Abstract

KIM, HYUNJUNG. Unit Root Tests in Panel Data: Weighted Symmetric Estimation and Maximum Likelihood Estimation. (Under the direction of David A. Dickey.)

There has been much interest in testing nonstationarity of panel data in the econometric literature. In the last decade, several tests based on the ordinary least squares and Lagrange multiplier method have been developed. In contrast to a unit root test in the univariate case, test statistics in panel data have Gaussian limiting distributions.

This dissertation considers weighted symmetric estimation and maximum likelihood estimation in the autoregressive model with individual effects. The asymptotic distributions have been derived as the number of individuals and time periods become large. The power study from Monte Carlo experiments shows that the proposed test statistics perform substantially better than those in previous studies even for small samples. As an example, we consider the real Gross Domestic Product per Capita for 12 countries.

Unit Root Tests in Panel Data: Weighted Symmetric Estimation and Maximum Likelihood Estimation

by

Hyunjung Kim

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Statistics

in the

GRADUATE SCHOOL
at
NC STATE UNIVERSITY
2001

Professor David A. Dickey
Chair of Advisory Committee

Professor Marcia L. Gumpertz

Professor Bibhuti B. Bhattacharyya

Professor Sastry G. Pantula

To my eternal beloved

and

my parents

Biography

Hyunjung Kim was born in Seoul, Korea, to parents Wonji Kim and Jaewoon Park on November 9, 1972. She earned her Bachelor's degree in Statistics in February 1995 and her Master's degree in Statistics in February 1997 at Ewha Womans University. She worked for Korea Economic Research Institute as a research assistant. Since August 1997, she has studied for the Ph.D degree in the Department of Statistics at North Carolina State University. While at NCSU, she has participated in several projects as a research assistant. Under the direction of Dr. David A. Dickey, she has worked on unit root tests in panel data.

Acknowledgements

I would like to express my gratitude to Dr. David A. Dickey. He has been my mentor and good example to me as a researcher. Without his kindly and sincere guidance, I could not complete this dissertation. I am also grateful to my committee members, Dr. Battacharyya, Dr. Gumpertz, and Dr. Pantula for their comments on my dissertation. Dr. Pantula has given me encouragement and valuable advice through my academic years. It has been a great pleasure for me to work with Dr. Brownie. Her supervision as well as financial support helped me to finish my graduate program without any difficulty. I will not forget the faculties and staff in our department for their enthusiastic teachings and kindness.

I also want to show my appreciation to the faculties in the department of statistics at Ewha Womans University. Especially, Dr. Dongwan Shin had provided me invaluable guidance during my master's program and led me to keep studying for Ph. D.

Recently, I realized that I have known many precious friends. Even though I do not mention their names, they are treasures to me that I have obtained through my life and will be remained in my deepest heart. I really appreciate all the saints who have spent time with me for last 4 years. Their love, fellowship and prayer in the Lord enliven me in every situation.

I can not say enough thanks and pay their love back to my parents. Their ceaseless love and support are the most important resources for me to finish my Ph.D. program.

Last but not least, this dissertation should be dedicated to my eternal company who became my everything.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Statement of the problem

Time series across states, industries, or randomly sampled individuals can be easily found in the macroeconomic and microeconomic literature. These kinds of data are called panel data, cross-section data, or longitudinal data. Among econometricians, there have been many studies about panel data with the autoregressive covariance structure over time. If it is assumed that there is no correlation across individuals, an autoregressive panel data model of order 1 with individual effects μ_i can be written as

$$y_{it} = \rho y_{it-1} + (1-\rho)\mu_i + e_{it}, \quad i = 1, 2, \cdots, N; t = 1, 2, \cdots, T, \tag{1.1}$$

where e_{it} follows a white noise process with mean zero and variance σ^2. Here, y_{it} can be interpreted as an observation which is obtained from the ith individual at time t. The individual effect μ_i can be assumed to be either a fixed or random effect which is independent of the white noise process e_{it}.

If $|\rho| < 1$, y_{it} follows a stationary process which has stable mean and variance even for large t. On the other hand, if $\rho = 1$, the variance of y_{it} increases as t becomes large. For this reason, as in a univariate time series, that is, when $N = 1$, there has been much interest in detecting a unit root in panel data. For testing $H_0 : \rho = 1$,

many econometricians have developed test statistics and the corresponding limiting distributions. Unlike a univariate time series, we need to consider two dimensions of indices, that is, (N, T) for deriving a limiting distribution.

In this thesis, we will develop unit root test statistics based on weighted symmetric estimation and maximum likelihood estimation for model (1.1). For weighted symmetric estimation, we also discuss higher order processes. For maximum likelihood estimation, we will consider random individual effects as well as fixed individual effects. Concerning white noise, we will assume the Gaussian distribution with mean zero and variance σ^2. Also, we will derive the limiting distributions for the test statistics under $H_0 : \rho = 1 \quad$ as $N, T \to \infty$. Through Monte Carlo simulation study, we will compare the power of our test statistics with some of previous studies. As an example, we will consider the real Gross Domestic Product per Capita for 12 countries.

1.2 Literature Review

Since the 1960's, there has been much literature about panel data models [e.g. Balestra and Nerlove (1966), Mundlak (1978), T.W. Anderson (1978), T.W. Anderson and Hsiao (1982)]. Anderson and Hsiao (1982) present several estimation methods for various regression models with the autoregressive order 1 covariance structure. Specifically, they focus on maximum likelihood estimation for the following stationary autoregressive panel data model.

$$y_{it} = \rho y_{it-1} + \delta' z_i + \gamma' x_{it} + \alpha_i + e_{it}, \quad i = 1, \cdots, N; t = 1, \cdots, T, \tag{1.2}$$

where

$$|\rho| < 1, E(\alpha_i) = E(e_{it}) = 0,$$

$$Var(e_{it}) = \sigma^2, Var(\alpha_i) = \lambda\sigma^2 \text{ for } \lambda > 0,$$

α_i and e_{it} are independent of each other, z_i is a vector of time-invariant exogenous variables and x_{it} is a vector of time-varying exogenous variables. For maximum likelihood estimation, the underlying distributions are Gaussian. They assume that the initial value y_{i0} may be fixed or random. For both cases, they allow that y_{i0} may be related to α_i or independent of α_i. They present consistency results for $\hat{\rho}$ and $\hat{\delta}$ when $N \to \infty$ for fixed T or $T \to \infty$ for fixed N.

Due to the fact that there exist nonstationary time series in many economic data sets, Quah (1994) introduces a unit root test for panel data. He considers the simplest panel data model,

$$y_{it} = \rho y_{it-1} + e_{it}, \tag{1.3}$$

where y_{i0} is a given random variable with $(2+\delta)$th finite moment and the e_{it}'s are independent identically distributed with mean zero, variance σ^2, and have bounded $(4+\delta)$th moment for $\delta > 0$. He assumes the cross-section dimension N and the time dimension T to be the same order of magnitude, that is, $N = N(T) = O(T)$. For model (1.3), he considers the ordinary least squares estimator $\hat{\rho}_{ols}$ of ρ. That is,

$$\hat{\rho}_{ols} = \Big[\sum_{i=1}^{N(T)}\sum_{t=1}^{T} {y_{it-1}}^2\Big]^{-1}\Big[\sum_{i=1}^{N(T)}\sum_{t=1}^{T} y_{it}y_{it-1}\Big]. \tag{1.4}$$

Under $H_0 : \rho = 1$, he shows

$$\frac{\sqrt{N(T)}}{\sqrt{2}}T(\hat{\rho}_{ols} - 1 - 2\frac{\mu}{\sigma^2}T^{-3/2}) \overset{D}{\Longrightarrow} N(0,1) \tag{1.5}$$

as $T \to \infty$,
where

$$\mu = \lim_{T\to\infty} E\Big[y_{i0}\Big(\frac{\sum_{t=1}^{T} e_{it}}{\sqrt{T}}\Big)\Big].$$

His result tells us that the unit root test statistic in panel data under $H_0 : \rho = 1$ has a Gaussian limiting distribution unlike its behavior in a univariate time series.

Levin and Lin (1992) generalize model (1.3) to have individual-specific effects and a time trend in their model. The model with individual-specific effects can be written as

$$y_{it} = \eta_i + \rho y_{it-1} + e_{it}, \tag{1.6}$$

where the e_{it} are independent and identically distributed with mean 0 and variance σ^2. Also, $E|e_{it}|^{2+\delta} < \infty$, for $\delta > 0$.
They consider the ordinary least squares estimator of ρ and the corresponding t-statistic in model (1.6). These are defined as

$$\hat{\rho}_{ols} = \Big[\sum_{i=1}^{N}\sum_{t=1}^{T}(y_{it} - \bar{y}_i)^2\Big]^{-1}\Big[\sum_{i=1}^{N}\sum_{t=1}^{T}(y_{it} - \bar{y}_i)(y_{it-1} - \bar{y}_i)\Big], \tag{1.7}$$

$$t_\rho = \Big[\hat{\sigma}_{ols}^{-2}\sum_{i=1}^{N}\sum_{t=1}^{T}(y_{it} - \bar{y}_i)^2\Big]^{1/2}\Big[\hat{\rho}_{ols} - 1\Big], \tag{1.8}$$

where

$$\bar{y}_i = \frac{1}{T}\sum_{t=1}^{T} y_{it},$$

and

$$\hat{\sigma}_{ols}^2 = \frac{1}{NT - N - 1}\sum_{i=1}^{N}\sum_{t=1}^{T}\Big[y_{it} - \bar{y}_i - \hat{\rho}_{ols}(y_{it-1} - \bar{y}_i)\Big]^2.$$

Under $H_0 : \rho = 1$, $\quad t_\rho$ has the following limiting distribution, based on the central limit theorem for triangular arrays (Billingsley 1986).

$$\sqrt{1.25}\Big[t_\rho - \frac{\sqrt{N}\mu_{1T}}{\sqrt{\mu_{2T}}}\Big] \xRightarrow{D} N(0,1), \tag{1.9}$$

as N and $T \to \infty$,

where

$$\mu_{1T} = E\Big[\frac{1}{T\sigma^2}\sum_{t=1}^{T}(y_{it-1} - \bar{y}_i)e_{it}\Big] = -\frac{1}{2} - \frac{1}{2T},$$

and

$$\mu_{2T} = E\Big[\frac{1}{T^2\sigma^2}\sum_{t=1}^{T}(y_{it-1} - \bar{y}_i)^2\Big] = \frac{1}{6} + \frac{5}{6T^2}.$$

If $\frac{\sqrt{N}}{T} \to 0$, then

$$\sqrt{1.25}t_\rho + \sqrt{1.875N} \overset{D}{\Longrightarrow} N(0,1). \tag{1.10}$$

In their 1993 paper, they allow the degree of persistence to vary across individuals. In other words, their individual-specific effects model (1.6) is generalized to the following model.

$$\Delta y_{it} = \eta_i + \delta_i y_{it-1} + \epsilon_{it}. \tag{1.11}$$

Here, ϵ_{it} is distributed independently across individuals and is a stationary invertible ARMA process as in Said and Dickey (1984). Also, this disturbance has finite fourth moments and bounded variation at frequency zero, that is,

$$E(\epsilon_{it}^2) + 2\sum_{j=1}^{\infty} E(\epsilon_{it}\epsilon_{it-j}) < M, \text{for } M > 0.$$

In the model (1.11), testing for a unit root is equivalent to testing $\delta_i = 0$ for all i. To remove time-specific aggregate effects, they subtract the cross-section average $\frac{1}{N}\sum_{i=1}^{N} y_{it}$ from the original data at time t. After this procedure, they perform the augmented Dickey-Fuller regression (1981) for each individual i. In other words, they consider the following regression model for each i.

$$\Delta y_{it} = \eta_i + \delta_i y_{it-1} + \sum_{j=1}^{p_i} \theta_{ij} \Delta y_{it-j} + e_{it} \tag{1.12}$$

Let $\hat{u}_{it}$ be the residual from the regression of Δy_{it} on η_i and $\sum_{j=1}^{p_i} \theta_{ij} \Delta y_{it-j}$. Also let $\hat{v}_{it-1}$ be the residual from the regression of y_{it-1} on η_i and $\sum_{j=1}^{p_i} \theta_{ij} \Delta y_{it-j}$. By using these residuals, they consider the following regression model for each i.

$$\hat{u}_{it} = \delta_i \hat{v}_{it-1} + e_{it}. \tag{1.13}$$

Dividing $\hat{u}_{it}$ and $\hat{v}_{it-1}$ by the regression standard error in the model (1.13), they obtain standardized residuals $\tilde{u}_{it}$ and $\tilde{v}_{it-1}$. They compute a combined estimator $\hat{\delta}$ and the corresponding t-test statistic from these standardized residuals. The definitions are

$$\hat{\delta} = \frac{\sum_{i=1}^{N} \sum_{t=2+p_i}^{T} \tilde{v}_{it-1} \tilde{u}_{it}}{\sum_{i=1}^{N} \sum_{t=2+p_i}^{T} \tilde{v}_{it-1}^2}, \tag{1.14}$$

and

$$t_\delta = \frac{\hat{\delta}}{SE(\hat{\delta})}, \tag{1.15}$$

where

$$SE(\hat{\delta}) = \hat{\sigma}_e \Big[\sum_{i=1}^{N} \sum_{t=2+p_i}^{T} \tilde{v}_{it-1}^2 \Big]^{-1/2},$$

$$\hat{\sigma}_e = \Big[\frac{1}{N\tilde{T}} \sum_{i=1}^{N} \sum_{t=2+p_i}^{T} (\tilde{u}_{it} - \hat{\delta} \tilde{v}_{it-1})^2 \Big]^{1/2},$$

and

$$\tilde{T} = T - \frac{1}{N} \sum_{i=1}^{N} p_i - 1.$$

Since t_δ in (1.15) goes to $-\infty$ for the individual-specific effects model, they introduce the following adjusted t-test statistic.

$$t_\delta^* = \frac{t_\delta - N\tilde{T}\hat{S}\hat{\sigma}_e^{-2}SE(\hat{\delta})\mu_{1\tilde{T}}^*}{\sigma_{1\tilde{T}}^*}, \tag{1.16}$$

where the average standard deviation ratio $\hat{S} = \frac{1}{N}\sum_{i=1}^{N}\hat{s}_i$ and $\hat{s}_i$ is the ratio of the long-run standard deviation to the innovation standard deviation for each i. The mean and variance adjustment factors, $\mu_{1\tilde{T}}^*, \sigma_{1\tilde{T}}^*$ can be found in Levin and Lin (1993). These are presented in a table computed by Monte Carlo simulation for various (N, T).
If $\frac{N}{T} \to 0$ then,

$$t_\delta^* \quad \stackrel{D}{\Longrightarrow} \quad N(0,1). \tag{1.17}$$

Im, Pesaran, and Shin (1997) propose the LM-bar test and t-bar test which are based on the averages of the Lagrange multiplier statistics and t-statistics, respectively. They consider the models (1.6) and (1.11) in Levin and Lin (1992 and 1993). However, the alternative hypothesis in this paper is that the fraction of the individual processes that are stationary is non-zero, whereas the alternative hypothesis in Levin and Lin (1993) is that all the individual processes are stationary. When errors in regressions are serially uncorrelated, that is, under the model (1.6), the LM-bar statistic is

$$\overline{LM}_{NT} = \frac{1}{N}\sum_{i=1}^{N} LM_{iT} \tag{1.18}$$

where

$$LM_{iT} = \frac{T\Delta\mathbf{y}_i'\mathbf{P}_i\Delta\mathbf{y}_i}{\Delta\mathbf{y}_i'\mathbf{M}_T\Delta\mathbf{y}_i}, \tag{1.19}$$

$$\mathbf{1}_T = \Big[1, 1, \cdots, 1\Big]_{T\times 1}',$$

$$\mathbf{M}_T = \mathbf{I}_T - \mathbf{1}_T(\mathbf{1}_T'\mathbf{1}_T)^{-1}\mathbf{1}_T',$$

$$\mathbf{P}_i = \mathbf{M}_T\mathbf{y}_{i,-1}(\mathbf{y}_{i,-1}'\mathbf{M}_T\mathbf{y}_{i,-1})^{-1}\mathbf{y}_{i,-1}'\mathbf{M}_T,$$

$$\mathbf{y}_{i,-1} = \Big[y_{i0}, y_{i1}, \cdots, y_{iT-1}\Big]',$$

and

$$\mathbf{\Delta y}_i = \Big[\Delta y_{i1}, \Delta y_{i2}, \cdots, \Delta y_{iT}\Big].$$

Then their standardized LM-bar statistic is

$$\Gamma_{\overline{LM}} = \frac{\sqrt{N}\Big[\overline{LM}_{NT} - E(LM_{iT})\Big]}{\sqrt{Var(LM_{iT})}}, \tag{1.20}$$

where $E(LM_{iT})$ and $Var(LM_{iT})$ are the mean and the variance of LM_{iT} under $H_0 : \rho = 1$.
Their t-bar statistic is

$$\overline{t}_{NT} = \frac{1}{N}\sum_{i=1}^{N} t_{iT}, \tag{1.21}$$

where t_{iT} is the t-statistic which corresponds to the estimator of δ_i from the augmented Dickey-Fuller regression model $\Delta y_{it} = \eta_i + \delta_i y_{it-1} + e_{it}$ for each i. For testing $\delta_i = 0$ for every i, they present the standardized t-bar statistic which is

$$\Gamma_{\overline{t}} = \frac{\sqrt{N}\Big[\overline{t}_{NT} - E(t_T|\delta_i = 0)\Big]}{\sqrt{Var(t_T|\delta_i = 0)}}. \tag{1.22}$$

They also construct the standardized LM-bar and t-bar statistics under the model (1.11). Under the model assumption of (1.6), they prove that $\Gamma_{\overline{LM}}$ and $\Gamma_{\bar{t}}$ have standard Gaussian limiting distributions as $N \to \infty$ and T is fixed. Also, they show that the standardized LM-bar and t-bar statistics which are constructed under the model (1.11) have standard Gaussian limiting distributions as N and T become large with a restriction of $\frac{N}{T} \to k$, a finite positive constant.

Karlsson and Löthgren (2000) compare power of Levin and Lin (1992)'s t- tests (hereafter LL tests) in (1.9) and (1.10) with Im, Peseran, and Shin (1997)'s tests (hereafter IPS tests) in (1.20) and (1.22). In their Monte Carlo experiments, they allow the proportion of stationary series in the panel to vary. The results show that IPS tests perform better than LL tests for large T. Also, the power increases as

- N increases,
- T increases,
- and the proportion of stationary series in the panel increases.

Especially, T is a critical factor in determining power.

Harris and Tzavalis (1999) also consider the unit root test statistics based on the pooled ordinary least squares estimator. In their paper, they derive limiting distributions for the normalized least squares estimators when N becomes large and T is finite. Under the model (1.6) and $H_0 : \rho = 1$,

$$\sqrt{N}(\hat{\rho} - 1 - b) \stackrel{D}{\Longrightarrow} N(0, C), \tag{1.23}$$

as $N \to \infty$ with fixed T,
where $\hat{\rho}$ is an ordinary least squares estimator of ρ under the model (1.6),

$$b = -\frac{3}{T+1},$$

and

$$C = \frac{3(17T^2 - 20T + 17)}{5(T-1)(T+1)^3}.$$

Phillips and Moon (1999) provide regression limit theory for nonstationary panel vector autoregressive models. They provide theorems which describe the relationship between sequential limits, where $T \to \infty$ is followed by $N \to \infty$ and joint limits, where both T and N become large simultaneously. Also, they develop the joint probability limit and joint central limit theorems for deriving a limiting distribution as $N, T \to \infty$. In their model, each subject has a vector of observations at each time and the goal is to test these for cointegration. For their models, they consider no cointegration, a heterogeneous cointegration relation, a homogeneous cointegration relation, and a near-homogeneous relation. They show that for each model, the pooled estimator is consistent and has a Gaussian limiting distribution when properly normalized.

Chapter 2

Simple and Weighted Symmetric Estimation

2.1 AR(1) process

In this chapter, we consider the simple symmetric and the weighted symmetric estimator in testing for a unit root in the panel data model. Let y_{it} be the first order autoregressive time series with fixed individual effects.

$$y_{it} = \eta_i + \rho y_{it-1} + e_{it}, \quad i = 1, 2, \cdots, N; t = 1, 2, \cdots, T, \tag{2.1}$$

where the e_{it}'s are independent and identically distributed as $N(0, \sigma^2)$ and y_{i0}'s are fixed. Under the stationarity condition, that is, $|\rho| < 1$, we can have a backward representation of the autoregressive time series. That is,

$$y_{it} = \eta_i + \rho y_{it+1} + u_{it},$$

where e_{it} and u_{it} have same covariance structure for each i. The discussion follows the logic of the univariate case as laid out by Dickey, Hasza, and Fuller (1984). In their paper, the following univariate seasonal model with seasonal period d is considered.

$$y_t = \sum_{i=1}^{d} \theta_i x_{it} + \alpha y_{t-d} + e_t; t = 1, 2, \cdots, n, \tag{2.2}$$

where

$$\begin{aligned} x_{it} &= 1 \quad \text{if} \quad t = i(\text{mod} \quad d) \\ &= 0 \quad \text{otherwise,} \end{aligned}$$

y_{-d+1},$y_{-d+2}, \cdots, y_0$ are fixed initial values and the e_t's are independent and identically distributed with mean 0 and variance σ^2.

The seasonal process in (2.2) can be partitioned into d component series such that

$$y_{dk+i} = \theta_i + \alpha y_{d(k-1)+i} + e_{dk+i}; i = 0, 1, \cdots, d-1,$$

and each component runs from $k = 1$ to m where the total number of observations is $n = dm$. This partition into d component series establishes a useful link between model (2.1) and (2.2). In other words, θ_i in the seasonal process corresponds to an individual intercept η_i in the panel model whereas the seasonal period d corresponds to the number of panels.

The weighted symmetric estimator of ρ in the model (2.1) is the estimator that minimizes

$$\sum_{i=1}^{N}\sum_{t=1}^{T} w_t \left[y_{it} - \eta_i - \rho y_{it-1}\right]^2 + \sum_{i=1}^{N}\sum_{t=0}^{T-1} (1 - w_{t+1}) \left[y_{it} - \eta_i - \rho y_{it+1}\right]^2 .$$

If the weight $w_t \equiv 1$, the estimator is the ordinary least squares estimator. The simple symmetric estimator which is discussed by Dickey, Hasza, and Fuller (1984) can be obtained by setting $w_t \equiv 0.5$. In weighted symmetric estimation for the first order autoregressive process, the weight is defined as follows.

$$w_t = \begin{array}{ll} 0, & t = 1 \\ \dfrac{t-1}{T} & , \quad t = 2, 3, \cdots, T. \end{array}$$

The weighted symmetric estimator of ρ can be obtained from Table 2.1.

Let $\beta' = [\eta_1, \eta_2, \cdots \eta_N, \rho]$.

Then the weighted symmetric estimator of β is

Table 2.1: Regression table for weighted symmetric estimation in AR(1) process

Weight	Dependent Variable	Parameter				
		η_1	η_2	$\cdots$	η_N	ρ
w_1	y_{11}	1	0	$\cdots$	0	y_{10}
w_2	y_{12}	1	0	$\cdots$	0	y_{11}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
w_T	y_{1T}	1	0	$\cdots$	0	y_{1T-1}
$1-w_T$	y_{1T-1}	1	0	$\cdots$	0	y_{1T}
$1-w_{T-1}$	y_{1T-2}	1	0	$\cdots$	0	y_{1T-1}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$1-w_1$	y_{10}	1	0	$\cdots$	0	y_{11}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
w_1	y_{N1}	0	0	$\cdots$	1	y_{N0}
w_2	y_{N2}	0	0	$\cdots$	1	y_{N1}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
w_T	y_{NT}	0	0	$\cdots$	1	y_{NT-1}
$1-w_T$	y_{NT-1}	0	0	$\cdots$	1	y_{NT}
$1-w_{T-1}$	y_{NT-2}	0	0	$\cdots$	1	y_{NT-1}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$1-w_1$	y_{N0}	0	0	$\cdots$	1	y_{N1}

$$\hat{\beta} = (\mathbf{X'WX})^{-1}(\mathbf{X'WY}),$$

where $\mathbf{X}$ is the (2NT)×(N+1) matrix that consists of columns corresponding to parameters, $\mathbf{Y}$ is the vector of the dependent variables, and $\mathbf{W}$ is the (2NT)×(2NT) diagonal matrix of weights. Then the simple symmetric estimator of ρ is

$$\hat{\rho}_s = Q_s^{-1}\Big[\sum_{i=1}^{N}\sum_{t=1}^{T} y_{it}y_{it-1} - \frac{1}{T}\sum_{i=1}^{N}(\sum_{t=2}^{T} y_{it-1} + 0.5y_{iT} + 0.5y_{i0})^2\Big] \tag{2.3}$$

where

$$Q_s = \sum_{i=1}^{N}(\sum_{t=2}^{T} {y_{it-1}}^2 + 0.5{y_{iT}}^2 + 0.5{y_{i0}}^2) - \frac{1}{T}\sum_{i=1}^{N}(\sum_{t=2}^{T} y_{it-1} + 0.5y_{iT} + 0.5y_{i0})^2.$$

The simple symmetric estimator of an individual intercept η_i is

$$\begin{aligned}\hat{\eta}_{is} &= \frac{1}{TQ_s}(\sum_{t=2}^{T} y_{it-1} + 0.5y_{iT} + 0.5y_{i0})\Big[\sum_{j=1}^{N}(\sum_{t=2}^{T} y_{jt-1}^2 + 0.5y_{jT}^2 + 0.5y_{j0}^2) \\ &\quad - \sum_{j=1}^{N}\sum_{t=1}^{T} y_{jt}y_{jt-1}\Big].\end{aligned}$$

On the other hand, the weighted symmetric estimator of ρ is

$$\begin{aligned}\hat{\rho}_{ws} &= Q_{ws}^{-1}\Big[\sum_{i=1}^{N}\sum_{t=1}^{T} y_{it}y_{it-1} - \frac{1}{T}\sum_{i=1}^{N}(\frac{T^2-1}{T^2}(\sum_{t=2}^{T} y_{it-1})^2 + \frac{T^2+T-2}{T^2}(\sum_{t=2}^{T} y_{it-1})y_{iT} \\ &\quad + \frac{T-1}{T^2}{y_{iT}}^2 + \frac{T+1}{T}y_{i0}\sum_{t=2}^{T} y_{it-1} + \frac{1}{T}y_{iT}y_{i0})\Big] \qquad (2.4)\end{aligned}$$

where

$$Q_{ws} = \sum_{i=1}^{N}\Big[\frac{T+1}{T}\sum_{t=2}^{T} {y_{it-1}}^2 + \frac{1}{T}{y_{iT}}^2\Big] - \frac{1}{T}\sum_{i=1}^{N}\Big[\frac{T+1}{T}\sum_{t=2}^{T} y_{it-1} + \frac{1}{T}y_{iT}\Big]^2.$$

Also, the weighted symmetric estimator of an individual intercept η_i is

$$\begin{aligned}\hat{\eta}_{iws} &= \frac{T-1}{T^2}(\sum_{t=2}^{T} y_{it-1} + y_{iT}) + \frac{1}{T}y_{i0} + \frac{1}{TQ_{ws}}\Big[\frac{T^2-1}{T^3}\sum_{j=1}^{N}(\sum_{t=2}^{T} y_{jt-1})^2 \\ &\quad + \frac{T^2+T-2}{T^3}\sum_{j=1}^{N}(\sum_{t=2}^{T} y_{jt-1})y_{jT} + \frac{T-1}{T^3}\sum_{j=1}^{N} y_{jT}^2 + \frac{T+1}{T^2}\sum_{j=1}^{N}(\sum_{t=2}^{T} y_{jt-1})y_{j0} \\ &\quad + \frac{1}{T^2}\sum_{j=1}^{N} y_{jT}y_{j0} - \sum_{j=1}^{N}\sum_{t=1}^{T} y_{jt}y_{jt-1}\Big](\frac{T+1}{T}\sum_{t=2}^{T} y_{it-1} + \frac{1}{T}y_{iT}).\end{aligned}$$

Based on the estimators in (2.3) and (2.4), test statistics for $H_0 : \rho = 1$ can be constructed. The normalized bias test statistics are

$$\sqrt{N}T(\hat{\rho}_s - 1 - b_s) = \frac{\frac{1}{\sqrt{N}}\sum_{i=1}^{N}[\sum_{t=1}^{T}(y_{it} - y_{it-1})y_{it-1} + \frac{1}{2}(y_{i0}^2 - {y_{iT}}^2) - b_s Q_{si}]/T}{\frac{1}{N}\sum_{i=1}^{N} Q_{si}/T^2} \quad (2.5)$$

and

$$\sqrt{N}T(\hat{\rho}_{ws} - 1 - b_{ws}) = \frac{\frac{1}{\sqrt{N}}\sum_{i=1}^{N}[A_i - b_{ws}Q_{wsi}]/T}{\frac{1}{N}\sum_{i=1}^{N} Q_{wsi}/T^2}, \quad (2.6)$$

where

$$b_s = -\frac{6T}{2T^2+1} = -\frac{3}{T} + O(1/T^2),$$

$$b_{ws} = -\frac{2T^2+T+2}{T^3+1} = -\frac{2}{T} + O(1/T^2),$$

$$Q_{si} = \sum_{t=2}^{T}{y_{it-1}}^2 + 0.5({y_{iT}}^2 + y_{i0}^2) - \frac{1}{T}[\sum_{t=2}^{T} y_{it-1} + 0.5(y_{iT} + y_{i0})]^2,$$

$$Q_{wsi} = \frac{T+1}{T}\sum_{t=2}^{T}{y_{it-1}}^2 + \frac{1}{T}{y_{iT}}^2 - \frac{1}{T}[\frac{T+1}{T}\sum_{t=2}^{T} y_{it-1} + \frac{1}{T}y_{iT}]^2,$$

and

$$\begin{aligned} A_i &= \frac{2(T+1)}{T^3}(\sum_{t=2}^{T} y_{it-1})^2 + \frac{-T^2+T+4}{T^3}(\sum_{t=2}^{T} y_{it-1})y_{iT} - \frac{T^2+T-2}{T^3}{y_{iT}}^2 \\ &+ \sum_{t=1}^{T}(y_{it} - y_{it-1})y_{it-1} + y_{i0}^2 - \frac{1}{T}\sum_{t=2}^{T}{y_{it-1}}^2 - \frac{1}{T^2}y_{i0}((T+1)\sum_{t=2}^{T} y_{it-1} + y_{iT}). \end{aligned}$$

Also, we consider the following studentized test statistics.

$$\hat{\tau}_s = \frac{\hat{\rho}_s - 1 - b_s}{\sqrt{\hat{\sigma}_s^2 Q_s^{-1}}}, \tag{2.7}$$

and

$$\hat{\tau}_{ws} = \frac{\hat{\rho}_{ws} - 1 - b_{ws}}{\sqrt{\hat{\sigma}_{ws}^2 Q_{ws}^{-1}}} \tag{2.8}$$

where

$$\hat{\sigma}_s^2 = \frac{1}{NT - N - 1}[\mathbf{Y'W}_s\mathbf{Y} - \mathbf{Y'W}_s\mathbf{X}(\mathbf{X'W}_s\mathbf{X})^{-1}\mathbf{X'W}_s\mathbf{Y}],$$

and

$$\hat{\sigma}_{ws}^2 = \frac{1}{NT - N - 1}[\mathbf{Y'W}_{ws}\mathbf{Y} - \mathbf{Y'W}_{ws}\mathbf{X}(\mathbf{X'W}_{ws}\mathbf{X})^{-1}\mathbf{X'W}_{ws}\mathbf{Y}]$$

Here, $\mathbf{W}_s$ and $\mathbf{W}_{ws}$ are the diagonal matrices of weights for simple symmetric and weighted symmetric estimations, respectively.

Now, we discuss the joint limiting distributions of the above test statistics as $N, T \to \infty$. To derive the joint limiting distributions of test statistics under $H_0 : \rho = 1$ as $N, T \to \infty$, we employ the following joint limit results from Phillips and Moon (1999). In their paper, they present the following theorems for random vectors instead of random variables.

Theorem 2.1: (Joint Probability Limits) Suppose the random variables Y_{iT} are independent across i for all T and integrable. Let $X_{N,T} = \frac{\sum_{i=1}^{N} Y_{iT}}{N}$ which is square integrable and μ_x=$\lim_{N,T \to \infty} E(X_{N,T})$=$\lim_{N,T \to \infty} \frac{1}{N} \sum_{i=1}^{N} E(Y_{iT})$ be finite. If $\lim_{N,T \to \infty} \frac{1}{N^2} \sum_{i=1}^{N} Var(Y_{iT}) = 0$, then

$$X_{N,T} \xrightarrow{P} \mu_x \quad \text{as} \quad N, T \to \infty.$$

Proof:

It is sufficient to show that $\lim_{N,T\to\infty} P\{|\frac{1}{N}\sum_{i=1}^{N}(Y_{iT}-E(Y_{iT}))|>\epsilon\}=0\ \forall\epsilon>0$. By Chebyshev's inequality,

$$\begin{aligned}\lim_{N,T\to\infty} P\{|\frac{1}{N}\sum_{i=1}^{N}(Y_{iT}-E(Y_{iT}))|>\epsilon\} &\leq \frac{1}{\epsilon^2}\lim_{N,T\to\infty} E|\frac{1}{N}\sum_{i=1}^{N}(Y_{iT}-E(Y_{iT}))|^2\\ &= \frac{1}{\epsilon^2}\lim_{N,T\to\infty}\frac{1}{N^2}\sum_{i=1}^{N}E|Y_{iT}-E(Y_{iT})|^2\\ &= \frac{1}{\epsilon^2}\lim_{N,T\to\infty}\frac{1}{N^2}\sum_{i=1}^{N}Var(Y_{iT})\\ &= 0.\end{aligned}$$

The first equality comes from the independence. If Y_{iT}'s are independent and identically distributed across i for all T, the condition for Theorem 2.1 reduces to the finiteness of $\lim_{T\to\infty} Var(Y_{iT})$. The following theorem gives us the condition for the joint limit Central Limit Theorem.

Theorem 2.2: (Joint Limit CLT) Suppose that Y_{iT} are independent and identically distributed with mean 0 and variance σ_T^2 across i for all T. Assume the following conditions hold.

i) Y_{iT}^2 are uniformly integrable in T.

ii) $\lim_{T\to\infty}\sigma_T^2=\sigma_\infty^2>0$.

Then,

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N}Y_{iT}\overset{D}{\Longrightarrow}N(0,\sigma_\infty^2)\quad\text{as}\quad N,T\to\infty.$$

Proof:

Phillips and Moon (1999) show that Lindeberg's theorem holds in panel data. In other words, let $Z_{i,N,T}=\dfrac{Y_{iT}}{\sqrt{\sum_{j=1}^{N}Var(Y_{jT})}}$ and $\forall\epsilon>0$,

$$\lim_{N,T\to\infty}\sum_{i=1}^{N}E[Z_{i,N,T}^2 1\{|Z_{i,N,T}|>\epsilon\}]=0.$$

Then as $N, T \to \infty$,

$$\sum_{i=1}^{N} Z_{i,N,T} \xrightarrow{D} N(0,1)$$

Therefore we need to prove that $\forall \epsilon > 0$,

$$\lim_{N,T \to \infty} \sum_{i=1}^{N} E[\frac{Y_{iT}^2}{N\sigma_T^2} 1\{|Y_{iT}| > \epsilon\sqrt{N\sigma_T^2}\}] = 0.$$

Since Y_{iT} are identically distributed across i for all T,

$$\begin{aligned} &\sum_{i=1}^{N} E[\frac{Y_{iT}^2}{N\sigma_T^2} 1\{|Y_{iT}| > \epsilon\sqrt{N\sigma_T^2}\}] \\ &= E[\frac{Y_{1T}^2}{\sigma_T^2} 1\{|Y_{1T}| > \epsilon\sqrt{N\sigma_T^2}\}]. \end{aligned} \tag{2.9}$$

By the assumption of uniformly integrability of Y_{1T}^2, (2.9) goes to 0 as $N, T \to \infty$. Since it satisfies the Lindeberg condition,

$$\frac{\sum_{i=1}^{N} Y_{iT}}{\sqrt{N\sigma_T^2}} \xrightarrow{D} N(0,1).$$

Since $\lim_{N,T \to \infty} \frac{\sigma_T^2}{\sigma_\infty^2} = 1$,

$$\frac{\sum_{i=1}^{N} Y_{iT}}{\sqrt{N\sigma_\infty^2}} \xrightarrow{D} N(0,1).$$

Now we apply these joint limit results to the test statistics in (2.5)-(2.8).

Theorem 2.3: Given model (2.1), under the null hypothesis $H_0 : \rho = 1$ and assuming $\eta_i = y_{i0} = 0$ for every i, then

$$i) \quad \sqrt{N}T(\hat{\rho}_s - 1 - b_s) \xrightarrow{D} N(0, 7.2) \quad \text{as} \quad N, T \to \infty,$$

$$ii) \quad \hat{\tau}_s = \frac{\hat{\rho}_s - 1 - b_s}{[\hat{\sigma}_s^2 Q_s^{-1}]^{0.5}} \stackrel{D}{\Longrightarrow} N(0, 1.2) \quad \text{as} \quad N, T \to \infty,$$

$$iii) \quad \sqrt{N}T(\hat{\rho}_{ws} - 1 - b_{ws}) \stackrel{D}{\Longrightarrow} N(0, 9) \quad \text{as} \quad N, T \to \infty,$$

$$iv) \quad \hat{\tau}_{ws} = \frac{\hat{\rho}_{ws} - 1 - b_{ws}}{[\hat{\sigma}_{ws}^2 Q_{ws}^{-1}]^{0.5}} \stackrel{D}{\Longrightarrow} N(0, 1.5) \quad \text{as} \quad N, T \to \infty.$$

Proof of i):

Under the assumptions, the normalized bias test statistic for $H_0 : \rho = 1$ is

$$\sqrt{N}T(\hat{\rho}_s - 1 - b_s) = \frac{\frac{1}{\sqrt{N}} \sum_{i=1}^{N} [\sum_{t=2}^{T} (y_{it} - y_{it-1}) y_{it-1} - 0.5 {y_{iT}}^2 - b_s Q_{si}]/T}{\frac{1}{N} \sum_{i=1}^{N} Q_{si}/T^2} \quad (2.10)$$

where $Q_{si} = \sum_{t=2}^{T} {y_{it-1}}^2 + 0.5 {y_{iT}}^2 - \frac{1}{T} [\sum_{t=2}^{T} y_{it-1} + 0.5 y_{iT}]^2$.

If we apply the joint probability limit result to the denominator and the joint limit CLT result to the numerator in (2.10), we find that the normalized bias test statistic has a Gaussian limiting distribution as $N, T \to \infty$. For convenience, the first and the second finite sample moments of random variables in the test statistics under the null hypothesis are shown in Appendix A.

Now,

$$E(Q_{si}/T^2) = \frac{2T^2 + 1}{12T^2} \sigma^2 \stackrel{T \to \infty}{\longrightarrow} \frac{1}{6} \sigma^2.$$

To prove that $\frac{1}{N} \sum_{i=1}^{N} Q_{si}/T^2 \stackrel{P}{\longrightarrow} \frac{1}{6} \sigma^2$, we need to show $\lim_{T \to \infty} Var(Q_{si}/T^2) < \infty$. Since

$$Var(Q_{si}/T^2) = \frac{8T^4 + 20T^2 + 17}{360T^4} \sigma^4 \stackrel{T \to \infty}{\longrightarrow} \frac{\sigma^4}{45} < \infty,$$

from Theorem 2.1,

$$\frac{1}{N}\sum_{i=1}^{N} Q_{si}/T^2 \xrightarrow{P} \frac{1}{6}\sigma^2$$

as $N, T \to \infty$.

Now, we will prove the numerator in (2.10) satisfies the conditions for Theorem 2.2. From the first and second moment results in Appendix A, we find that

$$E\Big[\frac{\sum_{t=2}^{T}(y_{it}-y_{it-1})y_{it-1} - 0.5{y_{iT}}^2 - b_s Q_{si}}{T}\Big] = 0,$$

and

$$\begin{aligned} Var\left[\frac{\sum_{t=2}^{T}(y_{it}-y_{it-1})y_{it-1} - 0.5{y_{iT}}^2 - b_s Q_{si}}{T}\right] &= \frac{\sigma^4}{10(2T^2+1)^2 T}\Big[8T^5 \\ &\quad -20T^2(T^2-T+1)+17T-5\Big]. \end{aligned}$$

Defining Z_{iT} as $[\sum_{t=2}^{T}(y_{it}-y_{it-1})y_{it-1} - 0.5{y_{iT}}^2 - b_s Q_{si}]/T$, we need to show that Z_{iT}^2 is uniformly integrable in T. If we consider the T-limit distribution of Z_{iT}^2,

$$Z_{iT}^2 \xRightarrow{D} \Big[-\frac{1}{2} + 3\Big(\int_0^1 B_i^2(t)dt - \big(\int_0^1 B_i(t)dt\big)^2\Big)\Big]^2\sigma^4 \stackrel{def}{=} Z_i^2$$

as $T \to \infty$,

where $B_i(t)$ is a Wiener process (Phillips,1987).

Since

$$\lim_{T\to\infty} E(Z_{iT}^2) = E(Z_i^2) = \frac{\sigma^4}{5},$$

Z_{iT}^2 are uniformly integrable in T.

Hence, by Theorem 2.2,

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N}[\sum_{t=2}^{T}(y_{it}-y_{it-1})y_{it-1} - 0.5{y_{iT}}^2 - b_s Q_{si}]/T \xRightarrow{D} N(0, \frac{\sigma^4}{5}).$$

Therefore,

$$\sqrt{N}T(\hat{\rho}_s - 1 - b_s) \stackrel{D}{\Longrightarrow} N(0, 7.2)$$

as $N, T \to \infty$ by Slutsky's Theorem.

Proof of ii):

The studentized test statistic from simple symmetric estimation is

$$\hat{\tau}_s = \frac{\frac{1}{\sqrt{N}}\sum_{i=1}^{N}[\sum_{t=2}^{T}(y_{it} - y_{it-1})y_{it-1} - 0.5{y_{iT}}^2 - b_s Q_{si}]/T}{\sqrt{\hat{{\sigma^2}_s}\frac{1}{N}\sum_{i=1}^{N}Q_{si}/T^2}}$$

Since $\hat{{\sigma^2}_s} \xrightarrow{P} \sigma^2$ and $\frac{1}{N}\sum_{i=1}^{N}Q_{si}/T^2 \xrightarrow{P} \frac{1}{6}\sigma^2$ as $N, T \to \infty$, the denominator in the studentized statistic converges to $\sigma^2\sqrt{\frac{1}{6}}$ in probability. Since the numerator converges to $N(0, \frac{\sigma^4}{5})$ in distribution from the Proof of i),

$$\hat{\tau}_s \stackrel{D}{\Longrightarrow} N(0, 1.2).$$

Proof of iii):

This proof is analogous to the proof of i). We apply joint limit results to the normalized bias test statistic in (2.6). If we look at the moments of the denominator in (2.6) first,

$$E(Q_{wsi}/T^2) = \frac{T^4 - T^3 + T - 1}{6T^4}\sigma^2 \to \frac{\sigma^2}{6},$$

and

$$Var(Q_{wsi}/T^2) = \frac{2T^8 - 4T^7 + 11T^6 - 8T^5 - 17T^4 + 34T^3 - T^2 - 22T + 5}{90T^8}\sigma^4.$$

Since $\lim_{T\to\infty} Var(Q_{wsi}/T^2) = \frac{\sigma^4}{45} < \infty$,

$$\frac{1}{N}\sum_{i=1}^{N} Q_{wsi}/T^2 \xrightarrow{P} \frac{\sigma^2}{6}.$$

as $N, T \to \infty$.

The first and second moments of the numerator in (2.6) are

$$E([A_i - b_{ws}Q_{wsi}]/T) = 0,$$

and

$$\begin{aligned} Var([A_i - b_{ws}Q_{wsi}]/T) &= \frac{\sigma^4\Big[15T^8 - 76T^7 + 167T^6 - 188T^5 + 84T^4 - 4T^3\Big]}{60T^4(T^2 - T + 1)^2} \\ &\quad + \frac{\sigma^4\Big[13T^2 - 2T - 9\Big]}{60T^4(T^2 - T + 1)^2}. \end{aligned}$$

As in the proof of (i), let us define $P_{iT} = [A_i - b_{ws}Q_{wsi}]/T$. To prove that P_{iT}^2 is uniformly integrable in T, we consider the T-limit distribution of P_{iT}^2.

$$P_{iT}^2 \xRightarrow{D} \Big[\frac{1}{2}(B_i^2(1) - 1) + \int_0^1 B_i^2(t)dt - (\int_0^1 B_i(t)dt)B_i(1)\Big]^2\sigma^4 \stackrel{def}{=} P_i^2.$$

as $T \to \infty$.

Since

$$\lim_{T\to\infty} E(P_{iT}^2) = E(P_i^2) = \frac{\sigma^4}{4},$$

P_{iT}^2 is uniformly integrable in T. Since it satisfies the condition of Theorem 2.2,

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N}[A_i - b_{ws}Q_{wsi}]/T \xRightarrow{D} N(0, \frac{\sigma^4}{4}).$$

Hence,

$$\sqrt{NT}(\hat{\rho}_{ws} - 1 - b_{ws}) \Longrightarrow^{D} N(0, 9)$$

as $N, T \to \infty$.

Proof of iv):

By arguments similar to those in (ii), we can prove that

$$\hat{\tau}_{ws} \Longrightarrow^{D} N(0, 1.5)$$

as $N, T \to \infty$.

Now we present the empirical percentiles of the test statistics in (2.5)-(2.8) from Monte Carlo simulations. We conduct simulations for the following combinations of N and T.

$$\begin{aligned}(N,T) &= (10,25),(10,50),(10,100),(25,25),(25,50),(25,100),\\ &\quad (50,25),(50,50),\text{and } (50,100).\end{aligned}$$

For each set, we generate y_{it} from the random walk model, that is, $y_{it} = y_{it-1} + e_{it}$ with 10,000 replicates, where the e_{it}'s are generated from N(0,1) and $y_{i0} = 0$ for each i . As a random number generator, we use the RANNOR function from SAS. The empirical percentiles of the normalized bias test statistics for the simple symmetric and the weighted symmetric estimators are shown in Table 2.2 and Table 2.3, respectively. Also, the empirical percentiles of the studentized test statistics are shown in Table 2.4 and Table 2.5, respectively. The percentiles for the limiting distribution in each table can be computed from the standard Gaussian distribution table. Numbers in the tables are unsmoothed. However a smoothing surface describes the 5 % critical values reasonably well. In order to smooth the 5 % critical values of the studentized test statistics, we compute the deviations of the 5 % empirical percentiles in finite samples from the limiting 5 % percentiles. Then we regress these deviations

on $\frac{1}{\sqrt{N}}$ and $\frac{1}{\sqrt{T}}$. For simple symmetric estimation, R^2 is 0.97 and the smoothed percentiles are given by $-1.81 - 0.4\frac{1}{\sqrt{N}} + 0.16\frac{1}{\sqrt{T}}$. For the weighted symmetric regression, R^2 is 0.86 and the smoothing function for the empirical percentiles is $-2.02 - 0.34\frac{1}{\sqrt{N}} + 0.22\frac{1}{\sqrt{T}}$.

Table 2.2: Empirical cumulative distribution of $\sqrt{NT}(\hat{\rho}_s - 1 - b_s)$

	Probability of a Smaller Value								
(N,T)	0.01	0.05	0.10	0.25	0.50	0.75	0.90	0.95	0.99
(10,25)	-8.45	-5.74	-4.31	-2.22	-0.27	1.28	2.55	3.12	4.10
(10,50)	-9.02	-5.83	-4.44	-2.30	-0.29	1.34	2.56	3.20	4.15
(10,100)	-9.00	-6.01	-4.63	-2.39	-0.38	1.32	2.55	3.19	4.28
(25,25)	-7.42	-5.18	-3.94	-2.10	-0.18	1.41	2.75	3.45	4.68
(25,50)	-7.87	-5.30	-4.08	-2.10	-0.18	1.47	2.78	3.54	4.74
(25,100)	-7.68	-5.36	-4.06	-2.12	-0.21	1.49	2.86	3.56	4.89
(50,25)	-7.24	-4.91	-3.73	-1.97	-0.16	1.47	2.81	3.58	5.05
(50,50)	-7.41	-4.97	-3.85	-1.99	-0.12	1.56	2.96	3.75	5.20
(50,100)	-7.20	-4.94	-3.75	-1.99	-0.12	1.59	2.95	3.76	5.10
(∞,∞)	-6.25	-4.43	-3.43	-1.80	0.00	1.80	3.43	4.43	6.25

($b_s = -6T/(2T^2 + 1)$.)

Table 2.3: Empirical cumulative distribution of $\sqrt{N}T(\hat{\rho}_{ws} - 1 - b_{ws})$

	Probability of a Smaller Value								
(N,T)	0.01	0.05	0.10	0.25	0.50	0.75	0.90	0.95	0.99
(10,25)	-9.39	-6.35	-4.78	-2.45	-0.38	1.50	3.06	3.89	5.20
(10,50)	-9.67	-6.39	-4.87	-2.59	-0.34	1.52	3.03	3.89	5.20
(10,100)	-9.57	-6.50	-4.91	-2.69	-0.37	1.57	3.02	3.87	5.24
(25,25)	-8.21	-5.76	-4.52	-2.38	-0.25	1.68	3.26	4.12	5.65
(25,50)	-8.54	-5.93	-4.51	-2.37	-0.18	1.68	3.28	4.21	5.69
(25,100)	-8.52	-5.79	-4.42	-2.31	-0.20	1.72	3.28	4.17	5.68
(50,25)	-8.22	-5.47	-4.26	-2.24	-0.17	1.73	3.34	4.31	5.89
(50,50)	-7.87	-5.53	-4.26	-2.24.	-0.12	1.80	3.44	4.39	5.97
(50,100)	-8.03	-5.39	-4.15	-2.22	-0.12	1.84	3.46	4.37	6.04
(∞,∞)	-6.99	-4.95	-3.84	-2.01	0.00	2.01	3.84	4.95	6.99

($b_{ws} = -(2T^2 + T + 2)/(T^3 + 1)$.)

Table 2.4: Empirical cumulative distribution of $\hat{\tau}_s$

	Probability of a Smaller Value								
(N,T)	0.01	0.05	0.10	0.25	0.50	0.75	0.90	0.95	0.99
(10,25)	-2.61	-1.90	-1.49	-0.83	-0.11	0.56	1.22	1.55	2.22
(10,50)	-2.69	-1.90	-1.51	-0.85	-0.12	0.59	1.22	1.60	2.25
(10,100)	-2.66	-1.93	-1.56	-0.88	-0.15	0.58	1.22	1.60	2.36
(25,25)	-2.54	-1.86	-1.46	-0.82	-0.08	0.61	1.25	1.61	2.30
(25,50)	-2.64	-1.88	-1.49	-0.81	-0.07	0.63	1.26	1.66	2.34
(25,100)	-2.57	-1.89	-1.48	-0.81	-0.08	0.64	1.30	1.66	2.43
(50,25)	-2.61	-1.84	-1.43	-0.78	-0.07	0.63	1.24	1.61	2.37
(50,50)	-2.63	-1.84	-1.46	-0.78	-0.05	0.66	1.30	1.69	2.45
(50,100)	-2.55	-1.83	-1.42	-0.78	-0.05	0.68	1.30	1.69	2.39
(∞,∞)	-2.55	-1.81	-1.40	-0.73	0.00	0.73	1.40	1.81	2.55

Table 2.5: Empirical cumulative distribution of $\hat{\tau}_{ws}$

	Probability of a Smaller Value								
(N,T)	0.01	0.05	0.10	0.25	0.50	0.75	0.90	0.95	0.99
(10,25)	-2.89	-2.10	-1.65	-0.91	-0.15	0.64	1.39	1.80	2.57
(10,50)	-2.88	-2.09	-1.67	-0.95	-0.13	0.66	1.39	1.85	2.62
(10,100)	-2.87	-2.11	-1.68	-0.99	-0.15	0.69	1.40	1.85	2.74
(25,25)	-2.77	-2.05	-1.65	-0.90	-0.10	0.70	1.40	1.83	2.60
(25,50)	-2.87	-2.08	-1.63	-0.91	-0.07	0.71	1.45	1.90	2.68
(25,100)	-2.86	-2.05	-1.62	-0.88	-0.08	0.73	1.46	1.90	2.76
(50,25)	-2.90	-2.01	-1.60	-0.86	-0.07	0.71	1.42	1.86	2.63
(50,50)	-2.80	-2.05	-1.61	-0.87	-0.05	0.75	1.48	1.93	2.72
(50,100)	-2.88	-2.00	-1.56	-0.87	-0.05	0.77	1.49	1.92	2.74
(∞, ∞)	-2.85	-2.02	-1.57	-0.82	0.00	0.82	1.57	2.02	2.85

2.2 Higher Order Processes

In this section, we consider the limiting distributions of the simple symmetric and the weighted symmetric estimators for the autoregressive higher order processes. If we generalize model (2.1) to a pth order autoregressive process,

$$y_{it} = \eta_i - \sum_{k=1}^{p} \alpha_k y_{it-k} + e_{it}, \quad i = 1, 2, \cdots, N; t = 1, 2, \cdots, T, \tag{2.11}$$

where the e_{it}'s are independent and identically distributed as $N(0, \sigma^2)$ and $y_{i,-p+1}$,$y_{i,-p+2}$,$\cdots$,y_{i0} are fixed. We consider the case when $p-1$ roots of

$$m^p + \sum_{k=1}^{p} \alpha_k m^{p-k} = 0$$

are less than one and there exists only one unit root. We can rewrite model (2.11) as

$$y_{it} = \eta_i + \theta_1 y_{it-1} + \sum_{k=2}^{p} \theta_k \Delta y_{it-k+1} + e_{it} \tag{2.12}$$

where

$$\theta_1 = -\sum_{k=1}^{p} \alpha_k,$$

$$\theta_l = \sum_{k=l}^{p} \alpha_k, \quad l = 2, 3, \cdots, p$$

and Δ signifies the first lag difference operator, that is, $\Delta y_{it} = y_{it} - y_{it-1}$. In model (2.12), the null hypothesis for testing a unit root is $\mathrm{H}_0 : \theta_1 = 1$. In this section, we derive the simple symmetric and weighted symmetric estimators when $p = 2$. Extension to general p is straightforward. From (2.12), a second order autoregressive model is

$$y_{it} = \eta_i + \theta_1 y_{it-1} + \theta_2 \Delta y_{it-1} + e_{it}. \tag{2.13}$$

As in the first order autoregressive process, a backward representation can be valid under the stationarity condition, that is,

$$y_{it} = \eta_i + \theta_1 y_{it+1} - \theta_2 \Delta y_{it+2} + u_{it},$$

where e_{it} and u_{it} have the same covariance structure for each i. For simple symmetric estimation, the regression uses weights $w_t = 0.5$. In weighted symmetric estimation of a pth order process, the weight is defined as follows.

$$w_t = \begin{array}{ll} 0, & t = 1, 2, \cdots, p, \\ \dfrac{t-p}{T-2p+2}, & t = p+1, p+2, \cdots, T-p+1, \\ 1, & t = T-p+2, T-p+3, \cdots, T. \end{array}$$

Fuller (1996) presents the regression table for weighted symmetric estimation in the univariate case. We generalize his table to panel data with fixed individual effects and obtain the weighted symmetric estimators from Table 2.6.
Let β' be a vector of parameters, that is, $\Big[\eta_1, \eta_2, \cdots, \eta_N, \theta_1, \theta_2\Big]$. Then the weighted symmetric estimator for β is

$$\hat{\beta} = (\mathbf{X'WX})^{-1}(\mathbf{X'WY}),$$

where $\mathbf{X}$ is the $2N(T-2) \times (N+2)$ matrix that consists of columns corresponding to $(N+2)$ parameters, $\mathbf{Y}$ is the vector of dependent variables, and $\mathbf{W}$ is the $2N(T-2) \times 2N(T-2)$ diagonal matrix of weights.

The simple symmetric estimator of θ_1 is

Table 2.6: Regression table for weighted symmetric estimation in AR(2) process

Weight	Dependent Variable	Parameter					
		η_1	η_2	$\cdots$	η_N	θ_1	θ_2
w_3	y_{13}	1	0	$\cdots$	0	y_{12}	Δy_{12}
w_4	y_{14}	1	0	$\cdots$	0	y_{13}	Δy_{13}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
w_T	y_{1T}	1	0	$\cdots$	0	y_{1T-1}	Δy_{1T-1}
$1-w_{T-1}$	y_{1T-2}	1	0	$\cdots$	0	y_{1T-1}	$-\Delta y_{1T}$
$1-w_{T-2}$	y_{1T-3}	1	0	$\cdots$	0	y_{1T-2}	$-\Delta y_{1T-1}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$1-w_2$	y_{11}	1	0	$\cdots$	0	y_{12}	$-\Delta y_{13}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
w_3	y_{N3}	0	0	$\cdots$	1	y_{N2}	Δy_{N2}
w_4	y_{N4}	0	0	$\cdots$	1	y_{N3}	Δy_{N3}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
w_T	y_{NT}	0	0	$\cdots$	1	y_{NT-1}	Δy_{NT-1}
$1-w_{T-1}$	y_{NT-2}	0	0	$\cdots$	1	y_{NT-1}	$-\Delta y_{NT}$
$1-w_{T-2}$	y_{NT-3}	0	0	$\cdots$	1	y_{NT-2}	$-\Delta y_{NT-1}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$1-w_2$	y_{N1}	0	0	$\cdots$	1	y_{N2}	$-\Delta y_{N3}$

$$
\begin{aligned}
\hat{\theta}_{1,s} &= \frac{1}{c_s}\Big[-\frac{1}{T-2}S_{22}\sum_{i=1}^{N}[(\sum_{t=3}^{T}y_{it-1})^2+\frac{1}{2}(\sum_{t=3}^{T}y_{it-1})(\Delta y_{iT}-\Delta y_{i2})] \\
&+ \frac{1}{2(T-2)}S_{12}\sum_{i=1}^{N}[(\sum_{t=3}^{T}y_{it-1})(\Delta y_{i2}-\Delta y_{iT})-\frac{1}{2}(\Delta y_{i2}-\Delta y_{iT})^2] \\
&+ S_{22}\sum_{i=1}^{N}[\sum_{t=3}^{T}y_{it}y_{it-1}+\frac{1}{2}(y_{i1}y_{i2}-y_{iT}y_{iT-1})] \\
&- S_{12}\sum_{i=1}^{N}[\sum_{t=3}^{T}y_{it}\Delta y_{it-1}+\frac{1}{2}(y_{i1}y_{i2}-y_{iT}y_{iT-1})]\Big], \qquad (2.14)
\end{aligned}
$$

where

$$S_{11} = \sum_{i=1}^{N}\sum_{t=3}^{T} {y_{it-1}}^2 - \frac{1}{T-2}\sum_{i=1}^{N}(\sum_{t=3}^{T} y_{it-1})^2,$$

$$\begin{aligned} S_{12} &= \sum_{i=1}^{N}\sum_{t=3}^{T} y_{it-1}\Delta y_{it-1} + \frac{1}{2}\sum_{i=1}^{N}(y_{i2}y_{i1} - y_{iT}y_{iT-1}) \\ &- \frac{1}{2(T-2)}\sum_{i=1}^{N}(\sum_{t=3}^{T} y_{it-1})(\Delta y_{i2} - \Delta y_{iT}), \end{aligned}$$

$$\begin{aligned} S_{22} &= \sum_{i=1}^{N}\sum_{t=3}^{T}(\Delta y_{it-1})^2 - \frac{1}{2}\sum_{i=1}^{N}[(\Delta y_{i2})^2 - (\Delta y_{iT})^2] \\ &- \frac{1}{4(T-2)}\sum_{i=1}^{N}(\Delta y_{i2} - \Delta y_{iT})^2, \end{aligned}$$

and

$$c_s = S_{11}S_{22} - {S_{12}}^2.$$

Also, the simple symmetric estimator of θ_2 is

$$\begin{aligned} \hat{\theta}_{2,s} &= \frac{1}{c_s}\Big[\frac{1}{T-2}S_{12}\sum_{i=1}^{N}[(\sum_{t=3}^{T} y_{it-1})^2 + \frac{1}{2}(\sum_{t=3}^{T} y_{it-1})(\Delta y_{iT} - \Delta y_{i2})] \\ &- \frac{1}{2(T-2)}S_{11}\sum_{i=1}^{N}[(\sum_{t=3}^{T} y_{it-1})(\Delta y_{i2} - \Delta y_{iT}) - \frac{1}{2}(\Delta y_{i2} - \Delta y_{iT})^2] \\ &- S_{12}\sum_{i=1}^{N}[\sum_{t=3}^{T} y_{it}y_{it-1} + \frac{1}{2}(y_{i1}y_{i2} - y_{iT}y_{iT-1})] \\ &+ S_{11}\sum_{i=1}^{N}[\sum_{t=3}^{T} y_{it}\Delta y_{it-1} + \frac{1}{2}(y_{i1}y_{i2} - y_{iT}y_{iT-1})]\Big]. \end{aligned} \tag{2.15}$$

On the other hand, the weighted symmetric estimator of θ_1 is

$$
\begin{aligned}
\hat{\theta}_{1,ws} &= \frac{1}{c_{ws}}\Big[-\frac{T-1}{(T-2)^2}W_{22}\sum_{i=1}^{N}[\frac{T-3}{T-2}(\sum_{t=3}^{T}y_{it-1})^2+(\sum_{t=3}^{T}y_{it-1})(y_{i1}+y_{iT})] \\
&- \frac{2}{(T-2)^2}W_{12}\sum_{i=1}^{N}[(\sum_{t=3}^{T}y_{it-1})(y_{i1}+y_{iT})+\frac{T-3}{T-2}(\sum_{t=3}^{T}y_{it-1})^2] \\
&+ \frac{T-1}{(T-2)^2}W_{12}\sum_{i=1}^{N}[(\Delta y_{i2}-\Delta y_{iT})(y_{i1}+y_{iT}+\frac{T-3}{T-2}\sum_{t=3}^{T}y_{it-1})] \\
&+ \frac{1}{T-2}W_{12}\sum_{i=1}^{N}[(y_{i1}+y_{iT})^2+\frac{T-3}{T-2}(\sum_{t=3}^{T}y_{it-1})(y_{i1}+y_{iT})] \\
&+ W_{22}\sum_{i=1}^{N}[y_{i1}y_{i2}+\sum_{t=3}^{T}y_{it}y_{it-1}] \\
&+ W_{12}\sum_{i=1}^{N}[\frac{T-1}{T-2}\sum_{t=3}^{T}y_{it}y_{it-2}-\sum_{t=3}^{T}y_{it}y_{it-1}-y_{i1}y_{i2}]\Big], \qquad (2.16)
\end{aligned}
$$

where

$$
W_{11} = \frac{T-1}{T-2}\sum_{i=1}^{N}\sum_{t=3}^{T}{y_{it-1}}^2-\frac{(T-1)^2}{(T-2)^3}\sum_{i=1}^{N}(\sum_{t=3}^{T}y_{it-1})^2,
$$

$$
\begin{aligned}
W_{12} &= \frac{1}{T-2}\sum_{i=1}^{N}[T\sum_{t=3}^{T}y_{it-1}\Delta y_{it-1}-\sum_{t=3}^{T}{y_{it-1}}^2-y_{iT}y_{iT-1}+(T-1)y_{i2}y_{i1}] \\
&+ \frac{T-1}{(T-2)^2}\sum_{i=1}^{N}[\frac{2}{T-2}(\sum_{t=3}^{T}y_{it-1})^2-\frac{T-1}{T-2}(\sum_{t=3}^{T}y_{it-1})(\Delta y_{i2}-\Delta y_{iT})] \\
&- \frac{T-1}{(T-2)^2}\sum_{i=1}^{N}(\sum_{t=3}^{T}y_{it-1})(y_{i1}+y_{iT}),
\end{aligned}
$$

$$
\begin{aligned}
W_{22} &= \frac{1}{T-2}\sum_{i=1}^{N}[(\Delta y_{i2})^2+(\Delta y_{iT})^2+T\sum_{t=3}^{T-1}(\Delta y_{it})^2] \\
&- \frac{1}{(T-2)^3}\sum_{i=1}^{N}[4(\sum_{t=3}^{T}y_{it-1})^2+(T-1)^2(\Delta y_{i2}-\Delta y_{iT})^2] \\
&- \frac{1}{T-2}\sum_{i=1}^{N}[(y_{i1}+y_{iT})^2-\frac{4(T-1)}{(T-2)^2}(\sum_{t=3}^{T}y_{it-1})(\Delta y_{i2}-\Delta y_{iT})] \\
&+ \frac{1}{(T-2)^2}\sum_{i=1}^{N}[4(\sum_{t=3}^{T}y_{it-1})(y_{i1}+y_{iT})-2(T-1)(\Delta y_{i2}-\Delta y_{iT})(y_{i1}+y_{iT})],
\end{aligned}
$$

and

$$c_{ws} = W_{11}W_{22} - W_{12}^2.$$

Also, the weighted symmetric estimator of θ_2 is

$$\begin{aligned}
\hat{\theta}_{2,ws} &= \frac{1}{c_{ws}}\Big[\frac{T-1}{(T-2)^2}W_{12}\sum_{i=1}^{N}[\frac{T-3}{T-2}(\sum_{t=3}^{T}y_{it-1})^2 + (\sum_{t=3}^{T}y_{it-1})(y_{i1}+y_{iT})] \\
&+ \frac{2}{(T-2)^2}W_{11}\sum_{i=1}^{N}[(\sum_{t=3}^{T}y_{it-1})(y_{i1}+y_{iT}) + \frac{T-3}{T-2}(\sum_{t=3}^{T}y_{it-1})^2] \\
&- \frac{T-1}{(T-2)^2}W_{11}\sum_{i=1}^{N}[(\Delta y_{i2}-\Delta y_{iT})(y_{i1}+y_{iT}+\frac{T-3}{T-2}\sum_{t=3}^{T}y_{it-1})] \\
&- \frac{1}{T-2}W_{11}\sum_{i=1}^{N}[(y_{i1}+y_{iT})^2 + \frac{T-3}{T-2}(\sum_{t=3}^{T}y_{it-1})(y_{i1}+y_{iT})] \\
&- W_{12}\sum_{i=1}^{N}[y_{i1}y_{i2} + \sum_{t=3}^{T}y_{it}y_{it-1}] \\
&- W_{11}\sum_{i=1}^{N}[\frac{T-1}{T-2}\sum_{t=3}^{T}y_{it}y_{it-2} - \sum_{t=3}^{T}y_{it}y_{it-1} - y_{i1}y_{i2}]\Big]. \qquad (2.17)
\end{aligned}$$

Based on simple symmetric estimators and weighted symmetric estimators in (2.14)-(2.17), we construct the studentized test statistics for a unit root. In the univariate case, it is known that the studentized test statistics in higher order processes have the same limiting distribution as in the first order autoregressive process. The following theorem shows that a similar invariance holds for the sequential limit distributions of the studentized test statistics in panel data.

Theorem 2.4: Given model (2.13), under the null hypothesis $\mathrm{H}_0 : \theta_1 = 1$ and assuming that $\eta_i = y_{i0} = y_{i,-1} = 0$ for each i, then

$$i) \quad \hat{\tau}(\theta_1)_s = \frac{\hat{\theta}_{1,s} - 1 - b_s^{(2)}(1-\hat{\theta}_{2,s})}{[\hat{\sigma}_s^2 S_{22}/c_s]^{0.5}} \Longrightarrow^{D} N(0, 1.2)$$

as $T \to \infty$ is followed by $N \to \infty$,

$$ii) \quad \hat{\tau}(\theta_1)_{ws} = \frac{\hat{\theta}_{1,ws} - 1 - b_{ws}^{(2)}(1 - \hat{\theta}_{2,ws})}{[\hat{\sigma}_{ws}^2 W_{22}/c_{ws}]^{0.5}} \stackrel{D}{\Longrightarrow} N(0, 1.5)$$

as $T \to \infty$ is followed by $N \to \infty$,

where

$$b_s^{(2)} = -\frac{6(T-2)}{2(T-2)^2 + 1},$$

and

$$b_{ws}^{(2)} = -\frac{2(T-2)^2 + (T-2) + 2}{(T-2)^3 + 1}.$$

Here, $b_s^{(2)}$ and $b_{ws}^{(2)}$ can be obtained by plugging $T-2$ instead of T into b_s and b_{ws} in the AR(1) process. This idea comes from the fact that we use $2N(T-2)$ observations in the AR(2) regression table. However, this does not affect our limiting distribution results. Also, $\hat{\sigma}_s^2$ and $\hat{\sigma}_{ws}^2$ are mean squared errors in symmetric and weighted symmetric estimations, respectively.

Proof:

Before we derive the sequential limit distributions of test statistics, we consider the T-limit distributions of several random variables which are used in our test statistics. Dickey (1976) presents the T-limit distribution results in higher order processes. When $\theta_1 = 1$ and $y_{i0} = y_{i,-1} = \eta_i = 0$ in model (2.13), we have the following equation.

$$(1 - \theta_2) y_{it} = \sum_{j=1}^{t} e_{ij} - \theta_2 \Delta y_{it}.$$

From the above equation and Dickey (1976)'s results, we can obtain the following T-limit results for each i .

$$\frac{\sum_{t=3}^{T} {y_{it-1}}^2}{T^2} \stackrel{T\to\infty}{\longrightarrow} \frac{\sigma^2 \int_0^1 B_i^2(t)dt}{(1-\theta_2)^2},$$

$$\frac{\sum_{t=3}^{T} y_{it-1}}{T^{3/2}} \overset{T\to\infty}{\longrightarrow} \frac{\sigma \int_0^1 B_i(t)dt}{1-\theta_2},$$

$$\frac{y_{iT}}{T^{1/2}} \overset{T\to\infty}{\longrightarrow} \frac{\sigma B_i(1)}{1-\theta_2},$$

$$\begin{aligned}\frac{\sum_{t=3}^{T} y_{it-1}\Delta y_{it-1}}{T} &= \frac{\sum_{t=3}^{T}(\Delta y_{it-1})^2 + y_{iT-1}^2}{2T} + O_p(1/\sqrt{T}) \\ &\overset{T\to\infty}{\longrightarrow} \frac{\sigma^2}{2}\Big[\frac{1}{1-{\theta_2}^2} + \frac{B_i^2(1)}{(1-\theta_2)^2}\Big],\end{aligned}$$

and

$$\begin{aligned}\frac{\sum_{t=3}^{T} y_{it-1}\Delta y_{it}}{T} &= \frac{y_{iT}^2 - \sum_{t=2}^{T}(\Delta y_{it})^2}{2T} + O_p(1/\sqrt{T}) \\ &\overset{T\to\infty}{\longrightarrow} \frac{\sigma^2}{2}[\frac{B_i^2(1)}{(1-\theta_2)^2} - \frac{1}{1-{\theta_2}^2}].\end{aligned}$$

Proof of i):

From the above T-limit results, we can assert that

$$\frac{S_{11}}{T^2} \overset{T\to\infty}{\longrightarrow} \frac{\sigma^2}{(1-\theta_2)^2}\sum_{i=1}^{N}\Big[\int_0^1 B_i^2(t)dt - (\int_0^1 B_i(t)dt)^2\Big],$$

$$\frac{S_{12}}{T^{3/2}} \overset{T\to\infty}{\longrightarrow} 0,$$

$$\frac{S_{12}}{T} \overset{T\to\infty}{\longrightarrow} \frac{N\sigma^2}{2(1-{\theta_2}^2)},$$

and

$$\frac{S_{22}}{T} \overset{T\to\infty}{\longrightarrow} \frac{N\sigma^2}{1-{\theta_2}^2}.$$

Therefore,

$$T^{-3}c_s \overset{T\to\infty}{\longrightarrow} \frac{\sigma^4 N\sum_{i=1}^{N}[\int_0^1 B_i^2(t)dt - (\int_0^1 B_i(t)dt)^2]}{(1-\theta_2)^2(1-{\theta_2}^2)}.$$

If we look at the studentized test statistic based on the symmetric estimator,

$$\hat{\tau}(\theta_1)_s = \frac{T[\hat{\theta}_{1,s} - 1 - b_s^{(2)}(1-\hat{\theta}_{2,s})]}{[\hat{\sigma}_s^2(T^{-1}S_{22})/(T^{-3}c_s)]^{0.5}}.$$

Let us consider the T-limit distribution of $T[\hat{\theta}_{1,s} - 1 - b_s^{(2)}(1-\hat{\theta}_{2,s})]$ first.

$$\begin{aligned} T[\hat{\theta}_{1,s} - 1 - b_s^{(2)}(1-\hat{\theta}_{2,s})] &= \frac{1}{T^{-3}c_s}\Big[- \frac{S_{22}}{2(T-2)}\frac{\sum_{i=1}^{N}(\sum_{t=3}^{T} y_{it-1})(\Delta y_{iT} - \Delta y_{i2})}{T^2} \\ &- \frac{S_{12}}{T^2}\frac{\sum_{i=1}^{N}(\Delta y_{i2} - \Delta y_{iT})^2}{4(T-2)} \\ &+ \frac{S_{22}}{T}\frac{\sum_{i=1}^{N}\sum_{t=3}^{T}(\Delta y_{it})y_{it-1}}{T} \\ &+ \frac{S_{22}}{T}\frac{\sum_{i=1}^{N}(y_{i1}y_{i2} - y_{iT}y_{iT-1})}{2T} \\ &- \frac{S_{12}}{T}\frac{\sum_{i=1}^{N}\sum_{t=3}^{T}\Delta y_{it}\Delta y_{it-1}}{T} - b_s^{(2)}T\frac{c_s}{T^3}(1-\hat{\theta}_{2,s})\Big]. \end{aligned}$$

Since $\hat{\theta}_{2,s} \to \theta_2$ as $T \to \infty$,

$$T[\hat{\theta}_{1,s} - 1 - b_s^{(2)}(1-\hat{\theta}_{2,s})] \overset{T\to\infty}{\longrightarrow} \frac{(1-\theta_2)\sum_{i=1}^{N}[-\frac{1}{2} + 3(\int_0^1 B_i^2(t)dt - (\int_0^1 B_i(t)dt)^2)]}{\sum_{i=1}^{N}(\int_0^1 B_i^2(t)dt - (\int_0^1 B_i(t)dt)^2)}.$$

Also, the denominator part has the following T-limit distribution.

$$\hat{\sigma}_s^2(T^{-1}S_{22})/(T^{-3}c_s) \overset{T\to\infty}{\longrightarrow} \frac{(1-\theta_2)^2}{\sum_{i=1}^{N}(\int_0^1 B_i^2(t)dt - (\int_0^1 B_i(t)dt)^2)}.$$

Therefore, the T-limit distribution of $\hat{\tau}(\theta_1)_s$ is

$$\frac{T[\hat{\theta}_{1,s} - 1 - b_s^{(2)}(1 - \hat{\theta}_{2,s})]}{[\hat{\sigma}_s^2(T^{-1}S_{22})/(T^{-3}c_s)]^{0.5}} \xrightarrow{T\to\infty} \frac{N^{-1/2}\sum_{i=1}^{N}[-\frac{1}{2} + 3(\int_0^1 B_i^2(t)dt - (\int_0^1 B_i(t)dt)^2)]}{\sqrt{N^{-1}\sum_{i=1}^{N}(\int_0^1 B_i^2(t)dt - (\int_0^1 B_i(t)dt)^2)}}.$$

From Appendix C, we know that

$$E\Big[-\frac{1}{2} + 3\Big(\int_0^1 B_i^2(t)dt - (\int_0^1 B_i(t)dt)^2\Big)\Big] = 0,$$

$$Var\Big[-\frac{1}{2} + 3\Big(\int_0^1 B_i^2(t)dt - (\int_0^1 B_i(t)dt)^2\Big)\Big] = \frac{1}{5},$$

and

$$E\Big[\int_0^1 B_i^2(t)dt - (\int_0^1 B_i(t)dt)^2\Big] = \frac{1}{6}.$$

If we apply the Central Limit Theorem and Strong Law of Large Numbers to our T-limit distribution, we find the same limit distribution as in the AR(1) case:

$$\hat{\tau}(\theta_1)_s \xRightarrow{D} N(0, 1.2)$$

as $T \to \infty$ is followed by $N \to \infty$.

Proof of ii):

From the T-limit results in the previous pages, we find that

$$\frac{W_{11}}{T^2} \xrightarrow{T\to\infty} \frac{\sigma^2}{(1-\theta_2)^2}\sum_{i=1}^{N}\Big[\int_0^1 B_i^2(t)dt - (\int_0^1 B_i(t)dt)^2\Big],$$

$$\frac{W_{12}}{T^{3/2}} \xrightarrow{T\to\infty} 0,$$

$$\frac{W_{12}}{T} \xrightarrow{T\to\infty} \frac{N\sigma^2}{2(1-{\theta_2}^2)} + \frac{\sigma^2\sum_{i=1}^{N}[\frac{1}{2}B_i^2(1) - \int_0^1 B_i^2(t)dt + 2(\int_0^1 B_i(t)dt)^2 - (\int_0^1 B_i(t)dt)B_i(1)]}{(1-\theta_2)^2},$$

$$\frac{W_{22}}{T} \xrightarrow{T\to\infty} \frac{N\sigma^2}{1-{\theta_2}^2},$$

and

$$T^{-3}c_{ws} \xrightarrow{T\to\infty} \frac{\sigma^4 N\sum_{i=1}^{N}[\int_0^1 B_i^2(t)dt - (\int_0^1 B_i(t)dt)^2]}{(1-\theta_2)^2(1-{\theta_2}^2)}.$$

Using similar arguments to those in the proof of i), we obtain the following result.

$$T[\hat{\theta}_{1,ws} - 1 - b_{ws}^{(2)}(1-\hat{\theta}_{2,ws})] \xrightarrow{T\to\infty} \frac{(1-\theta_2)}{\sum_{i=1}^{N}(\int_0^1 B_i^2(t)dt - (\int_0^1 B_i(t)dt)^2)}\sum_{i=1}^{N}\Big[\frac{1}{2}B_i^2(1) - \frac{1}{2} + \int_0^1 B_i^2(t)dt - (\int_0^1 B_i(t)dt)B_i(1)\Big],$$

and

$$\hat{\sigma}_{ws}^2(T^{-1}W_{22})/(T^{-3}c_{ws}) \xrightarrow{T\to\infty} \frac{(1-\theta_2)^2}{\sum_{i=1}^{N}(\int_0^1 B_i^2(t)dt - (\int_0^1 B_i(t)dt)^2)}.$$

Therefore, the T-limit of $\hat{\tau}(\theta_1)_{ws}$ is

$$\hat{\tau}(\theta_1)_{ws} \xrightarrow{T\to\infty} \frac{N^{-1/2}\sum_{i=1}^{N}[\frac{1}{2}(B_i^2(1)-1) + \int_0^1 B_i^2(t)dt - (\int_0^1 B_i(t)dt)B_i(1)]}{\sqrt{N^{-1}\sum_{i=1}^{N}(\int_0^1 B_i^2(t)dt - (\int_0^1 B_i(t)dt)^2)}}.$$

If we consider the moments of the numerator and denominator as in the proof of i),

$$E\Big[\frac{1}{2}(B_i^2(1)-1) + \int_0^1 B_i^2(t)dt - (\int_0^1 B_i(t)dt)B_i(1)\Big] = 0,$$

$$Var\Big[\frac{1}{2}(B_i^2(1)-1)+\int_0^1 B_i^2(t)dt-(\int_0^1 B_i(t)dt)B_i(1)\Big]=\frac{1}{4},$$

and

$$E\Big[\int_0^1 B_i^2(t)dt-(\int_0^1 B_i(t)dt)^2\Big]=\frac{1}{6}.$$

By the Central Limit Theorem and Strong Law of Large Numbers, we again have the same sequential limit distribution as in the AR(1) case:

$$\hat{\tau}(\theta_1)_{ws}\stackrel{D}{\Longrightarrow}N(0,1.5)$$

as $T\to\infty$ is followed by $N\to\infty$.

We find that the studentized test statistics in higher order processes have the same limiting distribution as in the first order autoregressive process. Even though we have different bias adjust terms like $b_s^{(2)}(1-\hat{\theta}_{2s})$ in the test statistics, those do not affect the limiting distributions. Through Monte Carlo simulation, we compute the empirical percentiles of $\hat{\tau}(\theta_1)_s$ and $\hat{\tau}(\theta_1)_{ws}$ under the model (2.13) and the unit root null hypothesis. We consider the same sets of (N,T) as in Table 2.2. Data are generated from $y_{it}=y_{it-1}+\theta_2\Delta y_{it-1}+e_{it}$ with 10,000 replicates where the e_{it}'s are generated from N(0,1) and we set the value of θ_2 equal to 0.5. Table 2.7 and Table 2.8 show the empirical percentiles for $\hat{\tau}(\theta_1)_s$ and $\hat{\tau}(\theta_1)_{ws}$, respectively. For small samples like $T=25$ and $T=50$, there are discrepancies between the first order autoregressive process and higher order processes. These discrepancies for small samples are due in part to the fact that we apply the bias adjustment factors from the first order autoregressive process to the test statistics in higher order processes. When T=100, the empirical percentiles are close to each other.

Table 2.7: Empirical cumulative distribution of $\hat{\tau}(\theta_1)_s$

	Probability of a Smaller Value								
(N,T)	0.01	0.05	0.10	0.25	0.50	0.75	0.90	0.95	0.99
(10,25)	-3.01	-2.26	-1.85	-1.23	-0.56	0.08	0.61	0.91	1.45
(10,50)	-2.92	-2.10	-1.70	-1.06	-0.36	0.34	0.91	1.25	1.88
(10,100)	-2.73	-2.02	-1.64	-1.00	-0.26	0.46	1.09	1.44	2.03
(25,25)	-3.18	-2.42	-2.05	-1.43	-0.78	-0.15	0.40	0.74	1.33
(25,50)	-2.89	-2.19	-1.82	-1.16	-0.45	0.24	0.84	1.17	1.87
(25,100)	-2.72	-2.06	-1.64	-1.02	-0.26	0.44	1.07	1.42	2.13
(50,25)	-3.31	-2.64	-2.30	-1.72	-1.06	-0.40	0.16	0.49	1.09
(50,50)	-3.01	-2.31	-1.92	-1.28	-0.61	0.08	0.72	1.09	1.79
(50,100)	-2.85	-2.10	-1.71	-1.03	-0.31	0.40	1.02	1.38	2.03
(∞,∞)	-2.55	-1.81	-1.40	-0.73	0.00	0.73	1.40	1.81	2.55

Table 2.8: Empirical cumulative distribution of $\hat{\tau}(\theta_1)_{ws}$

	Probability of a Smaller Value								
(N,T)	0.01	0.05	0.10	0.25	0.50	0.75	0.90	0.95	0.99
(10,25)	-3.00	-2.23	-1.79	-1.06	-0.31	0.44	1.10	1.45	2.12
(10,50)	-2.97	-2.19	-1.72	-1.01	-0.21	0.59	1.27	1.67	2.42
(10,100)	-2.96	-2.13	-1.73	-1.01	-0.21	0.62	1.34	1.75	2.55
(25,25)	-3.07	-2.24	-1.83	-1.15	-0.39	0.34	1.01	1.41	2.10
(25,50)	-2.96	-2.16	-1.77	-1.05	-0.24	0.56	1.27	1.66	2.46
(25,100)	-2.85	-2.11	-1.68	-0.97	-0.16	0.66	1.37	1.82	2.64
(50,25)	-3.10	-2.34	-1.91	-1.24	-0.49	0.25	0.92	1.32	2.02
(50,50)	-3.01	-2.24	-1.79	-1.07	-0.29	0.52	1.26	1.68	2.49
(50,100)	-2.93	-2.16	-1.72	-0.96	-0.15	0.65	1.34	1.77	2.60
(∞, ∞)	-2.85	-2.02	-1.57	-0.82	0.00	0.82	1.57	2.02	2.85

Chapter 3

Maximum Likelihood Estimation with Fixed Individual Effects

3.1 Maximum Likelihood Estimator for the Fixed Individual Effects Model

In this chapter, we consider the maximum likelihood estimation method in testing for $H_0 : \rho = 1$ in the panel data model with fixed individual effects. Let y_{it} be a univariate random variable for the ith individual unit at time t. The autoregressive model of order 1 with fixed individual effects can be written as follows.

$$\begin{aligned} y_{it} &= \mu_i + \epsilon_{it} \\ \epsilon_{it} &= \rho\epsilon_{it-1} + e_{it}, \quad i = 1, 2, \cdots, N; t = 1, 2, \cdots, T, \end{aligned} \tag{3.1}$$

where the e_{it}'s are independent identically distributed as $N(0, \sigma^2)$. We assume that the first observation y_{i1} is distributed as $N(\mu_i, \frac{\sigma^2}{1-\rho^2})$ for each i. Gonzalez and Dickey (1999) present the likelihood function in a single time series, that is, for $N = 1$.

Under the model (3.1), the log-likelihood function is

$$\begin{aligned} lnL &= \frac{N}{2}ln(1-\rho^2) - \frac{NT}{2}ln\sigma^2 - \frac{NT}{2}ln2\pi - \frac{1-\rho^2}{2\sigma^2}\sum_{i=1}^{N}(y_{i1}-\mu_i)^2 \\ &\quad - \frac{1}{2\sigma^2}\sum_{i=1}^{N}\sum_{t=2}^{T}[y_{it}-\mu_i-\rho(y_{it-1}-\mu_i)]^2. \end{aligned} \tag{3.2}$$

For given $|\rho| < 1$, the individual mean estimator, $\hat{\mu}_i$ which maximizes the likelihood function (3.2) is

$$\frac{y_{i1} + y_{iT} + (1-\rho)\sum_{t=2}^{T-1} y_{it}}{2 + (T-2)(1-\rho)}, \tag{3.3}$$

This mean estimator is same as the one in a single time series (Gonzalez and Dickey, 1999). Let $\hat{\rho}_{ml,fix}$ be the maximum likelihood estimator of ρ under the model (3.1). To find the limiting distribution of $T(\hat{\rho}_{ml,fix}-1)$, we reparametrize $T(1-\rho)$ as ζ. The individual mean estimator, $\hat{\mu}_i$ can be derived from the following regression table.

Dependent Variable	Parameter μ_1	μ_2	$\cdots$	μ_N
$\sqrt{\frac{\zeta}{T}(2-\frac{\zeta}{T})}y_{11}$	$\sqrt{\frac{\zeta}{T}(2-\frac{\zeta}{T})}$	0	$\cdots$	0
$y_{12}-y_{11}+\frac{\zeta}{T}y_{11}$	$\frac{\zeta}{T}$	0	$\cdots$	0
$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$
$y_{1T}-y_{1T-1}+\frac{\zeta}{T}y_{1T-1}$	$\frac{\zeta}{T}$	0	$\cdots$	0
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\sqrt{\frac{\zeta}{T}(2-\frac{\zeta}{T})}y_{N1}$	0	0	$\cdots$	$\sqrt{\frac{\zeta}{T}(2-\frac{\zeta}{T})}$
$y_{N2}-y_{N1}+\frac{\zeta}{T}y_{N1}$	0	0	$\cdots$	$\frac{\zeta}{T}$
$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$
$y_{NT}-y_{NT-1}+\frac{\zeta}{T}y_{NT-1}$	0	0	$\cdots$	$\frac{\zeta}{T}$

To reduce the number of unknown parameters, we substitute $\hat{\mu}_i$ for μ_i in the log-likelihood (3.2). The resulting profile log-likelihood is

$$lnL = \frac{N}{2}ln\Big(\frac{\zeta}{T}(2-\frac{\zeta}{T})\Big) - \frac{NT}{2}ln\sigma^2 - \frac{NT}{2}ln2\pi - \frac{1}{2\sigma^2}SS(\zeta), \tag{3.4}$$

where $SS(\zeta)$ is the sum of squares of errors from the previous regression table, that is,

$$\begin{aligned} SS(\zeta) &= -\frac{(T-1)\zeta^3 - 2T(T-2)\zeta^2 - 3T^2\zeta}{T^2(2T-2\zeta+\zeta T)}\sum_{i=1}^{N}y_{i1}^2 + \sum_{i=1}^{N}\sum_{t=2}^{T}(y_{it}-y_{it-1})^2 \\ &\quad +2\frac{\zeta}{T}\sum_{i=1}^{N}\sum_{t=2}^{T}(y_{it}-y_{it-1})y_{it-1} + \frac{\zeta^2}{T^2}\sum_{i=1}^{N}\sum_{t=2}^{T}{y_{it-1}}^2 \\ &\quad -\frac{1}{\zeta(2-\frac{\zeta}{T})+\frac{T-1}{T}\zeta^2}\Big[\frac{\zeta^2}{T}\sum_{i=1}^{N}{y_{iT}}^2 + \frac{\zeta^4}{T^3}\sum_{i=1}^{N}(\sum_{t=2}^{T}y_{it-1})^2 \\ &\quad +2\frac{\zeta^2}{T}(1-\frac{\zeta}{T})\sum_{i=1}^{N}y_{i1}y_{iT} + 2\frac{\zeta^3}{T^2}(1-\frac{\zeta}{T})\sum_{i=1}^{N}y_{i1}(\sum_{t=2}^{T}y_{it-1}) \\ &\quad +2\frac{\zeta^3}{T^2}\sum_{i=1}^{N}y_{iT}(\sum_{t=2}^{T}y_{it-1})\Big]. \end{aligned}$$

From (3.4), we know that the maximum likelihood estimator of σ^2 is $\dfrac{SS(\zeta)}{NT}$. By plugging this estimator into (3.4), we can concentrate on finding the maximum likelihood estimator of ζ, that is, $T(1-\rho)$. Since the maximum likelihood estimators of μ_i, σ^2, and ζ are linked to each other, iterative calculation is needed until the estimators converge. The corresponding profile log-likelihood becomes

$$lnL = \frac{N}{2}ln\Big(\frac{\zeta}{T}(2-\frac{\zeta}{T})\Big) - \frac{NT}{2}ln\frac{SS(\zeta)}{NT} - \frac{NT}{2}ln2\pi - \frac{NT}{2}. \tag{3.5}$$

Let $f_N(\zeta)$ be $\dfrac{1}{N}\Big[\dfrac{\partial(-2lnL)}{\partial\zeta}\Big]\dfrac{SS(\zeta)}{NT}$.

Then

$$f_N(\zeta) = -\frac{2(1-\frac{\zeta}{T})}{\zeta(2-\frac{\zeta}{T})}\frac{SS(\zeta)}{NT} + \frac{1}{N}\frac{\partial(SS(\zeta))}{\partial\zeta}.$$

Finding the value $\hat{\zeta}$ which makes $f_N(\hat{\zeta}) = 0$ is equivalent to finding the maximum likelihood estimator of ζ. Therefore we obtain the maximum likelihood estimator of ζ by solving the equation $f_N(\zeta) = 0$, where $f_N(\zeta)$ is a polynomial function of ζ. Also, we can express $f_N(\zeta)$ as a linear function of $\frac{\sum_{i=1}^N y_{i1}^2}{N}$, $\frac{\sum_{i=1}^N\sum_{t=2}^T (y_{it}-y_{it-1})^2}{NT}$, $\frac{\sum_{i=1}^N\sum_{t=2}^T (y_{it}-y_{it-1})y_{it-1}}{NT}$, $\frac{\sum_{i=1}^N\sum_{t=2}^T {y_{it-1}}^2}{NT^2}$, $\frac{\sum_{i=1}^N {y_{iT}}^2}{NT}$, $\frac{\sum_{i=1}^N(\sum_{t=2}^T y_{it-1})^2}{NT^3}$, $\frac{\sum_{i=1}^N y_{i1}y_{iT}}{N}$, $\frac{\sum_{i=1}^N y_{i1}(\sum_{t=2}^T y_{it-1})}{NT}$, and $\frac{\sum_{i=1}^N y_{iT}(\sum_{t=2}^T y_{it-1})}{NT^2}$. That is,

$$\begin{aligned} f_N(\zeta) &= c_1\frac{\sum_{i=1}^N y_{i1}^2}{N} + c_2\frac{\sum_{i=1}^N\sum_{t=2}^T (y_{it}-y_{it-1})^2}{NT} + c_3\frac{\sum_{i=1}^N\sum_{t=2}^T (y_{it}-y_{it-1})y_{it-1}}{NT} \\ &+ c_4\frac{\sum_{i=1}^N\sum_{t=2}^T {y_{it-1}}^2}{NT^2} + c_5\frac{\sum_{i=1}^N {y_{iT}}^2}{NT} + c_6\frac{\sum_{i=1}^N(\sum_{t=2}^T y_{it-1})^2}{NT^3} \\ &+ c_7\frac{\sum_{i=1}^N y_{i1}y_{iT}}{N} + c_8\frac{\sum_{i=1}^N y_{i1}(\sum_{t=2}^T y_{it-1})}{NT} + c_9\frac{\sum_{i=1}^N y_{iT}(\sum_{t=2}^T y_{it-1})}{NT^2}, \end{aligned}$$

where the coefficients and their T-limits are

$$\begin{aligned} c_1 &= \frac{1}{T^3(-2T+\zeta)(2T-2\zeta+\zeta T)^2}\Big[-2(T-1)^2(T-2)\zeta^4 \\ &\quad +2T(3T^3-16T^2+24T-12)\zeta^3 - 2T^2(2T^3-20T^2+39T-24)\zeta^2 \\ &\quad -2(8T^5-26T^4+20T^3)\zeta - 12T^4(T-1)\Big] \\ &\overset{T\to\infty}{\longrightarrow} 0, \end{aligned}$$

$$c_2 = -\frac{2(1-\frac{\zeta}{T})}{\zeta(2-\frac{\zeta}{T})} \overset{T\to\infty}{\longrightarrow} -\frac{1}{\zeta},$$

$$c_3 = \frac{2(T-2)\zeta - 4T(T-1)}{T(\zeta - 2T)} \overset{T\to\infty}{\longrightarrow} 2,$$

$$c_4 = \frac{2(T-1)\zeta^2 + (2T-4T^2)\zeta}{T(\zeta - 2T)} \overset{T\to\infty}{\longrightarrow} 2\zeta,$$

$$c_5 = \frac{2(T-2)\zeta^2 - 4T(T-2)\zeta + 4T^2(T-1)}{(2T-2\zeta+\zeta T)^2(\zeta-2T)} \overset{T\to\infty}{\longrightarrow} -\frac{2}{(2+\zeta)^2},$$

$$\begin{aligned} c_6 &= \frac{-2(T-2)(T-1)\zeta^4 + 4T(T^2-4T+2)\zeta^3 + 4T^2(3T-1)\zeta^2}{(2T-2\zeta+\zeta T)^2(\zeta-2T)} \\ &\overset{T\to\infty}{\longrightarrow} -\frac{2\zeta^2(3+\zeta)}{(2+\zeta)^2}, \end{aligned}$$

$$c_7 = \frac{2(T-2)^2\zeta^3 - 4T(T^2-6T+6)\zeta^2 - 24T^2(T-1)\zeta + 8T^3(T-1)}{T^2(2T-2\zeta+\zeta T)^2(\zeta-2T)} \overset{T\to\infty}{\longrightarrow} 0,$$

$$\begin{aligned} c_8 &= \frac{4(T-2)(T-1)\zeta^4 - 2T(5T^2-20T+12)\zeta^3 + 4T^2(T^2-11T+6)\zeta^2}{T^2(2T-2\zeta+\zeta T)^2(\zeta-2T)} \\ &+ \frac{8T^3(2T-1)\zeta}{T^2(2T-2\zeta+\zeta T)^2(\zeta-2T)} \overset{T\to\infty}{\longrightarrow} 0, \end{aligned}$$

and

$$c_9 = \frac{-2(T-2)^2\zeta^3 + 4T(T-4)(T-1)\zeta^2 + 8T^2(2T-1)\zeta}{(2T-2\zeta+\zeta T)^2(\zeta-2T)} \overset{T\to\infty}{\longrightarrow} -\frac{2\zeta(4+\zeta)}{(2+\zeta)^2}.$$

Since $f_N(\zeta)$ is a higher order polynomial of ζ, the Gauss Newton iteration method is appropriate to find $\hat{\zeta}$. For this procedure, we need to compute the first derivative of $f_N(\zeta)$ with respect to ζ. If we define this first derivative as $f_N'(\zeta)$,

$$
\begin{aligned}
f_N'(\zeta) &= d_1\frac{\sum_{i=1}^N y_{i1}^2}{N} + d_2\frac{\sum_{i=1}^N\sum_{t=2}^T (y_{it}-y_{it-1})^2}{NT} + d_3\frac{\sum_{i=1}^N\sum_{t=2}^T (y_{it}-y_{it-1})y_{it-1}}{NT} \\
&+ d_4\frac{\sum_{i=1}^N\sum_{t=2}^T {y_{it-1}}^2}{NT^2} + d_5\frac{\sum_{i=1}^N {y_{iT}}^2}{NT} + d_6\frac{\sum_{i=1}^N(\sum_{t=2}^T y_{it-1})^2}{NT^3} + d_7\frac{\sum_{i=1}^N y_{i1}y_{iT}}{N} \\
&+ d_8\frac{\sum_{i=1}^N y_{i1}(\sum_{t=2}^T y_{it-1})}{NT} + d_9\frac{\sum_{i=1}^N y_{iT}(\sum_{t=2}^T y_{it-1})}{NT^2},
\end{aligned}
$$

where the coefficients and their T-limits are

$$
\begin{aligned}
d_1 &= \frac{1}{(2T-2\zeta+\zeta T)^3(\zeta-2T)^2T^3}\big[-2(T-1)^2(T-2)^2\zeta^5 \\
&\quad +4T(T-1)^2(2T-7)(T-2)\zeta^4 \\
&\quad -2T^2(4T^4-48T^3+161T^2-194T+80)\zeta^3 \\
&\quad -4T^3(12T^3-72T^2+111T-56)\zeta^2-4T^4(29T^2-57T+38)\zeta \\
&\quad +8T^5(2T^2-5T+5)\big] \\
&\overset{T\to\infty}{\longrightarrow} 0,
\end{aligned}
$$

$$
d_2 = \frac{2(\zeta^2-2\zeta T+2T^2)}{\zeta^2(\zeta-2T)^2} \overset{T\to\infty}{\longrightarrow} \frac{1}{\zeta^2},
$$

$$
d_3 = \frac{4}{(\zeta-2T)^2} \overset{T\to\infty}{\longrightarrow} 0,
$$

$$
d_4 = \frac{2(T-1)\zeta^2-8T(T-1)\zeta+4T^2(2T-1)}{T(\zeta-2T)^2} \overset{T\to\infty}{\longrightarrow} 2,
$$

$$
\begin{aligned}
d_5 &= \frac{-2(T-2)^2\zeta^3+4T(2T^2-7T+6)\zeta^2}{(2T-2\zeta+\zeta T)^3(\zeta-2T)^2} \\
&\quad +\frac{-4T^2(T-2)(5T-3)\zeta+8T^3(2T^2-5T+1)}{(2T-2\zeta+\zeta T)^3(\zeta-2T)^2} \\
&\overset{T\to\infty}{\longrightarrow} \frac{4}{(2+\zeta)^3},
\end{aligned}
$$

$$
\begin{aligned}
d_6 &= \frac{1}{(2T-2\zeta+\zeta T)^3(\zeta-2T)^2}\Big[-2(T-1)(T-2)^2\zeta^5 \\
&\quad +4T(T-1)(T-2)(2T-7)\zeta^4-4T^2(2T^3-21T^2+53T-30)\zeta^3 \\
&\quad -32T^4(3T-1)\zeta\Big] \\
&\xrightarrow{T\to\infty} -\frac{2(\zeta^2+6\zeta+12)\zeta}{(2+\zeta)^3},
\end{aligned}
$$

$$
\begin{aligned}
d_7 &= \frac{-8(T-2)\zeta^3+16T(T^2-3)\zeta^2-8T^2(5T^2-3T-6)\zeta+16T^3(2T^2-1)}{T(2T-2\zeta+\zeta T)^3(\zeta-2T)^2} \\
&\xrightarrow{T\to\infty} 0,
\end{aligned}
$$

$$
\begin{aligned}
d_8 &= \frac{1}{T^2(2T-2\zeta+\zeta T)^3(\zeta-2T)^2}\Big[4(T-1)(T-2)^2\zeta^5 \\
&\quad -8T(T-1)(2T-7)(T-2)\zeta^4+4T^2(4T^3-43T^2+112T-68)\zeta^3 \\
&\quad +8T^3(12T^2-61T+38)\zeta^2+16T^4(17T-10)\zeta-32T^5(2T-1)\Big] \\
&\xrightarrow{T\to\infty} 0,
\end{aligned}
$$

and

$$
\begin{aligned}
d_9 &= \frac{1}{(2T-2\zeta+\zeta T)^3(\zeta-2T)^2}\Big[-4T(T-4)(T-2)\zeta^3-8T^2(7T-12)\zeta^2 \\
&\quad +16T^3(5T-6)\zeta-32T^4(2T-1)\Big] \\
&\xrightarrow{T\to\infty} -\frac{16}{(2+\zeta)^3}.
\end{aligned}
$$

In the following section, we discuss the joint limit distribution of the test statistic based on $\hat{\zeta}$, i.e, $T(1-\hat{\rho}_{ml,fix})$ under $H_0 : \rho = 1$.

3.2 Limiting Distribution

Before we derive the joint limit distribution, let us discuss the sequential limit distribution. From the previous section,

$$f_N(\zeta) \overset{T\to\infty}{\longrightarrow} \sigma^2\Big[-\frac{1}{\zeta}+2\zeta\frac{\sum_{i=1}^N(\int_0^1 B_i^2(t)dt-(\int_0^1 B_i(t)dt)^2)}{N} + \frac{\sum_{i=1}^N(B_i^2(1)-1)}{N}-2\frac{\sum_{i=1}^N(B_i(1)\int_0^1 B_i(t)dt-(\int_0^1 B_i(t)dt)^2)}{N} - \frac{2}{(2+\zeta)^2}\frac{\sum_{i=1}^N(2\int_0^1 B_i(t)dt-B_i(1))^2}{N}\Big].$$

We find that the T-limit distribution of $f_N(\zeta)$ can be expressed as a function of averages of T-limit random quantities. In other words,

$$f_N(\zeta) \overset{T\to\infty}{\longrightarrow} \sigma^2\Big[-\frac{1}{\zeta}+2\zeta A_N+B_N-\frac{2}{(2+\zeta)^2}C_N\Big]$$

where

$$A_N=\frac{\sum_{i=1}^N(\int_0^1 B_i^2(t)dt-(\int_0^1 B_i(t)dt)^2)}{N} \overset{N\to\infty}{\longrightarrow} \frac{1}{6},$$

$$B_N=\frac{\sum_{i=1}^N(B_i^2(1)-1)}{N}-2\frac{\sum_{i=1}^N(B_i(1)\int_0^1 B_i(t)dt-(\int_0^1 B_i(t)dt)^2)}{N} \overset{N\to\infty}{\longrightarrow} -\frac{1}{3},$$

and

$$C_N=\frac{\sum_{i=1}^N(2\int_0^1 B_i(t)dt-B_i(1))^2}{N} \overset{N\to\infty}{\longrightarrow} \frac{1}{3}.$$

Let us define this T-limit of $f_N(\zeta)$ as $M(A_N,B_N,C_N,\zeta)$. If $\zeta(A_N,B_N,C_N)$ denotes a solution of ζ such that $M(A_N,B_N,C_N,\zeta)=0$, then this would be the maximum likelihood estimator of ζ in the T-limit. We expand $\zeta(A_N,B_N,C_N)$ in a Taylor series about $\zeta(\frac{1}{6},-\frac{1}{3},\frac{1}{3})$ on the surface defined by $M(A_N,B_N,C_N,\zeta)=0$.

Since $\Big[A_N-\frac{1}{6},B_N+\frac{1}{3},C_N-\frac{1}{3}\Big]=O_p(1/\sqrt{N})$,

$$\begin{aligned}\zeta(A_N, B_N, C_N) &= \zeta(\frac{1}{6}, -\frac{1}{3}, \frac{1}{3}) + \frac{\partial \zeta}{\partial A_N}(A_N - \frac{1}{6}) + \frac{\partial \zeta}{\partial B_N}(B_N + \frac{1}{3}) \\ &\quad + \frac{\partial \zeta}{\partial C_N}(C_N - \frac{1}{3}) + O_p(1/N) \\ &= 2.37 - 8.991(A_N - \frac{1}{6}) - 1.896(B_N + \frac{1}{3}) \\ &\quad + 0.199(C_N - \frac{1}{3}) + O_p(1/N).\end{aligned}$$

From the T-limit moment results in Appendix C and the Central Limit Theorem,

$$\sqrt{N}\Big[T(\hat{\rho}_{ml,fix} - 1) + 2.37\Big] \stackrel{D}{\Longrightarrow} N(0, 6.719)$$

as $T \to \infty$ is followed by $N \to \infty$.

Now, we discuss the joint limit distribution of the test statistic based on $\hat{\zeta}$ where the order of N and T does not matter. From the previous section, we know that $f_N(\zeta)$ can be expressed as a linear function of random quantities which have the form of averages across i, for example, $\frac{\sum_{i=1}^{N} y_{i1}^2}{N}$. From the moment results in Appendix A and Weak Law of Large Number theory,

$$\begin{aligned}f_N(\zeta) \xrightarrow{P} &\Big[c_1 + c_2\frac{T-1}{T} + c_4\frac{T-1}{2T} + c_5 + c_6\frac{2T^3 - 3T^2 + T}{6T^3} \\ &+ c_7 + c_8\frac{T-1}{T} + c_9\frac{T-1}{2T}\Big]\sigma^2,\end{aligned}$$

as $N \to \infty$.

Let us define this N-probability limit as $f(\zeta)$. Then $f(\zeta)$ does not have randomness anymore. If $\zeta_0(T)$ denotes the value which satisfies $f(\zeta_0(T)) = 0$, we find that $\zeta_0(T)$ depends on T. Because $-1 < \hat{\rho}_{ml,fix} < 1$ and $\hat{\zeta} = T(1 - \hat{\rho}_{ml,fix})$, we restrict the range of $\zeta_0(T)$ such that it lies within $\Big[0,\ 2\mathrm{T}\ \Big]$ for $T > 2$. For several T's, we display the values of $\zeta_0(T)$. These values are rounded off to have 4 digits.

T	$\zeta_0(T)$
25	2.505
50	2.436
100	2.402
∞	2.37

From the sequential limit result, we know that $\hat{\zeta}$-$\zeta_0(T)$ is $O_p(1/\sqrt{N})$. Therefore, if we consider the Taylor series expansion of $f_N(\zeta)$ about $\zeta_0(T)$,

$$f_N(\hat{\zeta}) = f_N(\zeta_0(T)) + f_N'(\zeta_0(T))(\hat{\zeta} - \zeta_0(T)) + O_p(\frac{1}{N}).$$

Since $f_N(\hat{\zeta}) = 0$ by definition of $\hat{\zeta}$,

$$\sqrt{N}\Big[\hat{\zeta} - \zeta_0(T)\Big] \approx -\frac{\sqrt{N} f_N(\zeta_0(T))}{f_N'(\zeta_0(T))}. \tag{3.6}$$

Now we derive the joint limit distribution of $\sqrt{N}\Big[T(\hat{\rho}_{ml,fix} - 1) + \zeta_0(T)\Big]$ by using joint probability limit theory and the joint CLT from Chapter 2.

Theorem 3.1: Given model (3.1), let us assume that $y_{i0} = 0$ for every i. Then under $H_0 : \rho = 1$,

$$\sqrt{N}\Big[T(\hat{\rho}_{ml,fix} - 1) + \zeta_0(T)\Big] \xrightarrow{D} N(0, 6.719),$$

as $N, T \to \infty$.

Proof:

Moment results used in this theorem are from Appendix A. From (3.6), we can obtain the joint limit distribution of $\sqrt{N}\Big[T(\hat{\rho}_{ml,fix} - 1) + \zeta_0(T)\Big]$ by deriving the joint limit distribution
of $\dfrac{\sqrt{N} f_N(\zeta_0(T))}{f_N'(\zeta_0(T))}$. In other words, we apply joint probability theory to $f_N'(\zeta_0(T))$ and the joint CLT to $\sqrt{N} f_N(\zeta_0(T))$.

Now

$$f_N'(\zeta) = \frac{\sum_{i=1}^{N} f_i'(\zeta)}{N},$$

where

$$\begin{aligned} f_i'(\zeta) &= d_1 y_{i1}^2 + d_2 \frac{\sum_{t=2}^{T}(y_{it} - y_{it-1})^2}{T} + d_3 \frac{\sum_{t=2}^{T}(y_{it} - y_{it-1}) y_{it-1}}{T} + d_4 \frac{\sum_{t=2}^{T} {y_{it-1}}^2}{T^2} \\ &+ d_5 \frac{{y_{iT}}^2}{T} + d_6 \frac{(\sum_{t=2}^{T} y_{it-1})^2}{T^3} + d_7 y_{i1} y_{iT} + d_8 \frac{y_{i1}(\sum_{t=2}^{T} y_{it-1})}{T} \\ &+ d_9 \frac{y_{iT}(\sum_{t=2}^{T} y_{it-1})}{T^2}. \end{aligned}$$

If we take the expectation for $f_i'(\zeta_0(T))$,

T	$E\left[f_i'(\zeta_0(T))\right]$
25	0.445
50	0.485
100	0.506
∞	0.527

In order to show that

$$f_N'(\zeta_0(T)) \xrightarrow{P} 0.527$$

as $N, T \to \infty$, we need to prove $\lim_{T\to\infty} Var\left[f_i'(\zeta_0(T))\right] < \infty$. From the variance-covariance calculation results in Appendix A, a few values of the variance are

T	$Var\left[f_i'(\zeta_0(T))\right]$
25	0.073
50	0.082
100	0.087
∞	0.092

Therefore, from Theorem 2.1 (a joint probability limit theorem),

$$f_N'(\zeta_0(T)) \xrightarrow{P} 0.527 \tag{3.7}$$

as $N, T \to \infty$.

Now, if we look at $\sqrt{N} f_N(\zeta)$,

$$\sqrt{N} f_N(\zeta) = \frac{\sum_{i=1}^{N} f_i(\zeta)}{\sqrt{N}},$$

where

$$\begin{aligned} f_i(\zeta) &= c_1 y_{i1}^2 + c_2 \frac{\sum_{t=2}^{T} (y_{it} - y_{it-1})^2}{T} + c_3 \frac{\sum_{t=2}^{T} (y_{it} - y_{it-1}) y_{it-1}}{T} + c_4 \frac{\sum_{t=2}^{T} {y_{it-1}}^2}{T^2} \\ &+ c_5 \frac{{y_{iT}}^2}{T} + c_6 \frac{(\sum_{t=2}^{T} y_{it-1})^2}{T^3} + c_7 y_{i1} y_{iT} + c_8 \frac{y_{i1} (\sum_{t=2}^{T} y_{it-1})}{T} + c_9 \frac{y_{iT} (\sum_{t=2}^{T} y_{it-1})}{T^2}. \end{aligned}$$

By the definition of $\zeta_0(T)$, we know that $E\Big[f_i(\zeta_0(T))\Big] = 0$. The variances of $f_i(\zeta_0(T))$ for several T's are

T	$Var\Big[f_i(\zeta_0(T))\Big]$
25	1.355
50	1.595
100	1.727
∞	1.866

Now, we need to prove the uniformly integrability of $f_i^2(\zeta_0(T))$ in T. If we consider the T-limit distribution of $f_i^2(\zeta_0(T))$,

$$\begin{aligned} f_i^2(\zeta_0(T)) \xrightarrow{D} \Big[&- \frac{1}{\zeta_0(\infty)} + 2\zeta_0(\infty) \Big(\int_0^1 B_i^2(t)dt - \big(\int_0^1 B_i(t)dt \big)^2 \Big) + (B_i^2(1) - 1) \\ &- 2\Big(B_i(1) \int_0^1 B_i(t)dt - \big(\int_0^1 B_i(t)dt \big)^2 \Big) - 2 \frac{(2\int_0^1 B_i(t)dt - B_i(1))^2}{(\zeta_0(\infty) + 2)^2} \Big]^2 \sigma^4, \end{aligned}$$

as $T \to \infty$, where, as previously stated, $\zeta_0(\infty)$=2.37.

From Appendix C, the expectation of the right hand side in the above equation is 1.866. Since this value is equal to $\lim_{T\to\infty} E\Big[f_i^2(\zeta_0(T))\Big]$, $f_i^2(\zeta_0(T))$ is uniformly integrable in T. If we apply Slutsky's theorem to (3.6) using the result (3.7),

$$\sqrt{N}\Big[T(\hat{\rho}_{ml,fix}-1)+\zeta_0(T)\Big] \stackrel{D}{\Longrightarrow} N(0, 6.719),$$

as $N, T \to \infty$.

We conduct Monte Carlo simulation to obtain the empirical percentiles of $\sqrt{N}\Big[T(\hat{\rho}_{ml,fix}-1)+\zeta_0(T)\Big]$. These percentiles are shown in Table 3.1. We use the same sets of (N, T) as in Chapter 2. For each combination of (N, T), we generate y_{it} such that $y_{it} = y_{it-1} + e_{it}$, where the e_{it}'s are generated using the RANNOR function in SAS and $y_{i0} = 0$ for each i. We have 10,000 replicates. For the computation of $\hat{\rho}_{ml,fix}$, we run PROC MIXED in SAS, since PROC MIXED has the likelihood function for the autoregressive model of order 1. We use the same method as in Chapter 2 for smoothing the 5 % empirical percentiles in Table 3.1. The coefficient of determination is over 0.99 and the square root of the mean squared error is 0.08. Therefore, we can approximate the 5 % critical values with a linear function of $\frac{1}{\sqrt{N}}$ and $\frac{1}{\sqrt{T}}$, namely

$$\text{critical value} \approx -4.28 - 5.39\frac{1}{\sqrt{N}} + 0.37\frac{1}{\sqrt{T}}.$$

Table 3.1: Empirical cumulative distribution of $\sqrt{N}\Big[T(\hat{\rho}_{ml,fix} - 1) + \zeta_0(T)\Big]$

	Probability of a Smaller Value								
(N,T)	0.01	0.05	0.10	0.25	0.50	0.75	0.90	0.95	0.99
(10,25)	-8.87	-5.99	-4.54	-2.36	-0.33	1.31	2.55	3.11	4.16
(10,50)	-9.22	-5.96	-4.54	-2.32	-0.36	1.28	2.48	3.09	4.03
(10,100)	-9.29	-5.97	-4.55	-2.36	-0.37	1.24	2.44	3.04	3.93
(25,25)	-7.78	-5.33	-4.08	-2.13	-0.18	1.41	2.72	3.52	4.77
(25,50)	-7.97	-5.27	-4.07	-2.12	-0.17	1.48	2.78	3.48	4.78
(25,100)	-7.79	-5.19	-3.99	-2.09	-0.24	1.41	2.66	3.41	4.61
(50,25)	-7.32	-4.99	-3.85	-2.02	-0.18	1.59	3.00	3.71	5.13
(50,50)	-7.24	-4.91	-3.78	-1.94	-0.17	1.52	2.94	3.72	5.06
(50,100)	-7.34	-4.93	-3.73	-1.96	-0.19	1.50	2.92	3.69	5.04
(∞,∞)	-6.04	-4.28	-3.32	-1.74	0.00	1.74	3.32	4.28	6.04

Chapter 4

Maximum Likelihood Estimation with Random Individual Effects

4.1 Maximum Likelihood Estimator for the Random Individual Effects Model

In this chapter, we consider maximum likelihood estimation and resulting test for a unit root in a panel data model with random individual effects. The stationary autoregressive model of order 1 with random individual effects can be described as follows.

$$\begin{aligned} y_{it} &= \mu + \alpha_i + \epsilon_{it}, \\ \epsilon_{it} &= \rho\epsilon_{it-1} + e_{it}, \quad i = 1, 2, \cdots, N; t = 1, 2, \cdots, T, \end{aligned} \tag{4.1}$$

where the e_{it}'s are independent identically distributed as $N(0, \sigma^2)$, the α_i's are independent identically distributed as $N(0, \lambda\sigma^2)$, and λ is an arbitrary positive real number. We assume that the first observation y_{i1} is distributed as $N(\mu, \lambda\sigma^2 + \frac{\sigma^2}{1-\rho^2})$. Also, ϵ_{it} and α_i are assumed to be independent.

For $|\rho| < 1$, let the covariance matrix of the ith individual variables be $\sigma^2\mathbf{V}$. Then $\sigma^2\mathbf{V}$ is a $(T \times T)$ matrix which has the form $\sigma^2\Big[\mathbf{R} + \lambda\mathbf{J}\Big]$, where $\mathbf{J}$ is a $(T \times T)$

matrix all of whose elements are one and $\mathbf{R}$ is a toeplitz matrix whose first row is $\frac{1}{1-\rho^2}\Big[1, \rho, \rho^2, \cdots, \rho^{T-1}\Big]$.

Then the exact likelihood function for the model (4.1) is

$$L(\lambda, \sigma^2, \mu, \rho) = \frac{1}{(2\pi\sigma^2)^{\frac{NT}{2}}|\mathbf{V}|^{\frac{N}{2}}} exp\Big[-\sum_{i=1}^{N}\frac{\mathbf{Y}'_{i,\mu}\mathbf{V}^{-1}\mathbf{Y}_{i,\mu}}{2\sigma^2}\Big], \tag{4.2}$$

where

$$\mathbf{Y}_{i,\mu} = \Big[y_{i1} - \mu, y_{i2} - \mu, \cdots, y_{iT} - \mu\Big]'.$$

The maximum likelihood estimator of the overall mean parameter, $\hat{\mu}_{mle}$ can be obtained by the Generalized Least Squares(GLS) method. The GLS estimator for μ is

$$\hat{\mu} = \Big[\sum_{i=1}^{N}\mathbf{1}'\mathbf{V}^{-1}\mathbf{1}\Big]^{-1}\Big[\sum_{i=1}^{N}\mathbf{1}'\mathbf{V}^{-1}\mathbf{Y}_i\Big],$$

where $\mathbf{1}$ is a column vector of 1's with dimension T and

$$\mathbf{Y}_i = \Big[y_{i1}, y_{i2}, \cdots, y_{iT}\Big]'.$$

The inverse covariance matrix, $[\sigma^2\mathbf{V}]^{-1}$ can be expressed as $\frac{1}{\sigma^2}\Big[\mathbf{R}^{-1} - d\mathbf{L}\mathbf{L}'\Big]$ where

$$\mathbf{R}^{-1} = \begin{bmatrix} 1 & -\rho & 0 & 0 & \cdots & \cdots & 0 \\ -\rho & \rho^2+1 & -\rho & 0 & \vdots & \cdots & \vdots \\ 0 & -\rho & \rho^2+1 & -\rho & \vdots & \cdots & \vdots \\ 0 & 0 & \ddots & \ddots & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \cdots & -\rho & \rho^2+1 & -\rho \\ 0 & \cdots & \cdots & \cdots & 0 & -\rho & 1 \end{bmatrix},$$

$$d = \frac{\lambda(1-\rho)^2}{[(T-2)\lambda\rho^2 - 2(T-1)\lambda\rho + 1 + T\lambda]},$$

and

$$\mathbf{L'} = [1, 1-\rho, 1-\rho, \cdots, 1-\rho, 1].$$

For given $|\rho| < 1$, the maximum likelihood estimator of μ is

$$\begin{aligned} \hat{\mu}_{mle} &= \Big[\sum_{i=1}^{N} \mathbf{1'V^{-1}1}\Big]^{-1}\Big[\sum_{i=1}^{N} \mathbf{1'V^{-1}Y}_i\Big] \\ &= \frac{\sum_{i=1}^{N}\Big[y_{i1} + y_{iT} + (1-\rho)\sum_{t=2}^{T-1} y_{it}\Big]}{N\Big[2 + (T-2)(1-\rho)\Big]}. \end{aligned} \tag{4.3}$$

This has the form of the average of single subject mean estimators in Gonzalez (1999).

Let $\hat{\rho}_{ml,ran}$ denote the maximum likelihood estimator(MLE) of ρ in the random effects model. Since we are interested in finding the limit distribution of $T(\hat{\rho}_{ml,ran} - 1)$, we write $T(1-\rho) = \zeta$ as in Chapter 3. If we consider the determinant of $\mathbf{V}$,

$$|\mathbf{V}| = \frac{(T-2)\lambda\frac{\zeta^2}{T^2} + 2\lambda\frac{\zeta}{T} + 1}{\frac{\zeta}{T}\Big(2 - \frac{\zeta}{T}\Big)}.$$

Inserting the mean estimator, $\hat{\mu}_{mle}$ and $|\mathbf{V}|$, the profile log likelihood is

$$\begin{aligned} lnL(\lambda, \sigma^2, \zeta) &= -\frac{NT}{2}ln2\pi - \frac{N}{2}ln\Big[(T-2)\lambda\frac{\zeta^2}{T^2} + 2\lambda\frac{\zeta}{T} + 1\Big] \\ &\quad + \frac{N}{2}ln\Big(\frac{\zeta}{T}\Big(2 - \frac{\zeta}{T}\Big)\Big) - \frac{NT}{2}ln\sigma^2 - \frac{SS(\lambda,\zeta)}{2\sigma^2}, \end{aligned} \tag{4.4}$$

where

$$SS(\lambda,\zeta) = \sum_{i=1}^{N} \mathbf{Y}_i'\mathbf{V^{-1}Y}_i - \frac{(\sum_{i=1}^{N} \mathbf{1'V^{-1}Y}_i)^2}{\sum_{i=1}^{N} \mathbf{1'V^{-1}1}}$$

$$
\begin{aligned}
= & -\Big(1-\frac{\zeta}{T}\Big)^2\frac{\lambda\zeta^2(T-1)+2\lambda\zeta T+T^2}{den1}\sum_{i=1}^{N}y_{i1}^2 \\
- & \ 2\Big(1-\frac{\zeta}{T}\Big)\sum_{i=1}^{N}\sum_{t=2}^{T}(y_{it}-y_{it-1})y_{it-1}+\frac{\zeta^2}{T^2}\sum_{i=1}^{N}\sum_{t=2}^{T}{y_{it-1}}^2 \\
+ & \ \frac{\lambda\zeta^2(T-3)+2\lambda\zeta T+T^2}{den1}\sum_{i=1}^{N}{y_{iT}}^2-\frac{\lambda\zeta^4}{T^2den1}\sum_{i=1}^{N}(\sum_{t=2}^{T}y_{it-1})^2 \\
- & \ \frac{2\lambda(T-\zeta)\zeta^3}{T^2den1}\sum_{i=1}^{N}y_{i1}(\sum_{t=2}^{T}y_{it-1})-\frac{2\lambda(T-\zeta)\zeta^2}{Tden1}\sum_{i=1}^{N}y_{i1}y_{iT} \\
- & \ \frac{2\lambda\zeta^3}{Tden1}\sum_{i=1}^{N}y_{iT}(\sum_{t=2}^{T}y_{it-1})-\frac{\Big(1-\frac{\zeta}{T}\Big)^2\frac{\zeta^2}{T^2}}{den2}\frac{(\sum_{i=1}^{N}y_{i1})^2}{N} \\
- & \ \frac{\zeta^4}{T^4den2}\frac{(\sum_{i=1}^{N}\sum_{t=2}^{T}y_{it-1})^2}{N}-\frac{\zeta^2}{T^2den2}\frac{(\sum_{i=1}^{N}y_{iT})^2}{N} \\
- & \ \frac{2\Big(1-\frac{\zeta}{T}\Big)\frac{\zeta^3}{T^3}}{den2}\frac{(\sum_{i=1}^{N}y_{i1})(\sum_{i=1}^{N}\sum_{t=2}^{T}y_{it-1})}{N} \\
- & \ \frac{2\Big(1-\frac{\zeta}{T}\Big)\frac{\zeta^2}{T^2}}{den2}\frac{(\sum_{i=1}^{N}y_{i1})(\sum_{i=1}^{N}y_{iT})}{N}-\frac{2\frac{\zeta^3}{T^3}}{den2}\frac{(\sum_{i=1}^{N}y_{iT})(\sum_{i=1}^{N}\sum_{t=2}^{T}y_{it-1})}{den2},
\end{aligned}
$$

$$
den1=\lambda\zeta^2(T-2)+2\lambda\zeta T+T^2,
$$

and

$$
den2=\Big[(T-2)\lambda\frac{\zeta^2}{T^2}+2\lambda\frac{\zeta}{T}+1\Big]\Big[(T-2)\frac{\zeta^2}{T^2}+2\frac{\zeta}{T}\Big].
$$

Also, we see that $\dfrac{SS(\lambda,\zeta)}{NT}$ is the maximum likelihood estimator of σ^2. If we plug this estimator into the log-likelihood function (4.4), then we have the following profile log-likelihood function which is a function of only ζ and λ.

$$
lnL(\lambda,\zeta) \ = \ -\frac{NT}{2}ln2\pi-\frac{N}{2}ln\Big[(T-2)\lambda\frac{\zeta^2}{T^2}+2\lambda\frac{\zeta}{T}+1\Big]
$$

$$+\frac{N}{2}ln\Big(\frac{\zeta}{T}\Big(2-\frac{\zeta}{T}\Big)\Big)-\frac{NT}{2}ln\frac{SS(\lambda,\zeta)}{NT}-\frac{NT}{2}. \tag{4.5}$$

To find the maximum likelihood estimators of ζ and λ, we consider an iterative Gauss Newton method. For this procedure, we look at a vector of first derivatives and Hessian matrix. For convenience, we set $\big[\frac{\partial(-2lnL)}{\partial\zeta},\frac{\partial(-2lnL)}{\partial\lambda}\big]'$ as our objective function vector. Then

$$\frac{\partial(-2lnL)}{\partial\zeta}=N\frac{2\lambda(\zeta(T-2)+T)}{den1}+N\frac{2(\zeta-T)}{\zeta(2T-\zeta)}+\big[\frac{SS(\lambda,\zeta)}{NT}\big]^{-1}\frac{\partial(SS(\lambda,\zeta))}{\partial\zeta},$$

$$\frac{\partial(-2lnL)}{\partial\lambda}=N\frac{\zeta(\zeta(T-2)+2T)}{den1}+\big[\frac{SS(\lambda,\zeta)}{NT}\big]^{-1}\frac{\partial(SS(\lambda,\zeta))}{\partial\lambda},$$

$$\begin{aligned}\frac{\partial^2(-2lnL)}{\partial\zeta^2} &= N\frac{-2\lambda(\lambda\zeta^2(T-2)^2+2\lambda\zeta T(T-2)-T^2(T-2)+2\lambda T^2)}{den1^2}\\ &+ 2N\frac{\zeta^2-2\zeta T+2T^2}{\zeta^2(2T-\zeta)^2}+NT\frac{\frac{\partial^2(SS(\lambda,\zeta))}{\partial\zeta^2}SS(\lambda,\zeta)-\big[\frac{\partial(SS(\lambda,\zeta))}{\partial\zeta}\big]^2}{(SS(\lambda,\zeta))^2},\end{aligned}$$

$$\begin{aligned}\frac{\partial^2(-2lnL)}{\partial\lambda^2} &= N\frac{-\zeta^2(\zeta(T-2)+2T)^2}{den1^2}\\ &+ NT\frac{\frac{\partial^2(SS(\lambda,\zeta))}{\partial\lambda^2}SS(\lambda,\zeta)-\big[\frac{\partial(SS(\lambda,\zeta))}{\partial\lambda}\big]^2}{(SS(\lambda,\zeta))^2},\end{aligned}$$

and

$$\begin{aligned}\frac{\partial^2(-2lnL)}{\partial\lambda\partial\zeta} &= N\frac{2(\zeta(T-2)+T)T^2}{den1^2}\\ &+ NT\frac{\frac{\partial^2(SS(\lambda,\zeta))}{\partial\lambda\partial\zeta}SS(\lambda,\zeta)-\frac{\partial(SS(\lambda,\zeta))}{\partial\zeta}\frac{\partial(SS(\lambda,\zeta))}{\partial\lambda}}{(SS(\lambda,\zeta))^2}.\end{aligned}$$

Let $\hat{\theta}^{(j)}$ be the estimated values of $\left[\zeta^{(j)}, \lambda^{(j)}\right]'$ at the jth step of the Gauss Newton algorithm. Then to find maximum likelihood estimators, we iterate the following system for ζ and λ until convergence occurs.

$$\hat{\theta}^{(j+1)} = \hat{\theta}^{(j)} - \mathbf{H}^{-1}\mathbf{g},$$

where

$$\mathbf{H}= \begin{bmatrix} \dfrac{\partial^2(-2lnL)}{\partial\zeta^2} & \dfrac{\partial^2(-2lnL)}{\partial\lambda\partial\zeta} \\ \dfrac{\partial^2(-2lnL)}{\partial\lambda\partial\zeta} & \dfrac{\partial^2(-2lnL)}{\partial\lambda^2} \end{bmatrix},$$

and

$$\mathbf{g}= \begin{bmatrix} \dfrac{\partial(-2lnL)}{\partial\zeta} \\ \dfrac{\partial(-2lnL)}{\partial\lambda} \end{bmatrix}.$$

4.2 Sequential Limit Distribution

In this section, we discuss the sequential limit distribution of the test statistic based on $\hat{\zeta}$, i.e, $T(1-\hat{\rho}_{ml,ran})$ under $H_0 : \rho = 1$ as $T \to \infty$ and then $N \to \infty$. A difficulty arises in finding the joint limit distribution when we estimate λ and ζ at the same time. The problem comes from the fact that the determinant of the Hessian matrix is close to zero as $T \to \infty$. Therefore, we remove the effect of λ by taking the T-limit first. To derive the sequential limit distribution, we consider the T-limit distribution for finite fixed N.

We can express $SS(\lambda, \zeta)$ as

$$SS(\lambda,\zeta) = \sum_{i=1}^{N}\mathbf{Y}_i'\mathbf{V}^{-1}\mathbf{Y}_i - (\sqrt{T}\sum_{i=1}^{N}\mathbf{1}'\mathbf{V}^{-1}\mathbf{Y}_i)(T\sum_{i=1}^{N}\mathbf{1}'\mathbf{V}^{-1}\mathbf{1})^{-1}(\sqrt{T}\sum_{i=1}^{N}\mathbf{1}'\mathbf{V}^{-1}\mathbf{Y}_i).$$

We seek a limiting distribution for $SS(\lambda,\zeta)+2\sum_{i=1}^{N}\sum_{t=2}^{T}(y_{it}-y_{it-1})y_{it-1}-\sum_{i=1}^{N}{y_{iT}}^2$. Using the formula for $SS(\lambda,\zeta)$ following (4.4), we find

$$\begin{aligned}\sum_{i=1}^{N}\mathbf{Y}_i'\mathbf{V}^{-1}\mathbf{Y}_i \quad &+ \quad 2\sum_{i=1}^{N}\sum_{t=2}^{T}(y_{it}-y_{it-1})y_{it-1}-\sum_{i=1}^{N}{y_{iT}}^2\\ &\xrightarrow{T\to\infty}\\ -\sum_{i=1}^{N}y_{i1}^2 \quad &+ \quad \sum_{i=1}^{N}\Big[\zeta(B_i^2(1)-1)+\zeta^2\int_0^1 B_i^2(t)dt\Big]\sigma^2,\end{aligned}$$

$$T\sum_{i=1}^{N}\mathbf{1}'\mathbf{V}^{-1}\mathbf{1}\xrightarrow{T\to\infty}N(2\zeta+\zeta^2),$$

and

$$\sqrt{T}\sum_{i=1}^{N}\mathbf{1}'\mathbf{V}^{-1}\mathbf{Y}_i\xrightarrow{T\to\infty}\sum_{i=1}^{N}\Big[\zeta^2\int_0^1 B_i(t)dt+\zeta B_i(1)\Big]\sigma,$$

where $B_i(t), i=1,2,\cdots,N$ are independent standard multivariate Wiener processes. If we look at the T-limit distribution of $SS(\lambda,\zeta)+2\sum_{i=1}^{N}\sum_{t=2}^{T}(y_{it}-y_{it-1})y_{it-1}-\sum_{i=1}^{N}{y_{iT}}^2$, we have

$$\begin{aligned}SS(\lambda,\zeta)+2\sum_{i=1}^{N}\sum_{t=2}^{T}(y_{it}-y_{it-1})y_{it-1}-\sum_{i=1}^{N}{y_{iT}}^2 &\xrightarrow{T\to\infty} -\sum_{i=1}^{N}y_{i1}^2+\zeta\sigma^2\sum_{i=1}^{N}(B_i^2(1)-1)\\ &+\zeta^2\sigma^2\sum_{i=1}^{N}(\int_0^1 B_i^2(t)dt)-\frac{\sigma^2\zeta}{N(2+\zeta)}\Big[\sum_{i=1}^{N}(\zeta\int_0^1 B_i(t)dt+B_i(1))\Big]^2. \quad (4.6)\end{aligned}$$

Note that

$$\begin{aligned}\frac{SS(\lambda,\zeta)}{NT} &= \frac{\sum_{i=1}^{N}{y_{iT}}^2-2\sum_{t=2}^{T}(y_{it}-y_{it-1})y_{it-1}}{NT}+O_p(\frac{1}{\sqrt{N}T})\\ &= \frac{\sum_{i=1}^{N}\sum_{t=2}^{T}(y_{it}-y_{it-1})^2}{NT}+O_p(\frac{1}{\sqrt{N}T})\\ &\xrightarrow{p} \sigma^2.\end{aligned}$$

Then the T-limit distribution of $\dfrac{\partial(-2lnL)}{\partial\zeta}$ is

$$\begin{aligned}\frac{\partial(-2lnL)}{\partial\zeta} \quad \stackrel{T\to\infty}{\longrightarrow} \quad & -\frac{N}{\zeta}+\sum_{i=1}^{N}\Big[(B_i^2(1)-1)+2\zeta\int_0^1 B_i^2(t)dt\Big] \\ & -\frac{2}{(2+\zeta)^2N}\Big[\zeta^2(3+\zeta)(\sum_{i=1}^{N}\int_0^1 B_i(t)dt)^2 \\ & +\zeta(4+\zeta)(\sum_{i=1}^{N}\int_0^1 B_i(t)dt)(\sum_{i=1}^{N}B_i(1))+(\sum_{i=1}^{N}B_i(1))^2\Big]. \qquad (4.7)\end{aligned}$$

Thus we can see that the role of λ disappears in the T-limit. Furthermore, this T-limit distribution is same as in the fixed common mean effect model.

Now, based on this T-limit distribution, we consider a sequential limit distribution by taking the N-limit. In deriving the sequential limit, we use the same technique as in Chapter 3. Let us define the T-limit distribution of $N^{-1}\dfrac{\partial(-2lnL)}{\partial\zeta}$ as $g(D_N, E_N, F_N, I_N, \zeta)$. Then

$$\begin{aligned}g(D_N,E_N,F_N,I_N,\zeta) \quad = \quad & -\frac{1}{\zeta}+D_N+2\zeta E_N-\frac{2\zeta^2(3+\zeta)}{(2+\zeta)^2}F_N^2 \\ & -\frac{2\zeta(4+\zeta)}{(2+\zeta)^2}F_NI_N-\frac{2}{(2+\zeta)^2}I_N^2, \qquad (4.8)\end{aligned}$$

where

$$D_N=\frac{\sum_{i=1}^{N}(B_i^2(1)-1)}{N}\stackrel{N\to\infty}{\longrightarrow}0,$$

$$E_N=\frac{\sum_{i=1}^{N}(\int_0^1 B_i^2(t)dt)}{N}\stackrel{N\to\infty}{\longrightarrow}\frac{1}{2},$$

$$F_N=\frac{\sum_{i=1}^{N}(\int_0^1 B_i(t)dt)}{N}\stackrel{N\to\infty}{\longrightarrow}0,$$

and

$$I_N = \frac{\sum_{i=1}^{N}(B_i(1))}{N} \overset{N\to\infty}{\longrightarrow} 0.$$

We are interested in the surface, $g(D_N, E_N, F_N, I_N, \zeta) = 0$ to find the maximum likelihood estimator of ζ. If $\zeta(D_N, E_N, F_N, I_N)$ denotes a solution of $g(D_N, E_N, F_N, I_N, \zeta) = 0$, then $\zeta(0, \frac{1}{2}, 0, 0) = 1$. Using Taylor's series,

$$\begin{aligned}\zeta(D_N, E_N, F_N, I_N) &= \zeta(0, \frac{1}{2}, 0, 0) + \frac{\partial\zeta}{\partial D_N}D_N + \frac{\partial\zeta}{\partial E_N}(E_N - \frac{1}{2}) \\ &\quad + \frac{\partial\zeta}{\partial F_N}F_N + \frac{\partial\zeta}{\partial I_N}I_N + O_p(1/N) \\ &= 1 - \frac{1}{2}D_N - (E_N - \frac{1}{2}) + O_p(1/N), \qquad (4.9)\end{aligned}$$

where each partial derivative of $\zeta(D_N, E_N, F_N, I_N)$ is evaluated at $(\zeta = 1, D_N = 0, E_N = \frac{1}{2}, F_N = 0, I_N = 0)$. Then we have the following sequential limit result for $\sqrt{N}\big[\hat{\zeta} - 1\big]$.

Theorem 4.1: Given model (4.1), if we assume that $\mu = \epsilon_{i0} = 0$ for each i, then under $H_0 : \rho = 1$,

$$\sqrt{N}\Big[T(\hat{\rho}_{ml,ran} - 1) + 1\Big] \overset{D}{\Longrightarrow} N(0, 1.5)$$

as $T \to \infty$ is followed by $N \to \infty$.

Proof:

From (4.9), we know that

$$\sqrt{N}\big[\hat{\zeta} - 1\big] \overset{T\to\infty}{\longrightarrow} \frac{\sum_{i=1}^{N}\big[-\frac{1}{2}(B_i^2(1) - 1) - (\int_0^1 B_i^2(t)dt - \frac{1}{2})\big]}{\sqrt{N}} + O_p(1/\sqrt{N}).$$

If we look at the moment results of the T-limit random variables in Appendix C,

$$E\Big[-\frac{1}{2}(B_i^2(1)-1)-(\int_0^1 B_i^2(t)dt-\frac{1}{2})\Big]=0,$$

and

$$Var\Big[-\frac{1}{2}(B_i^2(1)-1)-(\int_0^1 B_i^2(t)dt-\frac{1}{2})\Big]=1.5.$$

Therefore, by the Central Limit Theorem,

$$\sqrt{N}\Big[\hat{\zeta}-1\Big]\stackrel{D}{\Longrightarrow}N(0,1.5)$$

as $T\to\infty$ is followed by $N\to\infty$.

Table 4.1 shows the empirical percentiles of $\sqrt{N}\Big[T(\hat{\rho}_{ml,ran}-1)+1\Big]$ when $\lambda=1$ under $H_0:\rho=1$. For this Monte Carlo simulation, we generate y_{it} such that $y_{it}=y_{it-1}+e_{it}$ where the e_{it}'s are generated from the RANNOR function in SAS. The initial observation y_{i0} is generated from $N(0,1)$ independently of e_{it}. We have 10,000 replicates for each combination of (N,T). Since SAS PROC MIXED implements the maximum likelihood algorithm for model (4.1), we can obtain $\hat{\rho}_{ml,ran}$ from PROC MIXED. We also present means of $T(\hat{\rho}_{ml,ran}-1)$ and variances of $\sqrt{N}T(\hat{\rho}_{ml,ran}-1)$ for each combination of (N,T) below the table of the empirical percentiles. As $T\to\infty$ is followed by $N\to\infty$, we find that means for $T(\hat{\rho}_{ml,ran}-1)$ become close to -1 and variances of $\sqrt{N}\Big[T(\hat{\rho}_{ml,ran}-1)+1\Big]$ converge to 1.5. This is consistent with our sequential limit result in Theorem 4.1. As in the previous chapter, we smooth the 5 % empirical percentiles in Table 4.1 by regressing on $\frac{1}{\sqrt{N}}$ and $\frac{1}{\sqrt{T}}$. For this regression, we use values after subtracting sequential limit 5 % percentiles, that is, -2.02 from the empirical percentiles. Therefore we do not include an intercept in the regression model. The adjusted R^2 is over 0.93 and the square root of mean squared error is 0.24. Hence the critical values can be explained well with a linear function of $\frac{1}{\sqrt{N}}$ and $\frac{1}{\sqrt{T}}$, that is, $-2.02-8.45\frac{1}{\sqrt{N}}+8.17\frac{1}{\sqrt{T}}$.

Table 4.1: Empirical cumulative distribution of $\sqrt{N}\left[T(\hat{\rho}_{ml,ran}-1)+1\right]$

	Probability of a Smaller Value								
(N,T)	0.01	0.05	0.10	0.25	0.50	0.75	0.90	0.95	0.99
(10,25)	-6.29	-3.54	-2.37	-1.01	0.00	0.70	1.17	1.42	1.77
(10,50)	-6.42	-3.67	-2.56	-1.23	-0.22	0.57	1.08	1.36	1.71
(10,100)	-6.10	-3.64	-2.66	-1.36	-0.28	0.51	1.04	1.30	1.71
(25,25)	-3.60	-2.05	-1.40	-0.50	0.30	1.00	1.55	1.80	2.25
(25,50)	-3.85	-2.45	-1.80	-0.85	0.05	0.80	1.35	1.65	2.10
(25,100)	-4.20	-2.70	-2.00	-1.00	-0.05	0.70	1.25	1.55	2.05
(50,25)	-2.47	-1.48	-0.92	-0.14	0.64	1.27	1.84	2.12	2.62
(50,50)	-3.11	-1.98	-1.41	-0.57	0.21	0.99	1.56	1.91	2.40
(50,100)	-3.39	-2.19	-1.63	-0.78	0.07	0.85	1.41	1.77	2.33
(∞,∞)	-2.85	-2.02	-1.57	-0.82	0.00	0.82	1.57	2.02	2.85

Mean of $T(\hat{\rho}_{ml,ran}-1)$			
(N,T)	25	50	100
10	-1.12	-1.17	-1.19
25	-0.97	-1.02	-1.05
50	-0.93	-0.98	-1.01

Variance of $\sqrt{N}\left[T(\hat{\rho}_{ml,ran}-1)+1\right]$			
(N,T)	25	50	100
10	2.7	2.7	2.6
25	1.5	1.75	1.75
50	1	1.5	1.5

Chapter 5

Power Study

In this chapter, we compare the powers of several unit root tests in panel data. The observations are generated by

$$y_{it} = \mu_i(1-\rho) + \rho y_{it-1} + e_{it}; \quad i = 1, \cdots, N, t = 1, \cdots, T$$

where μ_i and e_{it} are generated from the standard Gaussian distribution independently, and $y_{i0} = \mu_i$ for each i. As a random number generator, we use the RANNOR function in SAS. We consider 5 values for ρ, that is, 1.0, 0.98, 0.95, 0.90, and 0.85. Note that when ρ =1, we obtain the size for each test statistic. For our power study, we consider the following 8 sets of panel and time dimensions.

$$\begin{aligned}(N,T) &= (10,25),(10,50),(10,100),(25,25),(25,50),(25,100),\\ &\quad (50,25), \text{and } (50,50).\end{aligned}$$

For each set, 2,000 data sets are generated and we compute powers of the follwing test statistics.

LL=Levin and Lin's t-statistic in (1.8).
IPS=Im, Peseran, and Shin's t-bar statistic in (1.21).
$\hat{\tau}_s$=symmetric studentized test statistic in (2.7).

$\hat{\tau}_{ws}$=weighted symmetric studentized test statistic in (2.8).
MLEf=$\sqrt{N}\Big[T(\hat{\rho}_{ml,fix}-1)+\zeta_0(T)\Big]$ in chapter 3.
MLEr=$\sqrt{N}\Big[T(\hat{\rho}_{ml,ran}-1)+1\Big]$ in chapter 4.

As the 5 % critical values for our test statistics, we use the empirical percentiles in chapter 2,3, and 4. For Levin and Lin's (1992) t-statistic and Im, Peseran, and Shin's (1997) t-bar statistic, we use their 5 % empirical percentiles under $H_0 : \rho = 1$.

Concerning size, most of the statistics have proper sizes which are close to the nominal value, 0.05, whereas Levin and Lin's t-statstic has a size distortion problem as N becomes large. This is due to the fact that the critical values for Levin and Lin's t-statistic become much smaller for large N. As Karlsson and Löthgren (2000) mention in their paper, powers increase dramatically when T becomes large rather than when N becomes large. Especially, when $\rho < 0.98$, powers for test statistics based on the weighted symmetric estimator and the maximum likelihood estimator are much higher than LL and IPS test statistics even for small samples like T=25. Also, when we compare the power of our panel test statistics with those in Gonzalez (1999) that is, the single time series case, we find that panel test statistics are more powerful. For all alternative values when T=100, by the time N reaches 25 our tests have almost 100 % powers. So the T=100,N=50 case was not included.

Table 5.1: Size of unit root tests in panel data when $\rho = 1$

(N,T)		T=25	T=50	T=100
N=10	LL	0.050	0.060	0.055
	IPS	0.040	0.046	0.046
	$\hat{\tau}_s$	0.046	0.054	0.048
	$\hat{\tau}_{ws}$	0.046	0.056	0.050
	MLEf	0.046	0.054	0.048
	MLEr	0.053	0.046	0.057
N=25	LL	0.044	0.039	0.045
	IPS	0.049	0.049	0.050
	$\hat{\tau}_s$	0.052	0.040	0.044
	$\hat{\tau}_{ws}$	0.046	0.051	0.048
	MLEf	0.049	0.049	0.045
	MLEr	0.054	0.050	0.048
N=50	LL	0.032	0.034	0.042
	IPS	0.051	0.046	0.046
	$\hat{\tau}_s$	0.046	0.053	0.050
	$\hat{\tau}_{ws}$	0.048	0.047	0.045
	MLEf	0.046	0.058	0.045
	MLEr	0.044	0.054	0.049

Table 5.2: Power of unit root tests in panel data when $\rho = 0.98$

(N,T)		T=25	T=50	T=100
N=10	LL	10.2	13.6	26.6
	IPS	8.6	11.4	22.2
	$\hat{\tau}_s$	15.1	31.9	68.6
	$\hat{\tau}_{ws}$	17.1	40.0	86.8
	MLEf	16.4	36.2	81.5
	MLEr	17.4	42.6	90.3
N=25	LL	13.4	23.9	51.2
	IPS	11.2	21.4	42.2
	$\hat{\tau}_s$	26.8	65.3	97.8
	$\hat{\tau}_{ws}$	34.0	82.8	100
	MLEf	27.0	76.8	99.9
	MLEr	36.1	84.8	100
N=50	LL	16.7	41.6	
	IPS	16.5	33.4	
	$\hat{\tau}_s$	46.2	92.6	
	$\hat{\tau}_{ws}$	62.6	98.7	
	MLEf	55.3	97.8	
	MLEr	61.2	98.8	

Table 5.3: Power of unit root tests in panel data when $\rho = 0.95$

(N,T)		T=25	T=50	T=100
N=10	LL	17.4	34.6	83.1
	IPS	13.0	25.9	72.2
	$\hat{\tau}_s$	39.8	83.0	100
	$\hat{\tau}_{ws}$	53.1	95.4	100
	MLEf	43.5	91.0	100
	MLEr	53.4	96.6	100
N=25	LL	30.8	68.3	99.9
	IPS	24.5	56.2	99.9
	$\hat{\tau}_s$	79.9	99.7	100
	$\hat{\tau}_{ws}$	93.5	100	100
	MLEf	87.8	100	100
	MLEr	93.7	100	100
N=50	LL	49.9	93.8	
	IPS	42.3	83.6	
	$\hat{\tau}_s$	98.0	100	
	$\hat{\tau}_{ws}$	100	100	
	MLEf	99.4	100	
	MLEr	99.9	100	

Table 5.4: Power of unit root tests in panel data when $\rho = 0.9$

(N,T)		T=25	T=50	T=100
N=10	LL	38.0	84.8	100
	IPS	25.9	72.7	100
	$\hat{\tau}_s$	81.9	99.9	100
	$\hat{\tau}_{ws}$	94.7	100	100
	MLEf	86.7	100	100
	MLEr	93.9	100	100
N=25	LL	67.9	99.8	100
	IPS	52.4	98.7	100
	$\hat{\tau}_s$	99.7	100	100
	$\hat{\tau}_{ws}$	100	100	100
	MLEf	100	100	100
	MLEr	100	100	100
N=50	LL	93.4	100	
	IPS	84.9	100	
	$\hat{\tau}_s$	100	100	
	$\hat{\tau}_{ws}$	100	100	
	MLEf	100	100	
	MLEr	100	100	

Table 5.5: Power of unit root tests in panel data when $\rho = 0.85$

(N,T)		T=25	T=50	T=100
N=10	LL	62.2	99.7	100
	IPS	46.6	98.3	100
	$\hat{\tau}_s$	97.4	100	100
	$\hat{\tau}_{ws}$	99.8	100	100
	MLEf	99.0	100	100
	MLEr	99.8	100	100
N=25	LL	95.3	100	100
	IPS	86.0	100	100
	$\hat{\tau}_s$	100	100	100
	$\hat{\tau}_{ws}$	100	100	100
	MLEf	100	100	100
	MLEr	100	100	100
N=50	LL	98.9	100	
	IPS	99.0	100	
	$\hat{\tau}_s$	100	100	
	$\hat{\tau}_{ws}$	100	100	
	MLEf	100	100	
	MLEr	100	100	

Chapter 6

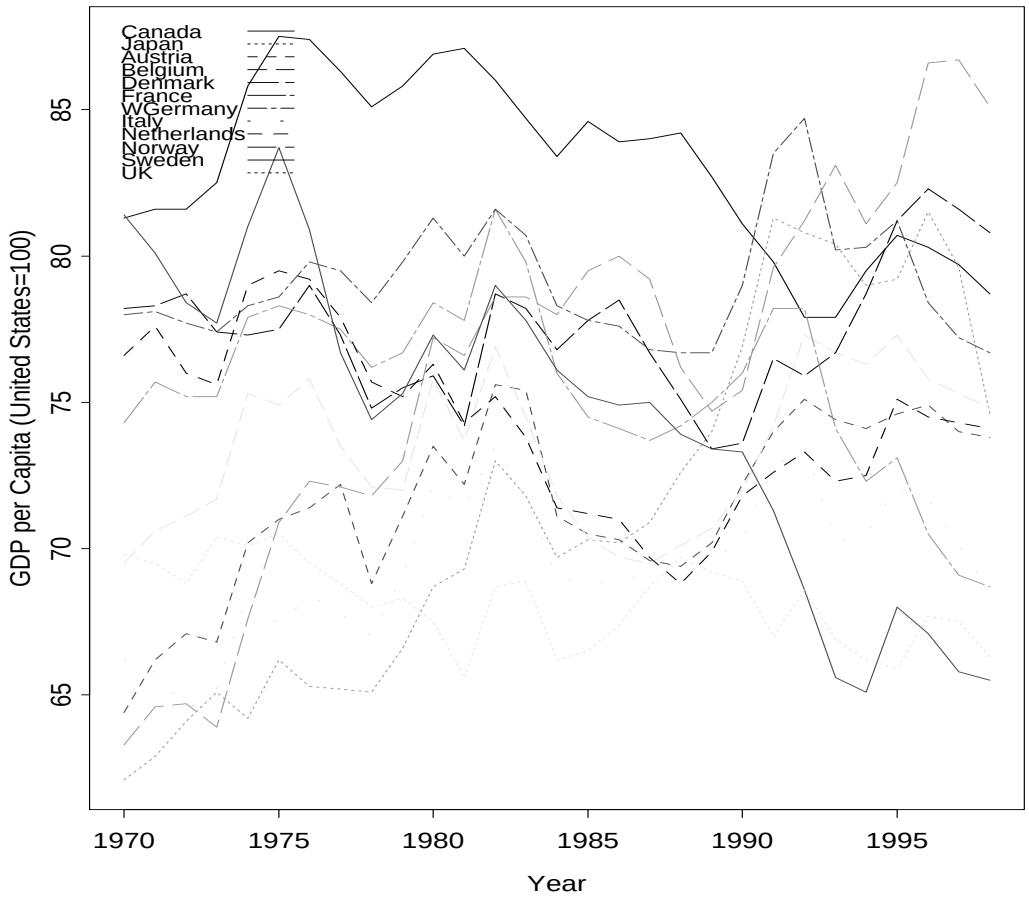
Example

As an example, we consider the Real Gross Domestic Product (real GDP) per Capita for 12 countries from 1970 to 1998. These historical data are available from the website of the United States Department of Labor (http://stats.bls.gov). The real GDP data are constructed using Elteto-Köves-Szulc (EKS) Purchasing Power Parities (PPPs) method. Since the price levels of the United States are set as 100, we do not include the USA data. The real GDP data are shown in Figure 6.1. We fit our data to the first order autoregressive model with individual fixed effects. The test statistics based on several estimation methods are shown in the below table. We use the same notations for the test statistics as in Chapter 5. Also, we present the 5 % critical value for each test statistic when N=10 and T=25.

Method	Test Statistic	Critical Value
LL	-4.76	-5.42
IPS	-1.64	-1.99
$\hat{\tau}_s$	0.02	-1.9
$\hat{\tau}_{ws}$	0.15	-2.1
MLEf	-0.39	-5.99

If we consider the one sided test at the 5 % level, we do not reject $H_0 : \rho = 1$ for all tests. Therefore, we fail to conclude that the real GDP per Capita follow the stationary processes for 12 countries.

Figure 6.1: Real GDP per Capita using EKS PPPs (United States=100)



Chapter 7

Summary

In this thesis, we have developed unit root test statistics based on weighted symmetric estimation and maximum likelihood estimation for the first order autoregressive panel model with individual effects. For the weighted symmetric estimation method, we have extended our model to the 2nd order autoregressive model. We have discussed the random individual effects case as well as the fixed individual effects case for the maximum likelihood estimation method.

For our test criteria, we have derived the limiting distribution. Unlike the results in the univariate case, we have Gaussian limiting distributions for unit root tests in panel data. A summary table of the T-limit and the joint limit distribution (the sequential limit in the random individual effects case) is given below.

Test	T-limit representation	Limiting distribution
$\hat{\tau}_s$	$\frac{\sum_{i=1}^{N}[-0.5+3(\int_0^1 B_i^2(t)dt-(\int_0^1 B_i(t)dt)^2)]}{\sum_{i=1}^{N}[\int_0^1 B_i^2(t)dt-(\int_0^1 B_i(t)dt)^2]}$	N(0,1.2)
$\hat{\tau}_{ws}$	$\frac{\sum_{i=1}^{N}[0.5(B_i^2(1)-1)+\int_0^1 B_i^2(t)dt-(\int_0^1 B_i(t)dt)B_i(1)]}{\sum_{i=1}^{N}[\int_0^1 B_i^2(t)dt-(\int_0^1 B_i(t)dt)^2]}$	N(0,1.5)
MLEf	no closed form	N(0,6.719)
MLEr	no closed form	N(0,1.5)

Through Monte Carlo experiments, we have computed the empirical percentiles

of our test statistics and obtained the smoothing functions for the 5 % critical values. The following table shows the smoothing functions and unsmoothed 5 % critical values when N=25 and T=50.

Test	5 % critical value (c.v.)	c.v. at N=25, T=50
$\hat{\tau}_s$	-1.81-0.4 $\frac{1}{\sqrt{N}}$+0.16 $\frac{1}{\sqrt{T}}$	-1.88
$\hat{\tau}_{ws}$	-2.02-0.34 $\frac{1}{\sqrt{N}}$+0.22 $\frac{1}{\sqrt{T}}$	-2.08
MLEf	-4.28-5.39 $\frac{1}{\sqrt{N}}$ + 0.37 $\frac{1}{\sqrt{T}}$	-5.27
MLEr	-2.02-8.45 $\frac{1}{\sqrt{N}}$+ 8.17 $\frac{1}{\sqrt{T}}$	-2.45

From our power study, we have found that our test statistics are quite a bit more powerful than those based on the ordinary least squares estimator. As has been pointed out in previous studies, a power improvement results from using panel data rather than a single time series. In the following table, we present the power of our test criteria for N=25 and T = 50 when ρ=0.95. To compare powers of unit root tests in panel data with those in a single time series, we consider the test statistics based on simple symmetric estimation in Dickey, Hasza and Fuller (1984), weighted symmetric estimation in Park and Fuller (1995) and maximum likelihood estimation in Gonzalez and Dickey (1999). For the computation of powers, we use 2,000 replicates.

Test	Power at N=25, T=50 ρ=0.95	Power at N=1, T=50 ρ=0.95
$\hat{\tau}_s$	99.7	12.3
$\hat{\tau}_{ws}$	100	13.7
MLEf	100	13.0
LL	68.3	
IPS	56.2	

References

Anderson, T. W. (1978), Repeated measurements on autoregressive processes. Journal of the American Statistical Association 73, 371-378.

Anderson, T. W. and Hsiao, C. (1982), Formulation and estimation of dynamic models using panel data. Journal of Econometrics 18, 47-82.

Balesta, P. and Nerlove, M. (1966), Pooling Cross Section and time series data in the estimation of a dynamic model: The demand for natural gas. Econometrica 34, 585-612.

Billingsley, P. (1999), *Convergence of Probability Measures* (2nd Edition). Wiley, New York, NY.

Billingsley, P. (1995), *Probability and Measure* (3rd Edition). Wiley, New York, NY.

Dickey, D. A. (1976), Estimation and hypothesis testing in nonstationary time series. Ph.D. thesis, Iowa State University, Ames, IA.

Dickey, D. A. and Fuller, W. A. (1979), Distribution of the estimators for autoregressive time series with a unit root. Journal of the American Statistical Association 74, 427-431.

Dickey, D. A.,Hasza, D. P., and Fuller, W. A. (1984), Testing for unit roots in seasonal time series. Journal of the American Statistical Association 79, 355-367.

Fuller, W. A. (1996), *Introduction to statistical time series* (2nd Edition). Wiley, New York, NY.

Gonzalez-Farias, G. and Dickey, D. A. (1999), Unit root test:an unconditional maximum likelihood approach. Bol.Soc.Mat,Mexicana 5, 199-221.

Harris, R. D. F. and Tzavalis, Elias. (1999), Inference for unit roots in dynamic panels where the time dimension is fixed. Journal of Econometrics 91,201-226.

Harville, D. A. (1997), *Matrix algebra from a statistician's perspective*. Springer, New York, NY.

Im, K. S.,Pesaran, M. H. and Shin, Y. (1997), Testing for unit roots in heterogeneous panels. Mimeo, Department of Applied Econometrics, University of Cambridge, Cambridge, UK.

Jolley, L. B. W. (1961), *Summation of series* (2nd edition). Dover Press, New York, NY.

Karlsson, S. and Löthgren, M. (2000), On the power and interpretation of panel unit root tests. Economics Letters 66, 249-255.

Levin, A. and Lin, C. F. (1992,1993), Unit root tests in panel data. Mimeo, University of California, San Diego, CA.

Mundlak, Y. (1978), On the pooling of time series and cross-section data. Econometrica 46, 69-86.

Park, H. J. and Fuller, W. A. (1995), Alternative estimators and unit root tests for the autoregressive process. Journal of Time Series Analysis 16, 415-429.

Pantula, S. G., Gonzalez-Farias, G., and Fuller, W. A. (1994), A comparison of Unit root test criteria. Jounal of Business and Economic Statistics 12, 449-459.

Phillips, P. C. B. (1987), Time series regression with a unit root. Econometrica 55, 277-310.

Phillips, P. C. B. and Moon, H. R. (1999), Linear regression limit theory for nonstationary panel data. Econometrica 67, 1057-1111.

Quah, D. (1994), Exploiting cross-section variation for unit root inference in dynamic data. Economics Letters 44,9-19.

Said, E. S. and Dickey, D. A. (1984), Testing for unit roots in autoregressive-moving average models of unknown order. Biometrika 71, 599-607.

SAS Institute Inc. (1997), *SAS/STAT software: changes and enhancements through release 6.12*. SAS Institute Inc., Cary, NC.

Serfling, R. J. (1980), *Approximation theorems of mathematical statistics*. Wiley, New York, NY.

Wolfinger, R., Tobias, Randy., and Sall, John. (1994), Computing Gaussian likelihoods and their derivatives for general linear mixed models. SIAM Journal of Science Computation 15, 1294-1310.

APPENDICES

Appendix A

Finite Sample Moments in Fixed Effects Case

In this section, we compute the finite sample moments of random variables with fixed effects under $H_0 : \rho = 1$. Harris and Tzavalis (1999) figure out the finite sample moments by expressing random variables as functions of white noise. We employ their method in computing the first and the second moments. First of all, the random variables in weighted symmetric estimation and maximum likelihood estimation with fixed effects can be expressed as functions of the following random variables under $H_0 : \rho = 1$. Let e_{it} be distributed as $N(0, \sigma^2)$.

- $\sum_{t=2}^{T} y_{it-1} = \sum_{t=1}^{T} (T-t) e_{it}$.
- $\sum_{t=2}^{T} (y_{it} - y_{it-1})^2 = \sum_{t=2}^{T} e_{it}^2$.
- $\sum_{t=2}^{T} (y_{it} - y_{it-1}) y_{it-1} = \sum_{t=1}^{T-1} \sum_{s=t+1}^{T} e_{it} e_{is}$.
- $\sum_{t=2}^{T} {y_{it-1}}^2 = \sum_{t=1}^{T} (T-t) e_{it}^2 + 2 \sum_{t=1}^{T-1} \sum_{s=t+1}^{T} (T-s) e_{it} e_{is}$.
- ${y_{iT}}^2 = \sum_{t=1}^{T} e_{it}^2 + 2 \sum_{t=1}^{T-1} \sum_{s=t+1}^{T} e_{it} e_{is}$.
- $(\sum_{t=2}^{T} y_{it-1}) y_{iT} = \sum_{t=1}^{T} (T-t) e_{it}^2 + \sum_{t=1}^{T-1} \sum_{s=t+1}^{T} (2T-s-t) e_{it} e_{is}$.

Based on previous results, we can find expectations,variances and covariances of those random variables.

- $E\Big[y_{i1}^2\Big] = E\Big[y_{i1}y_{iT}\Big] = \sigma^2$.
- $E\Big[\frac{\sum_{t=2}^{T}(y_{it}-y_{it-1})^2}{T}\Big] = \frac{T-1}{T}\sigma^2$.
- $E\Big[\frac{\sum_{t=2}^{T}(y_{it}-y_{it-1})y_{it-1}}{T}\Big] = 0$.
- $E\Big[\frac{\sum_{t=2}^{T}{y_{it-1}}^2}{T^2}\Big] = \frac{T-1}{2T}\sigma^2$.
- $E\Big[\frac{{y_{iT}}^2}{T}\Big] = \sigma^2$.
- $E\Big[\frac{(\sum_{t=2}^{T}y_{it-1})^2}{T^3}\Big] = \frac{2T^3-3T^2+T}{6T^3}\sigma^2$.
- $E\Big[\frac{y_{i1}(\sum_{t=2}^{T}y_{it-1})}{T}\Big] = \frac{T-1}{T}\sigma^2$.
- $E\Big[\frac{(\sum_{t=2}^{T}y_{it-1})y_{iT}}{T^2}\Big] = \frac{T-1}{2T}\sigma^2$.
- $Var\Big[y_{i1}^2\Big] = 2\sigma^4$.
- $Var\Big[\frac{\sum_{t=2}^{T}(y_{it}-y_{it-1})^2}{T}\Big] = \frac{2(T-1)}{T^2}\sigma^4$.
- $Var\Big[\frac{\sum_{t=2}^{T}(y_{it}-y_{it-1})y_{it-1}}{T}\Big] = \frac{T-1}{2T}\sigma^4$.
- $Var\Big[\frac{\sum_{t=2}^{T}{y_{it-1}}^2}{T^2}\Big] = \frac{(T-1)(T^2-T+1)}{3T^3}\sigma^4$.
- $Var\Big[\frac{{y_{iT}}^2}{T}\Big] = 2\sigma^4$.
- $Var\Big[\frac{(\sum_{t=2}^{T}y_{it-1})^2}{T^3}\Big] = \frac{1}{T^4}\Big[\frac{2T^4}{9}-\frac{2T^3}{3}+\frac{13T^2}{18}-\frac{T}{3}+\frac{1}{18}\Big]\sigma^4$.
- $Var\Big[y_{i1}y_{iT}\Big] = (T+1)\sigma^4$.
- $Var\Big[\frac{y_{i1}(\sum_{t=2}^{T}y_{it-1})}{T}\Big] = \frac{1}{T^2}\Big[\frac{T^3}{3}+\frac{T^2}{2}-\frac{11T}{6}+1\Big]\sigma^4$.
- $Var\Big[\frac{(\sum_{t=2}^{T}y_{it-1})y_{iT}}{T^2}\Big] = \frac{1}{T^2}\Big[\frac{7T^2}{12}-T+\frac{5}{12}\Big]\sigma^4$.

- $Cov\Big[y_{i1}^2, \dfrac{\sum_{t=2}^{T}(y_{it}-y_{it-1})^2}{T}\Big] = 0.$
- $Cov\Big[y_{i1}^2, \dfrac{\sum_{t=2}^{T}(y_{it}-y_{it-1})y_{it-1}}{T}\Big] = 0.$
- $Cov\Big[y_{i1}^2, \dfrac{\sum_{t=2}^{T}{y_{it-1}}^2}{T^2}\Big] = \dfrac{2(T-1)}{T^2}\sigma^4.$
- $Cov\Big[y_{i1}^2, \dfrac{{y_{iT}}^2}{T}\Big] = \dfrac{2}{T}\sigma^4.$
- $Cov\Big[y_{i1}^2, \dfrac{(\sum_{t=2}^{T}y_{it-1})^2}{T^3}\Big] = \dfrac{2(T-1)^2}{T^3}\sigma^4.$
- $Cov\Big[y_{i1}^2, y_{i1}y_{iT}\Big] = 2\sigma^4.$
- $Cov\Big[y_{i1}^2, \dfrac{y_{i1}(\sum_{t=2}^{T}y_{it-1})}{T}\Big] = \dfrac{2(T-1)}{T}\sigma^4.$
- $Cov\Big[y_{i1}^2, \dfrac{(\sum_{t=2}^{T}y_{it-1})y_{iT}}{T^2}\Big] = \dfrac{2(T-1)}{T^2}\sigma^4.$
- $Cov\Big[\dfrac{\sum_{t=2}^{T}(y_{it}-y_{it-1})^2}{T}, \dfrac{\sum_{t=2}^{T}(y_{it}-y_{it-1})y_{it-1}}{T}\Big] = 0.$
- $Cov\Big[\dfrac{\sum_{t=2}^{T}(y_{it}-y_{it-1})^2}{T}, \dfrac{\sum_{t=2}^{T}{y_{it-1}}^2}{T^2}\Big] = \dfrac{1}{T^3}\Big[T^2-3T+2\Big]\sigma^4.$
- $Cov\Big[\dfrac{\sum_{t=2}^{T}(y_{it}-y_{it-1})^2}{T}, \dfrac{{y_{iT}}^2}{T}\Big] = \dfrac{2(T-1)}{T^2}\sigma^4.$
- $Cov\Big[\dfrac{\sum_{t=2}^{T}(y_{it}-y_{it-1})^2}{T}, \dfrac{(\sum_{t=2}^{T}y_{it-1})^2}{T^3}\Big] = \dfrac{2T^3-9T^2+13T-6}{3T^4}\sigma^4.$
- $Cov\Big[\dfrac{\sum_{t=2}^{T}(y_{it}-y_{it-1})^2}{T}, y_{i1}y_{iT}\Big] = 0.$
- $Cov\Big[\dfrac{\sum_{t=2}^{T}(y_{it}-y_{it-1})^2}{T}, \dfrac{y_{i1}(\sum_{t=2}^{T}y_{it-1})}{T}\Big] = 0.$
- $Cov\Big[\dfrac{\sum_{t=2}^{T}(y_{it}-y_{it-1})^2}{T}, \dfrac{(\sum_{t=2}^{T}y_{it-1})y_{iT}}{T^2}\Big] = \dfrac{T^2-3T+2}{T^3}\sigma^4.$
- $Cov\Big[\dfrac{\sum_{t=2}^{T}(y_{it}-y_{it-1})y_{it-1}}{T}, \dfrac{\sum_{t=2}^{T}{y_{it-1}}^2}{T^2}\Big] = \dfrac{1}{T^2}\Big[\dfrac{T^2}{3}+\dfrac{2}{3}-T\Big]\sigma^4.$
- $Cov\Big[\dfrac{\sum_{t=2}^{T}(y_{it}-y_{it-1})y_{it-1}}{T}, \dfrac{{y_{iT}}^2}{T}\Big] = \dfrac{T-1}{T}\sigma^4.$

- $Cov\Big[\frac{\sum_{t=2}^{T}(y_{it}-y_{it-1})y_{it-1}}{T}, \frac{(\sum_{t=2}^{T}y_{it-1})^2}{T^3}\Big] = \frac{1}{T^3}\Big[\frac{T^3}{4} - \frac{5T^2}{6} + \frac{3T}{4} - \frac{1}{6}\Big]\sigma^4.$
- $Cov\Big[\frac{\sum_{t=2}^{T}(y_{it}-y_{it-1})y_{it-1}}{T}, y_{i1}y_{iT}\Big] = \frac{T-1}{T}\sigma^4.$
- $Cov\Big[\frac{\sum_{t=2}^{T}(y_{it}-y_{it-1})y_{it-1}}{T}, \frac{y_{i1}(\sum_{t=2}^{T}y_{it-1})}{T}\Big] = \frac{T^2-3T+2}{2T^2}\sigma^4.$
- $Cov\Big[\frac{\sum_{t=2}^{T}(y_{it}-y_{it-1})y_{it-1}}{T}, \frac{(\sum_{t=2}^{T}y_{it-1})y_{iT}}{T^2}\Big] = \frac{T^2-2T+1}{2T^2}\sigma^4.$
- $Cov\Big[\frac{\sum_{t=2}^{T}{y_{it-1}}^2}{T^2}, \frac{{y_{iT}}^2}{T}\Big] = \frac{1}{T^2}\Big[\frac{2T^2}{3} - T + \frac{1}{3}\Big]\sigma^4.$
- $Cov\Big[\frac{\sum_{t=2}^{T}{y_{it-1}}^2}{T^2}, \frac{(\sum_{t=2}^{T}y_{it-1})^2}{T^3}\Big] = \frac{1}{T^4}\Big[\frac{4T^4}{15} - \frac{2T^3}{3} + \frac{2T^2}{3} - \frac{T}{3} + \frac{1}{15}\Big]\sigma^4.$
- $Cov\Big[\frac{\sum_{t=2}^{T}{y_{it-1}}^2}{T^2}, y_{i1}y_{iT}\Big] = \frac{T-1}{T}\sigma^4.$
- $Cov\Big[\frac{\sum_{t=2}^{T}{y_{it-1}}^2}{T^2}, \frac{y_{i1}(\sum_{t=2}^{T}y_{it-1})}{T}\Big] = \frac{2T^2-3T+1}{3T^2}\sigma^4.$
- $Cov\Big[\frac{\sum_{t=2}^{T}{y_{it-1}}^2}{T^2}, \frac{(\sum_{t=2}^{T}y_{it-1})y_{iT}}{T^2}\Big] = \frac{5T^4-10T^3+7T^2-2T}{12T^4}\sigma^4.$
- $Cov\Big[\frac{{y_{iT}}^2}{T}, \frac{(\sum_{t=2}^{T}y_{it-1})^2}{T^3}\Big] = \frac{1}{T^2}\Big[\frac{T^2}{2} - T + \frac{1}{2}\Big]\sigma^4.$
- $Cov\Big[\frac{{y_{iT}}^2}{T}, y_{i1}y_{iT}\Big] = 2\sigma^4.$
- $Cov\Big[\frac{{y_{iT}}^2}{T}, \frac{y_{i1}(\sum_{t=2}^{T}y_{it-1})}{T}\Big] = \frac{T-1}{T}\sigma^4.$
- $Cov\Big[\frac{{y_{iT}}^2}{T}, \frac{(\sum_{t=2}^{T}y_{it-1})y_{iT}}{T^2}\Big] = \frac{T-1}{T}\sigma^4.$
- $Cov\Big[\frac{(\sum_{t=2}^{T}y_{it-1})^2}{T^3}, y_{i1}y_{iT}\Big] = \frac{T^3-2T^2+T}{T^3}\sigma^4.$
- $Cov\Big[\frac{(\sum_{t=2}^{T}y_{it-1})^2}{T^3}, \frac{y_{i1}(\sum_{t=2}^{T}y_{it-1})}{T}\Big] = \frac{1}{T^4}\Big[\frac{2T^4}{3} - \frac{5T^3}{3} + \frac{4T^2}{3} - \frac{T}{3}\Big]\sigma^4.$
- $Cov\Big[\frac{(\sum_{t=2}^{T}y_{it-1})^2}{T^3}, \frac{(\sum_{t=2}^{T}y_{it-1})y_{iT}}{T^2}\Big] = \frac{2T^3-5T^2+4T-1}{6T^3}\sigma^4.$

- $Cov\Big[y_{i1}y_{iT}, \dfrac{y_{i1}(\sum_{t=2}^{T} y_{it-1})}{T}\Big] = \dfrac{T^2+T-2}{2T}\sigma^4.$
- $Cov\Big[y_{i1}y_{iT}, \dfrac{(\sum_{t=2}^{T} y_{it-1})y_{iT}}{T^2}\Big] = \dfrac{3(T-1)}{2T}\sigma^4.$
- $Cov\Big[\dfrac{y_{i1}(\sum_{t=2}^{T} y_{it-1})}{T}, \dfrac{(\sum_{t=2}^{T} y_{it-1})y_{iT}}{T^2}\Big] = \dfrac{5T^2-9T+4}{6T^2}\sigma^4.$

Appendix B

Finite Sample Moments in Random Effects Case

Now we consider the finite sample moments of random variables in maximum likelihood estimation with random effects. In this case, each random variable has additional components due to the individual random effects unlike the usual random walk model. Let us assume that α_i has a Gaussian distribution with mean 0 and variance $\lambda\sigma^2$ and the e_{it}'s are independent of α_i and have the same distiribution as in Appendix A. Then, under $H_0 : \rho = 1$, the random variables of interest have the following expressions.

- $\sum_{t=2}^{T} y_{it-1} = (T-1)\alpha_i + \sum_{t=1}^{T}(T-t)e_{it}$.
- $\sum_{t=2}^{T}(y_{it} - y_{it-1})y_{it-1} = \alpha_i \sum_{t=2}^{T} e_{it} + \sum_{t=1}^{T-1}\sum_{s=t+1}^{T} e_{it}e_{is}$.
- $\sum_{t=2}^{T} {y_{it-1}}^2 = (T-1)\alpha_i^2 + 2\alpha_i \sum_{t=1}^{T}(T-t)e_{it} + \sum_{t=1}^{T}(T-t)e_{it}^2$
  $+ 2\sum_{t=1}^{T-1}\sum_{s=t+1}^{T}(T-s)e_{it}e_{is}$.
- ${y_{iT}}^2 = \alpha_i^2 + 2\alpha_i \sum_{t=1}^{T} e_{it} + \sum_{t=1}^{T} e_{it}^2 + 2\sum_{t=1}^{T-1}\sum_{s=t+1}^{T} e_{it}e_{is}$.
- $(\sum_{t=2}^{T} y_{it-1})y_{iT} = (T-1)\alpha_i^2 + \alpha_i \sum_{t=1}^{T}(2T-t-1)e_{it} + \sum_{t=1}^{T}(T-t)e_{it}^2$
  $+ \sum_{t=1}^{T-1}\sum_{s=t+1}^{T}(2T-s-t)e_{it}e_{is}$.

We can find the finite sample moments of these random variables from the previous expressions.

- $E\Big[y_{i1}^2\Big] = E\Big[y_{i1}y_{iT}\Big] = (\lambda+1)\sigma^2$.
- $E\Big[\frac{\sum_{t=2}^{T}(y_{it}-y_{it-1})y_{it-1}}{T}\Big] = 0$.
- $E\Big[\frac{\sum_{t=2}^{T}{y_{it-1}}^2}{T^2}\Big] = \frac{2(T-1)\lambda+T^2-T}{2T^2}\sigma^2$.
- $E\Big[\frac{{y_{iT}}^2}{T}\Big] = \frac{\lambda+T}{T}\sigma^2$.
- $E\Big[\frac{(\sum_{t=2}^{T}y_{it-1})^2}{T^3}\Big] = \frac{6(T-1)^2\lambda+2T^3-3T^2+T}{6T^3}\sigma^2$.
- $E\Big[\frac{y_{i1}(\sum_{t=2}^{T}y_{it-1})}{T}\Big] = \frac{(T-1)(\lambda+1)}{T}\sigma^2$.
- $E\Big[\frac{(\sum_{t=2}^{T}y_{it-1})y_{iT}}{T^2}\Big] = \frac{2(T-1)\lambda+T^2-T}{2T^2}\sigma^2$.
- $Var\Big[y_{i1}^2\Big] = 2(\lambda+1)^2\sigma^4$.
- $Var\Big[\frac{\sum_{t=2}^{T}(y_{it}-y_{it-1})y_{it-1}}{T}\Big] = \frac{2\lambda(T-1)+T^2-T}{2T^2}\sigma^4$.
- $Var\Big[\frac{\sum_{t=2}^{T}{y_{it-1}}^2}{T^2}\Big] = \frac{1}{3T^4}\Big[6(T-1)^2\lambda^2+(4T^3-6T^2+2T)\lambda+T^4-2T^3+2T^2-T\Big]\sigma^4$.
- $Var\Big[\frac{{y_{iT}}^2}{T}\Big] = \frac{2\lambda^2+4T\lambda+2T^2}{T^2}\sigma^4$.
- $Var\Big[\frac{(\sum_{t=2}^{T}y_{it-1})^2}{T^3}\Big] = \frac{1}{T^6}\Big[2(T-1)^4\lambda^2+(\frac{4T^5}{3}-\frac{14T^4}{3}+6T^3-\frac{10T^2}{3}+\frac{2T}{3})\lambda+\frac{2T^6}{9}-\frac{2T^5}{3}+\frac{13T^4}{18}-\frac{T^3}{3}+\frac{T^2}{18}\Big]\sigma^4$.
- $Var\Big[y_{i1}y_{iT}\Big] = \Big[2\lambda^2+(T+3)\lambda+(T+1)\Big]\sigma^4$.
- $Var\Big[\frac{y_{i1}(\sum_{t=2}^{T}y_{it-1})}{T}\Big] = \frac{1}{T^2}\Big[2(T-1)^2\lambda^2+(\frac{T^3}{3}+\frac{5T^2}{2}-\frac{35T}{6}+3)\lambda+\frac{T^3}{3}+\frac{T^2}{2}-\frac{11T}{6}+1\Big]\sigma^4$.

- $Var\Big[\frac{(\sum_{t=2}^{T} y_{it-1})y_{iT}}{T^2}\Big] = \frac{1}{T^4}\Big[2(T-1)^2\lambda^2 + (\frac{7T^3}{3} - \frac{9T^2}{2} + \frac{13T}{6})\lambda + \frac{7T^4}{12} - T^3 + \frac{5T^2}{12}\Big]\sigma^4.$
- $Cov\Big[y_{i1}^2, \frac{\sum_{t=2}^{T}(y_{it} - y_{it-1})y_{it-1}}{T}\Big] = 0.$
- $Cov\Big[y_{i1}^2, \frac{\sum_{t=2}^{T} {y_{it-1}}^2}{T^2}\Big] = \frac{2(T-1)(\lambda+1)^2}{T^2}\sigma^4.$
- $Cov\Big[y_{i1}^2, \frac{{y_{iT}}^2}{T}\Big] = \frac{2(\lambda+1)^2}{T}\sigma^4.$
- $Cov\Big[y_{i1}^2, \frac{(\sum_{t=2}^{T} y_{it-1})^2}{T^3}\Big] = \frac{2(T-1)^2(\lambda+1)^2}{T^3}\sigma^4.$
- $Cov\Big[y_{i1}^2, y_{i1}y_{iT}\Big] = 2(\lambda+1)^2\sigma^4.$
- $Cov\Big[y_{i1}^2, \frac{y_{i1}(\sum_{t=2}^{T} y_{it-1})}{T}\Big] = \frac{2(T-1)(\lambda+1)^2}{T}\sigma^4.$
- $Cov\Big[y_{i1}^2, \frac{(\sum_{t=2}^{T} y_{it-1})y_{iT}}{T^2}\Big] = \frac{2(T-1)(\lambda+1)^2}{T^2}\sigma^4.$
- $Cov\Big[\frac{\sum_{t=2}^{T}(y_{it} - y_{it-1})y_{it-1}}{T}, \frac{\sum_{t=2}^{T} {y_{it-1}}^2}{T^2}\Big] = \frac{1}{T^3}\Big[(T^2-3T+2)\lambda + \frac{T^3}{3} + \frac{2T}{3} - T^2\Big]\sigma^4.$
- $Cov\Big[\frac{\sum_{t=2}^{T}(y_{it} - y_{it-1})y_{it-1}}{T}, \frac{{y_{iT}}^2}{T}\Big] = \frac{2(T-1)\lambda + (T^2-T)}{T^2}\sigma^4.$
- $Cov\Big[\frac{\sum_{t=2}^{T}(y_{it} - y_{it-1})y_{it-1}}{T}, \frac{(\sum_{t=2}^{T} y_{it-1})^2}{T^3}\Big] = \frac{1}{T^4}\Big[(T^3 - 4T^2 + 5T - 2)\lambda + \frac{T^4}{4} - \frac{5T^3}{6} + \frac{3T^2}{4} - \frac{T}{6}\Big]\sigma^4.$
- $Cov\Big[\frac{\sum_{t=2}^{T}(y_{it} - y_{it-1})y_{it-1}}{T}, y_{i1}y_{iT}\Big] = \frac{(T-1)(\lambda+1)}{T}\sigma^4.$
- $Cov\Big[\frac{\sum_{t=2}^{T}(y_{it} - y_{it-1})y_{it-1}}{T}, \frac{y_{i1}(\sum_{t=2}^{T} y_{it-1})}{T}\Big] = \frac{(T^2-3T+2)(\lambda+1)}{2T^2}\sigma^4.$
- $Cov\Big[\frac{\sum_{t=2}^{T}(y_{it} - y_{it-1})y_{it-1}}{T}, \frac{(\sum_{t=2}^{T} y_{it-1})y_{iT}}{T^2}\Big] = \frac{1}{2T^3}\Big[(3T^2 - 7T + 4)\lambda + T^3 - 2T^2 + T\Big]\sigma^4.$

- $Cov\Big[\frac{\sum_{t=2}^{T} {y_{it-1}}^2}{T^2}, \frac{{y_{iT}}^2}{T}\Big] = \frac{1}{T^3}\Big[2(T-1)\lambda^2 + (2T^2-2T)\lambda + \frac{2T^3}{3} - T^2 + \frac{T}{3}\Big]\sigma^4.$
- $Cov\Big[\frac{\sum_{t=2}^{T} {y_{it-1}}^2}{T^2}, \frac{(\sum_{t=2}^{T} y_{it-1})^2}{T^3}\Big] = \frac{1}{T^5}\Big[2(T-1)^3\lambda^2 + (\frac{4T^4}{3} - \frac{10T^3}{3} + \frac{8T^2}{3} - \frac{2T}{3})\lambda + \frac{4T^5}{15} - \frac{2T^4}{3} + \frac{2T^3}{3} - \frac{T^2}{3} + \frac{T}{15}\Big]\sigma^4.$
- $Cov\Big[\frac{\sum_{t=2}^{T} {y_{it-1}}^2}{T^2}, y_{i1}y_{iT}\Big] = \frac{2(T-1)\lambda^2 + (T^2+T-2)\lambda + T^2 - T}{T^2}\sigma^4.$
- $Cov\Big[\frac{\sum_{t=2}^{T} {y_{it-1}}^2}{T^2}, \frac{y_{i1}(\sum_{t=2}^{T} y_{it-1})}{T}\Big] = \Big[\frac{6(T-1)^2\lambda^2 + (2T^3+3T^2-11T+6)\lambda}{3T^3} + \frac{2T^3-3T^2+T}{3T^3}\Big]\sigma^4.$
- $Cov\Big[\frac{\sum_{t=2}^{T} {y_{it-1}}^2}{T^2}, \frac{(\sum_{t=2}^{T} y_{it-1})y_{iT}}{T^2}\Big] = \Big[\frac{24(T-1)\lambda^2 + (20T^3-36T^2+16T)\lambda}{12T^4} + \frac{5T^4-10T^3+7T^2-2T}{12T^4}\Big]\sigma^4.$
- $Cov\Big[\frac{{y_{iT}}^2}{T}, \frac{(\sum_{t=2}^{T} y_{it-1})^2}{T^3}\Big] = \frac{1}{T^4}\Big[2(T-1)^2\lambda^2 + (2T^3-4T^2+2T)\lambda + \frac{T^4}{2} - T^3 + \frac{T^2}{2}\Big]\sigma^4.$
- $Cov\Big[\frac{{y_{iT}}^2}{T}, y_{i1}y_{iT}\Big] = \frac{2\lambda^2 + (2T+2)\lambda + 2T}{T}\sigma^4.$
- $Cov\Big[\frac{{y_{iT}}^2}{T}, \frac{y_{i1}(\sum_{t=2}^{T} y_{it-1})}{T}\Big] = \frac{2(T-1)\lambda^2 + (T^2+T-2)\lambda + T^2 - T}{T^2}\sigma^4.$
- $Cov\Big[\frac{{y_{iT}}^2}{T}, \frac{(\sum_{t=2}^{T} y_{it-1})y_{iT}}{T^2}\Big] = \frac{2(T-1)\lambda^2 + (3T^2-3T)\lambda + T^3 - T^2}{T^3}\sigma^4.$
- $Cov\Big[\frac{(\sum_{t=2}^{T} y_{it-1})^2}{T^3}, y_{i1}y_{iT}\Big] = \frac{2(T-1)^2\lambda^2 + (T^3-3T+2) + T^3 - 2T^2 + T}{T^3}\sigma^4.$
- $Cov\Big[\frac{(\sum_{t=2}^{T} y_{it-1})^2}{T^3}, \frac{y_{i1}(\sum_{t=2}^{T} y_{it-1})}{T}\Big] = \frac{1}{T^4}\Big[2(T-1)^3\lambda^2 + (\frac{2T^4}{3} + \frac{T^3}{3} - \frac{14T^2}{3} + \frac{17T}{3} - 4)\lambda + \frac{2T^4}{3} - \frac{5T^3}{3} + \frac{4T^2}{3} - \frac{T}{3}\Big]\sigma^4.$
- $Cov\Big[\frac{(\sum_{t=2}^{T} y_{it-1})^2}{T^3}, \frac{(\sum_{t=2}^{T} y_{it-1})y_{iT}}{T^2}\Big] = \frac{1}{6T^5}\Big[12(T-1)^3\lambda^2 + (10T^4-28T^3+26T^2-8T)\lambda + 2T^5 + 5T^4 + 4T^3 - T^2\Big]\sigma^4.$

- $Cov\Big[y_{i1}y_{iT}, \dfrac{y_{i1}(\sum_{t=2}^{T} y_{it-1})}{T}\Big] = \dfrac{4(T-1)\lambda^2 + (5T-6+T^2)\lambda + T^2 + T - 2}{2T}\sigma^4.$
- $Cov\Big[y_{i1}y_{iT}, \dfrac{(\sum_{t=2}^{T} y_{it-1})y_{iT}}{T^2}\Big] = \dfrac{2(T-1)^2\lambda^2 + (\dfrac{3T^2}{2} - 2 + \dfrac{T}{2})\lambda + \dfrac{3}{2}(T^2 - T)}{T^2}\sigma^4.$
- $Cov\Big[\dfrac{y_{i1}(\sum_{t=2}^{T} y_{it-1})}{T}, \dfrac{(\sum_{t=2}^{T} y_{it-1})y_{iT}}{T^2}\Big] = \dfrac{1}{6T^3}\Big[12(T-1)^2\lambda^2$
  $+ (5T^3 + 3T^2 - 20T + 12)\lambda + 5T^3 - 9T^2 + 4T\Big]\sigma^4.$

Appendix C

The Moments of the T-limit Random Variables

In this section, we derive the moments of the T-limit random quantities which are functions of standard Wiener processes $B(t)$. Dickey (1976) expresses these random quantites as infinite sums of standard Gaussian random variables. He defines

$$\eta' = \Big[\Psi, W, \Gamma\Big],$$

where

$$\Psi = \sum_{i=1}^{\infty} \sqrt{2}\gamma_i Z_i = B(1),$$

$$W = \sum_{i=1}^{\infty} \sqrt{2}\gamma_i^2 Z_i = \int_0^1 B(t)dt,$$

and

$$\Gamma = \sum_{i=1}^{\infty} \gamma_i^2 Z_i^2 = \int_0^1 B^2(t)dt.$$

where $\gamma_i = \dfrac{2(-1)^{i+1}}{(2i-1)\pi}$, the Z_i's are independent standard Gaussian random variables, and $B(t)$ is a Brownian motion. Based on these results from Dickey (1976), we can derive the expectation and covariance matrix of the vector $\Big[\Gamma, W^2, \Psi^2, \Psi W\Big]$.

Results

$$E\Big[\Gamma, W^2, \Psi^2, \Psi W\Big] = \Big[\frac{1}{2}, \frac{1}{3}, 1, \frac{1}{2}\Big],$$

and

$$Cov\Big[\Gamma, W^2, \Psi^2, \Psi W\Big] = \begin{bmatrix} \frac{1}{3} & \frac{4}{15} & \frac{2}{3} & \frac{5}{12} \\ \frac{4}{15} & \frac{2}{9} & \frac{1}{2} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{2} & 2 & 1 \\ \frac{5}{12} & \frac{1}{3} & 1 & \frac{7}{12} \end{bmatrix}.$$

Proof:

From Jolley(1961), we can find the values of $\sum_{i=1}^{\infty} \gamma_i^p$, for $p = 1...8$. Then

- $E(\Gamma) = \sum_{i=1}^{\infty} {\gamma_i}^2 E(Z_i^2) = \dfrac{1}{2}$,
- $E(W^2) = Var(W) = \sum_{i=1}^{\infty} 2{\gamma_i}^4 Var(Z_i) = \dfrac{1}{3}$,
- $E(\Psi^2) = Var(\Psi) = \sum_{i=1}^{\infty} 2{\gamma_i}^2 Var(Z_i) = 1$,
- $E(\Psi W) = Cov(\Psi, W) = \sum_{i=1}^{\infty} 2{\gamma_i}^3 Var(Z_i) = \dfrac{1}{2}$,
- $E(W^4) = \sum_{i=1}^{\infty} 4\gamma_i^8 E(Z_i^4) + 12 \sum_{i \neq j} \gamma_i^4 \gamma_j^4 E(Z_i^2 Z_j^2) = \dfrac{1}{3}$,
- $E(\Psi^2 W^2) = \sum_{i=1}^{\infty} 4\gamma_i^6 E(Z_i^4) + 4 \sum_{i \neq j} \gamma_i^2 \gamma_j^4 E(Z_i^2 Z_j^2) + 8 \sum_{i \neq j} \gamma_i^3 \gamma_j^3 E(Z_i^2 Z_j^2) = \dfrac{5}{6}$,
- $Var(\Gamma) = \sum_{i=1}^{\infty} {\gamma_i}^4 Var(Z_i^2) = \dfrac{1}{3}$,

- $Var(W^2) = E(W^4) - (E(W^2))^2 = \frac{2}{9}$,
- $Var(\Psi^2) = \sum_{i=1}^{\infty} 4{\gamma_i}^4 Var(Z_i^2) + \sum_{i \neq j} 8\gamma_i^2\gamma_j^2 E(Z_i^2 Z_j^2) = 2$,
- $Var(\Psi W) = E(\Psi^2 W^2) - (E(\Psi W))^2 = \frac{7}{12}$,
- $Cov(\Gamma, W^2) = 2\sum_{i=1}^{\infty} {\gamma_i}^6 Var(Z_i^2) = \frac{4}{15}$,
- $Cov(\Gamma, \Psi^2) = 2\sum_{i=1}^{\infty} {\gamma_i}^4 Var(Z_i^2) = \frac{2}{3}$,
- $Cov(\Gamma, \Psi W) = 2\sum_{i=1}^{\infty} {\gamma_i}^5 Var(Z_i^2) = \frac{5}{12}$,
- $Cov(W^2, \Psi^2) = 4\sum_{i=1}^{\infty} {\gamma_i}^6 Var(Z_i^2) + 8\sum_{i \neq j} \gamma_i^3\gamma_j^3 Var(Z_i Z_j) = \frac{1}{2}$,
- $Cov(W^2, \Psi W) = 4\sum_{i=1}^{\infty} {\gamma_i}^7 Var(Z_i^2) + 8\sum_{i \neq j} \gamma_i^3\gamma_j^4 Var(Z_i Z_j) = \frac{1}{3}$,
- $Cov(\Psi^2, \Psi W) = 4\sum_{i=1}^{\infty} {\gamma_i}^5 Var(Z_i^2) + 8\sum_{i \neq j} \gamma_i^2\gamma_j^3 Var(Z_i Z_j) = 1$.