estimator in dynamic panel data models. *economics Letters*, *95*(1), 32-38.

Holtz-Eakin, D., Newey, W., & Rosen, H. S. (1988). Estimating vector autoregressions with panel data. *Econometrica: Journal of the Econometric Society*, 1371-1395.

Kiviet, J. F. (1995). On bias, inconsistency, and efficiency of various estimators in dynamic panel data models, *Journal of Econometrics* 68, 53-78.

Mileva, E. (2007). Using Arellano-Bond dynamic panel GMM estimators in Stata. *Economics Department, Fordham University*, 1-10.

Ogaki, M. (1993). 17 Generalized method of moments: Econometric applications. *Handbook of statistics*, *11*, 455-488.

Pakes, A., & Griliches, Z. (1984). Estimating distributed lags in short panels with an application to the specification of depreciation patterns and capital stock constructs. *The Review of Economic Studies*, *51*(2), 243-262.

Phillips, P. C., & Durlauf, S. N. (1986). Multiple time series regression with integrated processes. *The Review of Economic Studies*, *53*(4), 473-495.

Roodman, D. (2009). A note on the theme of too many instruments. *Oxford Bulletin of Economics and statistics*, *71*(1), 135-158.

Santos, L. A., & Barrios, E. B. (2011). Small sample estimation in dynamic panel data models: A simulation study. *Open Journal of Statistics*, *1*(2), 58.

Schwarz, C., & Kripfganz, S. (2015). *Estimation of linear dynamic panel data models with time-invariant regressors* (No. 1838).

Sun, X., & Ashley, R. A. (2014). A Transformed System GMM Estimator for Dynamic Panel Data Models. *Available at SSRN 2398717*.

Windmeijer, F. (2005). A finite sample correction for the variance of linear efficient two-step GMM estimators. *Journal of econometrics*, *126*(1), 25-51.

Wooldridge, J. M. (2001). Applications of generalized method of moments estimation. *The Journal of Economic Perspectives*, *15* (4), 87-100.

Youssef, A. H., & Abonazel, M. R. (2015). Alternative GMM estimators for first-order autoregressive panel model: an improving efficiency approach. *Communications in Statistics-Simulation and Computation*, (just-accepted), 00-00.