Abstract

HUH, SEUNGHO. Sample Size Determination and Stationarity Testing in the Presence of Trend Breaks. (Under the direction of Professor David A. Dickey.)

Traditionally it is believed that most macroeconomic time series represent stationary fluctuations around a deterministic trend. However, simple applications of the Dickey-Fuller test have, in many cases, been unable to show that major macroeconomic variables are stationary univariate time series structure. One possible reason for non-rejection of unit roots is that the simple mean or linear trend function used by the tests are not sufficient to describe the deterministic part of the series. To address this possibility, unit root tests in the presence of trend breaks have been studied by several researchers.

In our work, we deal with some issues associated with unit root testing in time series with a trend break.

The performance of various unit root test statistics is compared with respect to the break induced size distortion problem. We examine the effectiveness of tests based on symmetric estimators as compared to those based on the least squares estimator. In particular, we show that tests based on the weighted symmetric estimator not only eliminate the spurious rejection problem but also have reasonably good power properties when modified to allow for a break.

We suggest alternative test statistics for testing the unit root null hypothesis in the presence of a trend break. Our new test procedure, which we call the "bisection" method, is based on the idea of subgrouping. This is simpler than other methods since the necessity of searching for the break is avoided.

Using stream flow data from the US Geological Survey, we perform a temporal analysis of some hydrologic variables. We first show that the time series for the target variables are stationary, then focus on finding the sample size necessary to

detect a mean change if one occurs. Three different approaches are used to solve this problem : OLS, GLS and a frequency domain method. A cluster analysis of stations is also performed using these sample sizes as data. We investigate whether available geographic variables can be used to predict cluster membership.

SAMPLE SIZE DETERMINATION AND STATIONARITY TESTING IN THE PRESENCE OF TREND BREAKS

by

SEUNGHO HUH

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 D. A. Dickey, Chair

Professor B. B. Bhattacharyya

Professor M. L. Gumpertz

Professor S. G. Pantula

To my wife, son, parents and parents-in-law

Biography

Seungho Huh was born in Seoul, Korea in March 21, 1964. In 1987, he graduated from Seoul National University with a Bachelor's degree in Statistics. After receiving his Master's degree in Statistics at Seoul National University in 1989, he started to work for Korea Telecom as a member of the technical staff. He married Seungshin in 1993 and had a son, Jungwoo, in 1996. Since the fall semester of 1996, he has been studying for the Ph.D. degree at North Carolina State University, under the direction of Dr. David A. Dickey.

Acknowledgements

I would be grateful to my advisor Dr. Dickey from the bottom of my heart. My research became fruitful thanks to his helpful instruction and advice. Without meeting him, I must have been frustrated in writing this dissertation. He also showed me what a scholar should look like. I was always impressed with his enthusiasm for both teaching and research.

I would like to thank my committee members, Dr. Bhattacharyya, Dr. Gumpertz and Dr. Pantula, for their thoughtful comments and suggestions that greatly improved the quality of this dissertation. Another memorable thing during my stay in the department was Dr. Pantula's devotional concern for students as the Director of Graduate Program.

My wife, Seungshin, has always been encouraging and supporting me during the whole period of study. If it had not been for her, my hard days of schoolwork would have been worse. Always on my side was Seungshin and my son Jungwoo when I was worn out. I am sure they would be happier than me with obtaining my Ph.D. degree.

Lastly, but not least importantly, this honor is surely attributed to my parents and parents-in-law who brought up me and Seungshin with their whole hearts. Throughtout the years of my studying abroad, they affectionately supported me from a distance. They deserve to be congratulated on their son or son-in-law's achievement.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Literature review

In the field of time series analysis, there are many good references and textbooks available among which are Priestley (1981), Wei (1990), Box, Jenkins and Reinsel (1994), Hamilton (1994) and Fuller (1996). Of those mentioned, only Box, Jenkins and Reinsel (1994), Hamilton (1994) and Fuller (1996) address the topic of unit roots.

The analysis of time series with a unit root became popular following, among others, the work of Dickey (1976). He shows that under the null hypothesis $H_0 : \rho = 1$ in the model $Y_t = \rho Y_{t-1} + e_t$ ($Y_0 = 0$, $e_t \sim NI(0, \sigma^2)$), $n(\hat{\rho} - 1) \xrightarrow{d} 0.5(T^2 - 1)/G$ where $\hat{\rho}$ is the least squares estimator of ρ, $T = \sum_{i=1}^{\infty} \sqrt{2}\gamma_i U_i$, $G = \sum_{i=1}^{\infty} \gamma_i^2 U_i^2$, $\gamma_i = 2(-1)^{i+1}/\pi(2i-1)$ and $U_i \sim NI(0,1)$. Dickey (1976) finds the distribution of $n(\hat{\rho}_\mu - 1)$ and $n(\hat{\rho}_\tau - 1)$ where $\hat{\rho}_\mu$ and $\hat{\rho}_\tau$ are the least squares estimators of ρ when the fitted model contains a nonzero mean term and a linear trend term respectively. He also tabulates the distribution of t type statistics under the above 3 models. Critical values of these distribuions can be found in Dickey (1976) and Fuller (1996). See Dickey and Fuller (1979) for more information. The random variables $T = W(1)$ and $G = \int_0^1 W^2(t)dt$ can be expressed as functionals of a standard Wiener process $W(t)$.

Said and Dickey (1984) consider testing for unit roots in the general ARMA(p, q)

model of unknown orders p and q. They show that fitting a high order autoregressive model is an appropriate way to test for a unit root in a model of unknown order. Their method is commonly known as the augmented Dickey-Fuller (ADF) test although ADF is also mentioned in Dickey and Fuller (1979) for pure autoregressions..

Gonzalez-Farias (1992) considers maximum likelihood estimation of the parameters in autoregressive time series. Her paper studies the maximizers of the exact stationary likelihood function when the input data are from a unit root process. She proposes a new unit root test based on these estimators and derives its limiting distribution. Dickey, Hasza and Fuller (1984) discuss the properties of another type of estimator known as the simple symmetric estimator. Park and Fuller (1995) study the weighted symmetric estimator which is another in the class of symmetric estimators. Pantula, Gonzalez-Farias and Fuller (1994) compare the performance of these and other unit root test criteria and determine that the weighted symmetric estimator and the unconditional maximum likelihood estimator provide the most powerful tests against the stationary alternative.

Nelson and Plosser (1982) investigate a collection of important macroeconomic time series asking if they are better characterized as stationary fluctuations around a deterministic trend or as nonstationary processes that have no tendency to return to a deterministic path. They apply the augmented Dickey-Fuller test to 14 major macroeconomic time series. Their study, which finds that most macroeconomic variables have a univariate time series structure with a unit root, is followed by a series of empirical analyses with similar findings. Several researchers have tried to explain these rather interesting results, believing that most macroeconomic time series represent stationary fluctuations around a deterministic trend.

Two possible reasons for non-rejection of unit roots are that (1) there are not enough observations to endow the tests with sufficient power and (2) the simple mean or linear trend function used by the tests are not sufficient to describe the deterministic part of the series. To address the second possibility, the topics of unit root testing and trend breaks are combined by Perron (1989). In the pioneering work of Perron (1989), he considers the null hypothesis that a time series has a unit root with possibly nonzero drift against the alternative that the process is trend stationary. Allowing, under both the null and alternative hypotheses, for the presence of a one-time change in the trend function, Perron (1989) derives test statistics which can distinguish the two hypotheses when such a break is present. He applies these tests to the Nelson and Plosser (1982) data set and rejects the unit root hypothesis for 11 out of the 14 series. He argues that the failure of the usual Dickey-Fuller test to reject the unit root null hypothesis reflects not an actual unit root, but instead that the data are trend stationary around a broken trend. This has come to be called the 'Perron phenomenon'.

Since Perron assumes the break point is known a priori and treated as exogenous, criticism has emerged from some authors such as Banerjee, Lumsdaine and Stock (1992), Christiano (1992) and Zivot and Andrews (1992).

The most notable of them is Christiano (1992) who argues that the choice of break dates has to be viewed as being correlated with the data. He presents some algorithms for selecting the break date endogenously. A bootstrap approach that takes into account pretest data examination is used to test the null hypothesis of no trend break.

Banerjee, Lumsdaine and Stock (1992) also treat the break date as unknown a priori and suggest recursive and sequential tests of the unit root null hypothesis. Their theoretical results are similar to those of Zivot and Andrews (1992).

Zivot and Andrews (1992) transform Perron's (1989) unit root test, which is conditional on structural change at a known point in time, into an unconditional unit root test in which the break point is estimated rather than fixed. They check all possible break points and take the minimum t statistic. The data series considered by Perron (1989) are reanalyzed using their estimated break point test statistic.

Perron (1994) is a good summarizing reference for the issue of unit roots and structural change. He presents works by himself and many other researchers.

Perron's (1997) work is closely related to that of Banerjee et al (1992) and that of Zivot and Andrews (1992) in that similar procedures and series are analyzed. He first reexamines his findings from Perron (1989). Unlike his previous study, Perron (1997) assumes that the date of a possible break is not fixed a priori but instead is unknown. He considers various methods to select the break point and the truncation lag parameter. Most of the rejections reported in Perron (1989) are confirmed using his new approach.

Additional tests for a unit root allowing for a break at an unknown time are given by Vogelsang and Perron (1998). They focus on the additive outlier approach where the break is sudden, as opposed to the innovational outlier approach where the change occurs slowly over time.

Recently, Leybourne, Mills and Newbold (1998) discover and analyze what they call the 'converse Perron phenomenon'. They show that routine application of the Dickey-Fuller test can lead to a severe problem of spurious rejection of the null hypothesis when the data are generated from a random walk with a trend-break near the beginning. Leybourne et al (1998) also show that this is not an end-effects problem in the ordinary least squares estimation since their results hold asymptotically for the break size growing at an appropriate rate. For a structural break of fixed size, on the other hand, Amsler and Lee (1995) suggest that the asymptotic distributions of the usual Dickey-Fuller tests under $\rho = 1$ are unaffected and so the spurious rejection problem exists only in finite samples. Chapter 2 of this dissertation is closely related to Leybourne et al (1998) in that we compare the performance of the symmetric estimator with that of the ordinary least squares estimator in the problem of spurious rejection.

1.2 Outline of research

In our work, we concentrate on analyzing time series data with a trend-break. Three somewhat independent papers, Chapters 2, 3 and 4, are related by their focus on the trend-break issue.

In Chapter 2, we compare the performance of various unit root test statistics in the break induced size distortion problem. Leybourne et al (1998) show that, if the data generating process is a random walk with a break near the beginning, standard Dickey-Fuller (1979) tests based on the least squares estimator have serious size distortion. In Chapter 2, we examine the performance of alternative tests based on symmetric estimators. In particular we show that tests based on the weighted symmetric estimator not only eliminate spurious rejection in these problem data sets but also have reasonably good power properties when modified to allow for a break.

In Chapter 3, we suggest alternative test statistics for testing the unit root null hypothesis in the presence of a trend-break. Our new test procedure which we call the "bisection" method is based on the idea of subgrouping. The idea here is to split the data in half and look at the minimum of the resulting two unit root test statistics. This avoids the necessity of searching for the break. It uses all the data in the sense that the minimum is chosen, but clearly is not efficient in its use of the data. We anticipate paying a price in power for a gain in simplicity. Considering some data generating processes, we display empirical size and power results from simulation. We also apply our bisection method to the well-known Nelson and Plosser (1982) data set and compare the results with those of others. The simple bisection method rejects unit roots in several, but not all, of the series for which the more complicated search methods reject.

In Chapter 4, we analyze stream flow data from the US Geological Survey. We perform a temporal analysis of some hydrologic variables of interest to USGS scientists. We are particularly interested in the required sample size to detect a break, or shift, of the mean in the presence of autocorrelation. We apply three alternative approaches to obtain the sample size. The first approach is a simple method using ordinary least squares theory that assumes a known error variance. The second approach uses generalized least squares and is more realistic in that it treats the error variance as unknown. The third approach is a frequency domain method for which identification of the autocorrelation structure is not needed. Chapter 4 finishes with a cluster analysis, performed to obtain clusters based on the sample sizes. The cluster pattern does not suggest any particular physiographic segregation. This is consistent with a regression analysis which shows no influence of available physiographic variables on the sample sizes.

Chapter 2

Comparison of Break Induced Size Distortions among Unit Root Test Statistics

2.1 Introduction

Recently there has been much interest in testing for a unit root in a time series that has a trend break. Leybourne, Mills and Newbold (1998) study the behavior of standard unit root tests when the data generating process contains a trend break not accounted for by the fitted model. For data consisting of a random walk with a shift in level, they report empirical sizes less than the nominal level when the shift is not too near the beginning of the series. However, if the shift is near the beginning of the series, they report too many rejections of the unit root null hypothesis. Thus a unit root process with an early level shift is too often declared stationary. Leybourne et al (1998) call this the 'converse Perron phenomenon' in contrast to the well known 'Perron phenomenon' investigated by Perron (1989).

If we switch from tests based on the least squares estimator to more powerful tests based on the simple symmetric or weighted symmetric estimators (see Pantula, Gonzalez-Farias and Fuller (1994)), we find that this phenomenon decreases dramatically, seeming to disappear in the weighted symmetric case.

In section 2.2, we examine the simple random walk model with a break in level. We consider the simple symmetric estimator $\tilde{\rho}_s$ and the weighted symmetric estimator $\tilde{\rho}_w$ of the lag 1 autoregressive coefficient ρ. Empirical sizes for the pivotal statistics τ_s and τ_w associated with $\tilde{\rho}_s$ and $\tilde{\rho}_w$, respectively, are presented to compare with those for τ_μ, the ordinary least squares test statistic. To gain some insight into the empirical results, we examine the expected values of the quadratic forms constituting the test statistics. This analysis includes $\tau_{s,\tau}$ and $\tau_{w,\tau}$, the linear trend adjusted versions of τ_s and τ_w, respectively, to compare with τ_τ, the linear trend adjusted version of τ_μ. An augmented test with lagged first differences is introduced to extend the results to more general cases.

In section 2.3, we consider the random walk model with a break in trend, using $\tau_{s,\tau}$ and $\tau_{w,\tau}$ as test statistics to obtain empirical sizes. The results are compared with those for τ_τ. Section 2.4 presents some power results for a test based on the weighted symmetric estimator, modified to allow for a break. Finally we make some concluding remarks in section 2.5.

2.2 Data with a break in level

In this section, following Leybourne et al (1998), we consider a data generating process (DGP)

$$\text{Model I} \quad : \quad Y_t = m\sigma I(t > c) + X_t, \quad X_t = X_{t-1} + e_t, \quad t = 1, 2, ..., n$$

where the e_t are normal independent $(0,\sigma^2)$ random variables, m is the size of the break as a multiple of σ and $c = n\lambda$ is the time of the break. The estimators of the lag 1 autoregressive coefficient ρ are invariant to σ so we can assume $\sigma = 1$. This model corresponds to 'Model (A) with $\mu = 0$ under the null hypothesis' of Perron (1989). Figure 2.1 presents some typical time series data generated from Model I.

2.2.1 Using mean adjusted statistics

As in Leybourne et al (1998), to test the unit root null hypothesis, we consider

$$\hat{\rho}_\mu = \frac{\sum_{t=2}^{n}(Y_t - \bar{Y})(Y_{t-1} - \bar{Y})}{\sum_{t=2}^{n}(Y_{t-1} - \bar{Y})^2} \tag{2.1}$$

which is an ordinary least squares (OLS) estimator* of ρ in the non-zero mean first order autoregressive process

$$Y_t = \mu + \rho Y_{t-1} + e_t. \tag{2.2}$$

We use, as a test statistic, the associated pivotal statistic

$$\tau_\mu = \frac{\hat{\rho}_\mu - 1}{\text{s.e.}} = \frac{\hat{\rho}_\mu - 1}{\sqrt{[\sum_{t=2}^{n}(Y_{t-1} - \bar{Y})^2]^{-1} s^2}} \tag{2.3}$$

where

$$s^2 = \frac{1}{n-3}\sum_{t=2}^{n}[Y_t - \bar{Y} - \hat{\rho}_\mu(Y_{t-1} - \bar{Y})]^2.$$

Using the test statistic τ_μ, Leybourne et al (1998) reported the empirical sizes of nominal 5% level tests for various values of break size, m, and break time, c. They found that, as m increases, an increasingly severe phenomenon of 'spurious rejection' of the unit root null hypothesis emerges. They also indicated that the earlier is the break, the greater is the rejection rate. Our empirical sizes for τ_μ are in close agreement with those reported by Leybourne et al (1998).

We consider alternative symmetric estimators $\tilde{\rho}_s$ and $\tilde{\rho}_w$ (Fuller, 1996) of ρ in (2.2) and their associated pivotal statistics τ_s and τ_w. As will be seen, their performance is often quite superior to that of τ_μ.

For the process (2.2), the simple symmetric (SS) estimator can be written as

$$\tilde{\rho}_s = \frac{\sum_{t=2}^{n} y_t y_{t-1}}{\sum_{t=2}^{n-1} y_t^2 + \frac{1}{2}(y_1^2 + y_n^2)}$$

*This is from the least squares regression of $Y_t - \bar{Y}$ on $Y_{t-1} - \bar{Y}$. The regression of Y_t on 1, Y_{t-1} gives a slightly different estimator, differing trivially from (2.1).

where $y_t = Y_t - \bar{Y}$. The pivotal statistic for the SS estimator is

$$\tau_s = \frac{\tilde{\rho}_s - 1}{\text{s.e.}} = \frac{\tilde{\rho}_s - 1}{\sqrt{\tilde{\sigma}_s^2[\sum_{t=2}^{n-1} y_t^2 + \frac{1}{2}(y_1^2 + y_n^2)]^{-1}}} \tag{2.4}$$

where

$$\begin{aligned}\tilde{\sigma}_s^2 &= \frac{1}{n-2}[\sum_{t=2}^{n} \frac{1}{2}(y_t - \tilde{\rho}_s y_{t-1})^2 + \sum_{t=1}^{n-1} \frac{1}{2}(y_t - \tilde{\rho}_s y_{t+1})^2] \\ &= \frac{1}{n-2}[\sum_{t=2}^{n}(y_t - \tilde{\rho}_s y_{t-1})^2 + \frac{1}{2}(1 - \tilde{\rho}_s^2)(y_1^2 - y_n^2)].\end{aligned}$$

The weighted symmetric (WS) estimator for the process (2.2) is

$$\tilde{\rho}_w = \frac{\sum_{t=2}^{n} y_t y_{t-1}}{\sum_{t=2}^{n-1} y_t^2 + \frac{1}{n}\sum_{t=1}^{n} y_t^2} \tag{2.5}$$

and the associated pivotal statistic is

$$\tau_w = \frac{\tilde{\rho}_w - 1}{\text{s.e.}} = \frac{\tilde{\rho}_w - 1}{\sqrt{\tilde{\sigma}_w^2(\sum_{t=2}^{n-1} y_t^2 + \frac{1}{n}\sum_{t=1}^{n} y_t^2)^{-1}}} \tag{2.6}$$

where

$$\begin{aligned}\tilde{\sigma}_w^2 &= \frac{1}{n-2}[\sum_{t=2}^{n} w_t(y_t - \tilde{\rho}_w y_{t-1})^2 + \sum_{t=1}^{n-1}(1 - w_{t+1})(y_t - \tilde{\rho}_w y_{t+1})^2] \\ &= \frac{1}{n-2}[\sum_{t=2}^{n}(y_t - \tilde{\rho}_w y_{t-1})^2 + (1 - \tilde{\rho}_w^2)(y_1^2 - \frac{1}{n}\sum_{t=1}^{n} y_t^2)].\end{aligned}$$

and $w_t = \frac{t-1}{n}$ for $t = 1, \cdots, n$. See Fuller (1996) for detailed discussion of symmetric estimators.

In Table 2.1, we show empirical sizes of nominal 5% level tests using τ_μ, τ_s and τ_w. We consider the same values of m and c as those used by Leybourne et al (1998). Simulations are based on 5,000 replications at each (m, c) combination so the standard error of an empirical size is less than $\sqrt{.25/5000} = .007$. The sample size per replication is $n = 100$. Critical values used are those for the mean removed case in Fuller (1996). Data are generated with $X_0 = 0$ so that $X_1 = e_1 \sim N(0, 1)$.

Our τ_μ simulations match those of Leybourne et al (1998), however we find that the 'spurious rejection' problem is far less severe in the symmetric estimators. For

τ_s, the proportion of rejections shows symmetry around $c = 50$ and gets larger as m increases . The rejection rate is higher for c values in the first or last part than for those in the middle of the data. The τ_s rejection rate is much smaller than that of τ_μ for early breaks but larger for late breaks.

For τ_w, the rejection rate is almost invariant to the break size m and shows little size distortion. Most importantly, compared to τ_μ, τ_w gives satisfactory performance in that the largest empirical size is around 0.056 (m=5) and most of the others are less than the nominal level 0.05. Figure 2.2 shows empirical sizes of the three tests using mean adjusted statistics for model I with $m = 10$.

2.2.1.1 Effect of ignoring the break

To obtain some insight into the empirical results, we examine some statistical properties of the test statistics when the break is ignored. As suggested by Amsler and Lee (1995), the asymptotic distributions of the usual Dickey-Fuller tests under $\rho = 1$ are unaffected by a structural break of fixed size. We now show that this also holds good for the tests based on the symmetric estimators. Thus the spurious rejection problem exists only in finite samples and the effect of ignoring the break may differ between the OLS estimator and the symmetric estimators.

Let $x_t = X_t - \bar{X}$ in Model I where $\bar{X} = \frac{1}{n}\sum_{t=1}^{n} X_t$. Then we know that

$$\bar{Y} = \bar{X} + (1-\lambda)m$$

and

$$y_t = \begin{cases} x_t - (1-\lambda)m & \text{if } t \le c = n\lambda \\ x_t + \lambda m & \text{if } t > c = n\lambda. \end{cases}$$

Recall that the OLS estimator of ρ is

$$\hat{\rho}_\mu = \frac{\sum_{t=2}^{n} y_t y_{t-1}}{\sum_{t=2}^{n} y_{t-1}^2}$$

and

$$\hat{\rho}_\mu - 1 = \frac{\sum_{t=2}^n y_{t-1}(y_t - y_{t-1})}{\sum_{t=2}^n y_{t-1}^2} \equiv \frac{N_\mu}{D_\mu} \tag{2.7}$$

By some straightforward algebra, the numerator of $\hat{\rho}_\mu - 1$ can be written as

$$\begin{aligned}\sum_{t=2}^n y_{t-1}(y_t - y_{t-1}) &= \sum_{t=2}^n x_{t-1}(x_t - x_{t-1}) - (1-\lambda)m \sum_{t=2}^{n\lambda+1} e_t + \lambda m \sum_{t=n\lambda+2}^{n} e_t \\ &\quad + m x_{n\lambda} - (1-\lambda)m^2.\end{aligned}$$

Therefore

$$\frac{1}{n}\sum_{t=2}^n y_{t-1}(y_t - y_{t-1}) = \frac{1}{n}\sum_{t=2}^n x_{t-1}(x_t - x_{t-1}) + O_p(\frac{1}{\sqrt{n}}).$$

The denominator of $\hat{\rho}_\mu - 1$ can be written as

$$\begin{aligned}\sum_{t=2}^n y_{t-1}^2 &= \sum_{t=2}^n x_{t-1}^2 - 2m \sum_{t=2}^{n\lambda+1} x_{t-1} + 2\lambda m \sum_{t=2}^n x_{t-1} \\ &\quad + (1-\lambda)^2 m^2 n\lambda + \lambda^2 m^2 (n - n\lambda - 1).\end{aligned}$$

Using the fact that

$$\frac{1}{n^2}\sum_{t=2}^{n\lambda+1} x_{t-1} = O_p(\frac{1}{\sqrt{n}})$$

and

$$\frac{1}{n^2}\sum_{t=2}^{n} x_{t-1} = O_p(\frac{1}{n\sqrt{n}}),$$

we have

$$\frac{1}{n^2}\sum_{t=2}^n y_{t-1}^2 = \frac{1}{n^2}\sum_{t=2}^n x_{t-1}^2 + O_p(\frac{1}{\sqrt{n}}). \tag{2.8}$$

Combining the results,

$$\begin{aligned}n(\hat{\rho}_\mu - 1) &= \frac{\frac{1}{n}\sum_{t=2}^n y_{t-1}(y_t - y_{t-1})}{\frac{1}{n^2}\sum_{t=2}^n y_{t-1}^2} \\ &= \frac{\frac{1}{n}\sum_{t=2}^n x_{t-1}(x_t - x_{t-1}) + O_p(\frac{1}{\sqrt{n}})}{\frac{1}{n^2}\sum_{t=2}^n x_{t-1}^2 + O_p(\frac{1}{\sqrt{n}})} \\ &= \frac{\frac{1}{n}\sum_{t=2}^n x_{t-1}(x_t - x_{t-1})}{\frac{1}{n^2}\sum_{t=2}^n x_{t-1}^2} + O_p(\frac{1}{\sqrt{n}}) \\ &\equiv n(\hat{\rho}_{\mu,0} - 1) + O_p(\frac{1}{\sqrt{n}})\end{aligned} \tag{2.9}$$

where $\hat{\rho}_{\mu,0}$ stands for the OLS estimator of ρ with no break.

Recall that the SS estimator of ρ is

$$\tilde{\rho}_s = \frac{\sum_{t=2}^{n} y_t y_{t-1}}{\sum_{t=2}^{n-1} y_t^2 + \frac{1}{2}(y_1^2 + y_n^2)}$$

and

$$\begin{aligned} \tilde{\rho}_s - 1 &= \frac{\sum_{t=2}^{n} y_t y_{t-1} - \sum_{t=2}^{n-1} y_t^2 - \frac{1}{2}(y_1^2 + y_n^2)}{\sum_{t=2}^{n-1} y_t^2 + \frac{1}{2}(y_1^2 + y_n^2)} \\ &= \frac{-\frac{1}{2}\sum_{t=2}^{n}(y_t - y_{t-1})^2}{\sum_{t=2}^{n-1} y_t^2 + \frac{1}{2}(y_1^2 + y_n^2)} \equiv \frac{N_s}{D_s} \end{aligned} \qquad (2.10)$$

by rearranging the terms in the numerator. The numerator of $\tilde{\rho}_s - 1$ can be rewritten as

$$-\frac{1}{2}\sum_{t=2}^{n}(y_t - y_{t-1})^2 = -\frac{1}{2}\sum_{t=2}^{n}(x_t - x_{t-1})^2 - m(x_{n\lambda+1} - x_{n\lambda}) - \frac{m^2}{2}.$$

Therefore

$$\begin{aligned} \frac{1}{n}\{-\frac{1}{2}\sum_{t=2}^{n}(y_t - y_{t-1})^2\} &= -\frac{1}{2n}\sum_{t=2}^{n}(x_t - x_{t-1})^2 - \frac{m}{n}e_{n\lambda+1} - \frac{m^2}{2n} \\ &= \frac{1}{n}\{-\frac{1}{2}\sum_{t=2}^{n}(x_t - x_{t-1})^2\} + O_p(\frac{1}{n}). \end{aligned} \qquad (2.11)$$

We notice that the effect of ignoring the break does not depend on the break time $c = n\lambda$ here.

Since

$$y_1^2 = x_1^2 - 2(1-\lambda)mx_1 + (1-\lambda)^2 m^2$$

and

$$y_n^2 = x_n^2 + 2\lambda m x_n + \lambda^2 m^2,$$

we can easily show that

$$\frac{1}{n^2}y_1^2 = \frac{1}{n^2}x_1^2 + O_p(\frac{1}{n\sqrt{n}}) \quad \text{and} \quad \frac{1}{n^2}y_n^2 = \frac{1}{n^2}x_n^2 + O_p(\frac{1}{n\sqrt{n}}). \qquad (2.12)$$

By (2.8) and (2.12), we have

$$\frac{1}{n^2}\{\sum_{t=2}^{n-1} y_t^2 + \frac{1}{2}(y_1^2 + y_n^2)\} = \frac{1}{n^2}\{\sum_{t=2}^{n-1} x_t^2 + \frac{1}{2}(x_1^2 + x_n^2)\} + O_p(\frac{1}{\sqrt{n}}) \qquad (2.13)$$

Combining (2.11) and (2.13),

$$\begin{aligned} n(\tilde{\rho}_s - 1) &= \frac{-\frac{1}{2n}\sum_{t=2}^{n}(y_t - y_{t-1})^2}{\frac{1}{n^2}\{\sum_{t=2}^{n-1} y_t^2 + \frac{1}{2}(y_1^2 + y_n^2)\}} \\ &= \frac{-\frac{1}{2n}\sum_{t=2}^{n}(x_t - x_{t-1})^2 + O_p(\frac{1}{n})}{\frac{1}{n^2}\{\sum_{t=2}^{n-1} x_t^2 + \frac{1}{2}(x_1^2 + x_n^2)\} + O_p(\frac{1}{\sqrt{n}})} \\ &= \frac{-\frac{1}{2n}\sum_{t=2}^{n}(x_t - x_{t-1})^2}{\frac{1}{n^2}\{\sum_{t=2}^{n-1} x_t^2 + \frac{1}{2}(x_1^2 + x_n^2)\}} + O_p(\frac{1}{\sqrt{n}}) \\ &\equiv n(\tilde{\rho}_{s,0} - 1) + O_p(\frac{1}{\sqrt{n}}) \end{aligned} \tag{2.14}$$

where $\tilde{\rho}_{s,0}$ stands for the SS estimator of ρ with no break.

As to the WS estimator of ρ, we recall that

$$\tilde{\rho}_w = \frac{\sum_{t=2}^{n} y_t y_{t-1}}{\sum_{t=2}^{n-1} y_t^2 + \frac{1}{n}\sum_{t=1}^{n} y_t^2}$$

and

$$\tilde{\rho}_w - 1 = \frac{\sum_{t=2}^{n} y_t y_{t-1} - \sum_{t=2}^{n-1} y_t^2 - \frac{1}{n}\sum_{t=1}^{n} y_t^2}{\sum_{t=2}^{n-1} y_t^2 + \frac{1}{n}\sum_{t=1}^{n} y_t^2} \equiv \frac{N_w}{D_w}. \tag{2.15}$$

By some straightforward algebra, the numerator of $\tilde{\rho}_w - 1$ can be written as

$$\sum_{t=2}^{n} y_t y_{t-1} - \sum_{t=2}^{n-1} y_t^2 - \frac{1}{n}\sum_{t=1}^{n} y_t^2 = -\frac{1}{2}\sum_{t=2}^{n}(y_t - y_{t-1})^2 + \frac{1}{2}(y_1^2 + y_n^2) - \frac{1}{n}\sum_{t=1}^{n} y_t^2.$$

By (2.12), we can show that

$$\frac{1}{n}y_1^2 = \frac{1}{n}x_1^2 + O_p(\frac{1}{\sqrt{n}}) \quad \text{and} \quad \frac{1}{n}y_n^2 = \frac{1}{n}x_n^2 + O_p(\frac{1}{\sqrt{n}}). \tag{2.16}$$

Therefore

$$\begin{aligned} \frac{1}{n}(\sum_{t=2}^{n} y_t y_{t-1} - \sum_{t=2}^{n-1} y_t^2 - \frac{1}{n}\sum_{t=1}^{n} y_t^2) &= \frac{1}{n}\{-\frac{1}{2}\sum_{t=2}^{n}(x_t - x_{t-1})^2 + \frac{1}{2}(x_1^2 + x_n^2) - \frac{1}{n}\sum_{t=1}^{n} x_t^2\} \\ &\quad + O_p(\frac{1}{\sqrt{n}}) \end{aligned} \tag{2.17}$$

by (2.8), (2.11) and (2.16).

On the other hand,

$$\frac{1}{n^2}(\sum_{t=2}^{n-1} y_t^2 + \frac{1}{n}\sum_{t=1}^{n} y_t^2) = \frac{1}{n^2}(\sum_{t=2}^{n-1} x_t^2 + \frac{1}{n}\sum_{t=1}^{n} x_t^2) + O_p(\frac{1}{\sqrt{n}}) \tag{2.18}$$

by (2.8). Combining (2.17) and (2.18),

$$
\begin{aligned}
n(\tilde{\rho}_w - 1) &= \frac{\frac{1}{n}(\sum_{t=2}^{n} y_t y_{t-1} - \sum_{t=2}^{n-1} y_t^2 - \frac{1}{n}\sum_{t=1}^{n} y_t^2)}{\frac{1}{n^2}(\sum_{t=2}^{n-1} y_t^2 + \frac{1}{n}\sum_{t=1}^{n} y_t^2)} \qquad (2.19)\\
&= \frac{\frac{1}{n}\{-\frac{1}{2}\sum_{t=2}^{n}(x_t - x_{t-1})^2 + \frac{1}{2}(x_1^2 + x_n^2) - \frac{1}{n}\sum_{t=1}^{n} x_t^2\} + O_p(\frac{1}{\sqrt{n}})}{\frac{1}{n^2}(\sum_{t=2}^{n-1} x_t^2 + \frac{1}{n}\sum_{t=1}^{n} x_t^2) + O_p(\frac{1}{\sqrt{n}})}\\
&= \frac{\frac{1}{n}(\sum_{t=2}^{n} x_t x_{t-1} - \sum_{t=2}^{n-1} x_t^2 - \frac{1}{n}\sum_{t=1}^{n} x_t^2) + O_p(\frac{1}{\sqrt{n}})}{\frac{1}{n^2}(\sum_{t=2}^{n-1} x_t^2 + \frac{1}{n}\sum_{t=1}^{n} x_t^2) + O_p(\frac{1}{\sqrt{n}})}\\
&= \frac{\frac{1}{n}(\sum_{t=2}^{n} x_t x_{t-1} - \sum_{t=2}^{n-1} x_t^2 - \frac{1}{n}\sum_{t=1}^{n} x_t^2)}{\frac{1}{n^2}(\sum_{t=2}^{n-1} x_t^2 + \frac{1}{n}\sum_{t=1}^{n} x_t^2)} + O_p(\frac{1}{\sqrt{n}})\\
&\equiv n(\tilde{\rho}_{w,0} - 1) + O_p(\frac{1}{\sqrt{n}})
\end{aligned}
$$

where $\tilde{\rho}_{w,0}$ stands for the WS estimator of ρ with no break.

Thus the effect of an added break is $O_p(1/\sqrt{n})$ in all cases with the numerator of $n(\tilde{\rho}_s - 1)$ having an even faster convergence rate, $O_p(1/n)$.

The studentized test statistics can be written as

$$
\begin{aligned}
\tau_\mu &= \frac{\hat{\rho}_\mu - 1}{\text{s.e.}} = \frac{n(\hat{\rho}_\mu - 1)\sqrt{\frac{1}{n^2}\sum_{t=2}^{n} y_{t-1}^2}}{\sqrt{s^2}} \qquad (2.20)\\
&= \frac{\{n(\hat{\rho}_{\mu,0} - 1) + O_p(\frac{1}{\sqrt{n}})\}\sqrt{\frac{1}{n^2}\sum_{t=2}^{n} x_{t-1}^2 + O_p(\frac{1}{\sqrt{n}})}}{\sqrt{s^2}}\\
&\equiv \tau_{\mu,0} + O_p(\frac{1}{\sqrt{n}}),
\end{aligned}
$$

$$
\begin{aligned}
\tau_s &= \frac{\tilde{\rho}_s - 1}{\text{s.e.}} = \frac{n(\tilde{\rho}_s - 1)\sqrt{\frac{1}{n^2}\{\sum_{t=2}^{n-1} y_t^2 + \frac{1}{2}(y_1^2 + y_n^2)\}}}{\sqrt{\tilde{\sigma}_s^2}} \qquad (2.21)\\
&= \frac{\{n(\tilde{\rho}_{s,0} - 1) + O_p(\frac{1}{\sqrt{n}})\}\sqrt{\frac{1}{n^2}\{\sum_{t=2}^{n-1} x_t^2 + \frac{1}{2}(x_1^2 + x_n^2)\} + O_p(\frac{1}{\sqrt{n}})}}{\sqrt{\tilde{\sigma}_s^2}}\\
&\equiv \tau_{s,0} + O_p(\frac{1}{\sqrt{n}})
\end{aligned}
$$

and

$$
\tau_w = \frac{\tilde{\rho}_w - 1}{\text{s.e.}} = \frac{n(\tilde{\rho}_w - 1)\sqrt{\frac{1}{n^2}(\sum_{t=2}^{n-1} y_t^2 + \frac{1}{n}\sum_{t=1}^{n} y_t^2)}}{\sqrt{\tilde{\sigma}_w^2}} \qquad (2.22)
$$

$$
\begin{aligned}
&= \frac{\{n(\tilde{\rho}_{w,0}-1)+O_p(\frac{1}{\sqrt{n}})\}\sqrt{\frac{1}{n^2}(\sum_{t=2}^{n-1}x_t^2+\frac{1}{n}\sum_{t=1}^{n}x_t^2)+O_p(\frac{1}{\sqrt{n}})}}{\sqrt{\tilde{\sigma}_w^2}} \\
&\equiv \tau_{w,0}+O_p(\frac{1}{\sqrt{n}})
\end{aligned}
$$

where $\tau_{\mu,0}$, $\tau_{s,0}$ and $\tau_{w,0}$ represent the test statistics with no break. We use the fact that s^2, $\tilde{\sigma}_s^2$ and $\tilde{\sigma}_w^2$ are consistent regardless of the break under $\rho = 1$.

Although the effect of ignoring the break in the numerator of $n(\tilde{\rho}_s - 1)$ is $O_p(\frac{1}{n})$, the effect is $O_p(\frac{1}{\sqrt{n}})$ in all of the studentized test statistics. The spurious rejection problem disappears at the same rate for all tests, as n becomes larger.

2.2.1.2 Analysis of expectations

We next analyze the expectations of the quadratic forms constituting the test statistics to gain some more clues with respect to the empirical results. These are only clues, as the power also depends on the shapes of the distributions' tails.

Define $\hat{\rho}_\mu - 1 \equiv N_\mu/D_\mu$, $\tilde{\rho}_s - 1 \equiv N_s/D_s$ and $\tilde{\rho}_w - 1 \equiv N_w/D_w$. Some straightforward but tedious algebra gives

$$
\begin{aligned}
\mathrm{E}[N_\mu] &= n\sigma^2[-\frac{1}{2}+\frac{1}{2n}-\frac{m^2(1-\lambda)}{n}] \\
\mathrm{E}[D_\mu] &= n^2\sigma^2[\frac{1}{6}-\frac{1}{3n}+\frac{1}{3n^2}-\frac{1}{6n^3}+\frac{m^2\lambda(1-\lambda)}{n}-\frac{m^2\lambda^2}{n^2}] \\
\mathrm{E}[N_s] &= n\sigma^2(-\frac{1}{2}+\frac{1}{2n}-\frac{m^2}{2n}) \\
\mathrm{E}[D_s] &= n^2\sigma^2[\frac{1}{6}-\frac{1}{3n}+\frac{1}{3n^2}-\frac{1}{6n^3}+\frac{m^2\lambda(1-\lambda)}{n}-\frac{m^2\lambda^2+m^2(1-\lambda)^2}{2n^2}] \\
\mathrm{E}[N_w] &= n\sigma^2[-\frac{1}{3}+\frac{1}{3n^2}-\frac{2m^2\lambda(1-\lambda)}{n}] \\
\mathrm{E}[D_w] &= n^2\sigma^2[(\frac{1}{6}-\frac{2}{3n}+\frac{5}{6n^2}-\frac{1}{3n^3})(\frac{n+1}{n})+\frac{2}{3n^2}-\frac{1}{n^3}+\frac{1}{3n^4} \\
&\quad +\frac{m^2\lambda(1-\lambda)(n+1)-m^2\lambda^2-m^2(1-\lambda)^2}{n^2}]
\end{aligned}
$$

where $\lambda = c/n$ is the proportion of observations before the break.

Figure 2.3 presents the expectations of the quadratic forms, their ratios and pivotal statistics constructed by replacing quadratic forms with their expectations.

The expectations of the denominators above are concave functions of λ and symmetric around $\lambda = 0.5$. They are very similar as we can see in the second column of Figure 2.3. This is expected because the denominators can be written as

$$\begin{aligned} D_\mu &= \sum_{t=2}^{n-1} y_t^2 + y_1^2 \\ D_s &= \sum_{t=2}^{n-1} y_t^2 + \frac{1}{2}(y_1^2 + y_n^2) \\ D_w &= \sum_{t=2}^{n-1} y_t^2 + \frac{1}{n}\sum_{t=1}^{n} y_t^2 \end{aligned}$$

The major differences are in the numerators. $\mathrm{E}[N_\mu]$ is an increasing function of λ and $\mathrm{E}[N_s]$ does not depend on λ. $\mathrm{E}[N_w]$ is a convex function of λ and symmetric around $\lambda = 0.5$. Notice that lower values of the numerators suggest higher rates of rejecting $H_0 : \rho = 1$ as we are using left tailed tests.

If we consider the ratio of the expectations of the numerators and the denominators, $\mathrm{E}[N_\mu]/\mathrm{E}[D_\mu]$ is a concave function of λ which is negative and not symmetric around $\lambda = 0.5$. Its absolute value is larger for λ near 0 than for λ near 1. As λ increases from 0 to 1, its absolute value decreases at first and then increases. This corresponds to the empirical results of high τ_μ rejection rates for early breaks.

$\mathrm{E}[N_s]/\mathrm{E}[D_s]$ is a symmetric concave function of λ. Its absolute values are larger for λ near 0 or 1 than for λ near 0.5. $\mathrm{E}[N_w]/\mathrm{E}[D_w]$ is also a symmetric concave function of λ, becoming nearly constant as n increases. The absolute values of $\mathrm{E}[N_s]/\mathrm{E}[D_s]$ are much larger than those of $\mathrm{E}[N_w]/\mathrm{E}[D_w]$ for λ near 0 or 1.

This also corresponds to the empirical results that τ_s and τ_w show symmetry around $\lambda = 0.5$ with τ_s giving higher rejection rates than τ_w for λ near 0 or 1.

All three standard errors in (2.3), (2.4) and (5.1) are of the form

$$\text{s.e.} = \sqrt{\frac{\text{MSE}}{D}}$$

where MSE stands for the mean square error and D denotes D_μ, D_s or D_w. Recalling that the mean square error is a consistent estimator of σ^2, we define an approximate standard error, a.s.e., as

$$\text{a.s.e.} \equiv \sqrt{\frac{\sigma^2}{\mathrm{E}[D]}}.$$

We find that a.s.e. is a symmetric convex function of λ and that $(\mathrm{E}[N_\mu]/\mathrm{E}[D_\mu])/\text{a.s.e.}$ and $(\mathrm{E}[N_s]/\mathrm{E}[D_s])/\text{a.s.e.}$ show similar shapes to $\mathrm{E}[N_\mu]/\mathrm{E}[D_\mu]$ and $\mathrm{E}[N_s]/\mathrm{E}[D_s]$ respectively.

On the other hand, $(\mathrm{E}[N_w]/\mathrm{E}[D_w])/\text{a.s.e.}$ has a little different shape from that of $\mathrm{E}[N_w]/\mathrm{E}[D_w]$. The former is convex whereas the latter is nearly flat but concave. These graphs are displayed as 'pivotal statistics' in the right column of Figure 2.3.

2.2.2 Using linear trend adjusted statistics

Leybourne et al (1998) also considered $\hat{\rho}_\tau$ which is the OLS estimator of ρ in the model

$$Y_t = \mu + \beta t + \rho Y_{t-1} + e_t. \tag{2.23}$$

Alternatively the first order autoregressive model around a linear trend can be written as

$$Y_t - \mu - \beta t = \rho[Y_{t-1} - \mu - \beta(t-1)] + e_t. \tag{2.24}$$

By simple algebra, the model (2.24) becomes

$$\begin{aligned} Y_t &= [\mu(1-\rho) + \beta\rho] + \beta(1-\rho)t + \rho Y_{t-1} + e_t \\ &\equiv \gamma_0 + \gamma_1 t + \rho Y_{t-1} + e_t \end{aligned}$$

which is the same model as (2.23). We denote the pivotal statistic associated with $\hat{\rho}_\tau$ as τ_τ. $\hat{\rho}_\tau$ and τ_τ can be obtained if we replace $(Y_t - \bar{Y})$ with $(Y_t - \hat{a} - \hat{b}t)$ in (2.1) and (2.3), respectively, where $\hat{a}$ and $\hat{b}$ are OLS estimates of a and b in the simple linear regression

$$Y_t = a + bt + e_t. \tag{2.25}$$

Leybourne et al (1998) found that τ_τ gives a similar size distortion pattern to that of τ_μ when Model I is used to generate the data. Our empirical sizes for τ_τ in Table 2.2 are in close agreement with theirs. The magnitude of the spurious rejection problem is more severe for τ_τ than for τ_μ for the smallest λ. The problem evaporates more rapidly for τ_τ than for τ_μ as λ increases.

Denote the linear trend adjusted versions of τ_s and τ_w as $\tau_{s,\tau}$ and $\tau_{w,\tau}$ respectively. $\tau_{s,\tau}$ and $\tau_{w,\tau}$ can be obtained from (2.4) and (2.5), respectively, by replacing $(Y_t - \bar{Y})$ with the residual from (2.25). Empirical sizes for $\tau_{s,\tau}$ and $\tau_{w,\tau}$ are given in Table 2.2. In the simulation, the number of replications is 5,000 and the sample size per replication is $n = 100$. Critical values used are those for the linear trend removed case in Fuller (1996).

As in the mean adjusted cases in section 2.2.1, the spurious rejection problem is less severe for $\tau_{s,\tau}$ and $\tau_{w,\tau}$ than for τ_τ. Also $\tau_{s,\tau}$ and $\tau_{w,\tau}$ show similar patterns of empirical sizes to τ_s and τ_w, respectively. $\tau_{s,\tau}$ gives a slightly higher rejection rate than τ_s. Like τ_w, $\tau_{w,\tau}$ gives satisfactory performance and retains size close to the nominal 5% level. Figure 2.4 displays the empirical sizes for model I with $m = 10$ for the linear trend adjusted statistics.

2.2.3 Using lagged first differences

To extend the results for the first order process to more general processes, we consider an augmented test based on the WS estimator since it shows better performance than the other estimators in the previous results. We introduce lagged first differences as in the usual Augmented Dickey-Fuller test based on the OLS estimator. Our test statistic is denoted as $\tau_{w,a}$.

Using the data arrangement and weights in Table 10.1.1, Fuller (1996), with $p = 2$, we perform the weighted regression estimation of the autoregressive model written as

$$Y_t = \theta_1 Y_{t-1} + \theta_2 Z_{t-1} + e_t$$

where $Z_{t-1} = Y_{t-1} - Y_{t-2}$. For reference, we reproduce the table for $p = 2$ here.

Weight	Dependent Variable	θ_1	θ_2
w_3	Y_3	Y_2	$Y_2 - Y_1$
w_4	Y_4	Y_3	$Y_3 - Y_2$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
w_n	Y_n	Y_{n-1}	$Y_{n-1} - Y_{n-2}$
1-w_{n-1}	Y_{n-2}	Y_{n-1}	$Y_{n-1} - Y_n$
1-w_{n-2}	Y_{n-3}	Y_{n-2}	$Y_{n-2} - Y_{n-1}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
1-w_2	Y_1	Y_2	$Y_2 - Y_3$

The estimator of $\boldsymbol{\theta} = (\theta_1, \theta_2)'$ can be obtained as

$$\hat{\boldsymbol{\theta}} = (\tilde{\theta}_1, \tilde{\theta}_2)' = (\boldsymbol{X}'\boldsymbol{W}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{W}\boldsymbol{Y}$$

where $\boldsymbol{X}$ is the $(2n - 2p) \times p$ matrix below the headings θ_1, θ_2 in the table, $\boldsymbol{Y}$ is the $(2n - 2p)$-dimensional column vector called the dependent variable and $\boldsymbol{W}$ is the $(2n - 2p)$ diagonal matrix whose elements are given in the "Weight" column. The hypothesis of a unit root is tested by testing the hypothesis that $\theta_1 = 1$. Then the test statistic $\tau_{w,a}$ can be written as

$$\tau_{w,a} = \frac{\tilde{\theta}_1 - 1}{\text{s.e.}} \tag{2.26}$$

whose limiting distribution, using the proper s.e., is the same as that of τ_w in (5.1). Because each observation gives rise to 2 rows in Fuller's Table, the standard error printed by a typical regression package needs to be multiplied by $\sqrt{2}$ in (2.26) in order to use the critical values from Fuller (1996).

Empirical sizes for Model I as a DGP are shown in Table 2.3 and Figure 2.5. The same values of m and c as before are used. The number of replications is 1,000 and the sample size is $n = 100$. As a whole, the performance of the WS estimator in the augmented case is as good as in the simple case. In particular, $\tau_{w,a}$ maintains size close to the nominal 5% level regardless of the break size.

2.3 Data with a break in trend

In this section, we consider another DGP

$$\text{Model II} \quad : \quad Y_t = Y_{t-1} + m\sigma I(t > c) + e_t, \qquad t = 1, 2, ..., n$$

where the e_t are normal independent $(0,\sigma^2)$ random variables and we can assume $\sigma = 1$ without loss of generality as in Model I. This model corresponds to 'Model (B) under the null hypothesis' of Perron (1989). Since Model II can be rewritten as

$$\begin{aligned} Y_t &= Y_{t-1} + e_t, & t &= 1, \cdots, c \\ Y_t &= Y_{t-1} + m + e_t \\ &= Y_c + m(t-c) + \sum_{j=c+1}^{t} e_j, & t &= c+1, \cdots, n, \end{aligned} \tag{2.27}$$

it might be called a random walk with a break in drift where $mI(t > c)$ is the drift parameter. When $m \neq 0$, the time trend will dominate the long-run behavior of Y_t in (2.27) because $\sum_{j=c+1}^{t} e_j = O_p(\sqrt{n})$ and hence $\sum_{j=c+1}^{t} e_j$ is small in probability relative to t. Therefore we consider the linear trend adjusted versions of statistics to test the unit root null hypothesis. Time series data generated from Model II with various values of m are displayed in Figure 2.6.

Leybourne et al (1998) reported empirical sizes for Model II using the test statistic τ_τ. Columns labelled τ_τ in Table 2.4 contain our empirical sizes for τ_τ, which are close to those of Leybourne et al (1998). They found that, as m increases, the range of values of λ for which serious problems are found increases. They also concluded that frequent spurious rejections of the null hypothesis occur for low values of λ but the pattern for Model II is a little different than that for Model I. The spurious rejection problem was most severe for the lowest values of λ in Model I. In contrast, for Model II, it increases at first in severity with increasing λ, but then rapidly diminishes as λ increases further.

We use test statistics $\tau_{s,\tau}$ and $\tau_{w,\tau}$ to obtain the proportion of rejections of the unit root null hypothesis. We consider the same values of m and c as those used by Leybourne et al (1998). These results are also presented in Table 2.4. We notice that, for Model II, $\tau_{s,\tau}$ and $\tau_{w,\tau}$ give similar patterns of empirical sizes to each other. Although they do not show size distortion problems, they are severely undersized for a wide range of c which might suggest lower power against the alternative hypothesis. We comment on this in section 2.4. As far as spurious rejection is concerned, $\tau_{w,\tau}$ again presents much better empirical sizes than τ_τ. Empirical sizes for Model II with $m = 2$ are shown in Figure 2.7.

2.4 Empirical powers allowing for a break

Thus far we have considered the effects of ignoring an existing break and have found the weighted symmetric estimator performs well, avoiding spurious declararations of stationarity. If a series consists of a level shift plus stationary errors and one tests for stationarity without modelling the level shift, it is unclear whether accepting or rejecting the null hypothesis is desirable - neither is correct so power should not be a concern here. However, if we modified the WS estimator to accomodate a model with a break, we would want that test to have good power. We investigate this now.

Let D_t be a dummy variable representing a structural break, that is $D_t = I(t > c)$. We first perform regression estimation of the model

$$Y_t = \beta_0 + \beta_1 D_t + e_t$$

and produce the residuals $r_t = Y_t - \hat{Y}_t = Y_t - \hat{\beta}_0 - \hat{\beta}_1 D_t$. Replacing $(Y_t - \bar{Y})$ with r_t in (2.5) and (2.6), we obtain a WS estimator and associated pivotal statistic, respectively, allowing for a break. For the limiting distribution of these statistics, see Appendix A.

Table 2.5 presents empirical size and power results with respect to Model I as a DGP. Various values of n, λ and ρ are used where $\lambda = c/n$ is the break time in D_t. Although we are assuming a known break, it makes no difference in the results for the modified estimator whether there is a break or not in the DGP. Both critical values from a simulation method and from Fuller (1996) are applied. The number of replications is 1,000.

For all values of n, empirical sizes ($\rho = 1$) are too big if we use the critical values from Fuller (1996). This is no surprise since most previous research on trend breaks shows that critical values must be adjusted for breaks. For instance, Perron (1989) showed that the critical values under various models are noticeably smaller than the standard Dickey-Fuller critical value.

The values for τ_w with no fitted break are shown under the label $\lambda = 0$ in Table 2.5. Previous researchers have used least squares estimation for trend break cases and, in Table 2.5, we include with underlines the powers obtained using OLS estimation. See Tables 4-6 in Pantula et al (1994) for the empirical powers of τ_μ and τ_w when no break is present.

Using critical values from simulation, we obtain reasonably good empirical powers. For given n and ρ, as λ changes from .02 to .5, the power decreases. In general, when there really is no break, the modified test that is invariant to breaks gives less power than τ_w with no fitted break. When there is a break ($\lambda > 0$), the modified test based on the WS estimator gives better power than the modified test based on the OLS estimator. It is seen that substantial improvement in power occurs with the WS estimator especially for larger λ values.

2.5 Conclusion

In this paper, we compare the empirical sizes of unit root tests based on the OLS estimator with those of tests based on alternative symmetric estimators. We find that the symmetric estimators have much less severe size distortion problems than the OLS estimator in both the random walk with a break in level (Model I) and the random walk with a break in trend (Model II).

The WS estimator in particular shows practically no size distortion and, even for large m, nearly retains the nominal level. The performance of the WS estimator in the augmented case is as good as in the simple case. The WS estimator also generates reasonably good empirical power when modified to allow for a break.

Obviously we do not want to reject the unit root null hypothesis when it is true but there is a break. Therefore we recommend the use of tests based on the WS estimator to avoid the 'converse Perron phenomenon'. Further, if the break is modeled, the WS estimator has good power properties.

Table 2.1: Empirical sizes for Model I using mean adjusted statistics

			m=2.5			m=5			m=10
c	τ_μ	τ_s	τ_w	τ_μ	τ_s	τ_w	τ_μ	τ_s	τ_w
1	0.0906	0.0574	0.0486	0.2174	0.0926	0.0526	0.5444	0.2562	0.0514
2	0.0946	0.0626	0.0510	0.2008	0.0866	0.0564	0.5104	0.1964	0.0476
3	0.0716	0.0486	0.0476	0.1674	0.0672	0.0426	0.4696	0.1394	0.0382
4	0.0740	0.0518	0.0518	0.1584	0.0658	0.0498	0.4308	0.1104	0.0404
5	0.0716	0.0522	0.0520	0.1378	0.0596	0.0484	0.3828	0.0778	0.0362
6	0.0638	0.0528	0.0534	0.1206	0.0566	0.0490	0.3426	0.0650	0.0378
7	0.0602	0.0516	0.0532	0.1118	0.0520	0.0502	0.2860	0.0532	0.0344
8	0.0610	0.0514	0.0506	0.0946	0.0498	0.0488	0.2314	0.0456	0.0354
9	0.0622	0.0508	0.0504	0.0882	0.0490	0.0476	0.1864	0.0424	0.0364
10	0.0592	0.0478	0.0460	0.0796	0.0436	0.0432	0.1398	0.0340	0.0302
12	0.0560	0.0558	0.0530	0.0648	0.0478	0.0500	0.0898	0.0342	0.0382
14	0.0526	0.0516	0.0498	0.0558	0.0476	0.0472	0.0548	0.0312	0.0366
16	0.0502	0.0520	0.0536	0.0468	0.0476	0.0526	0.0398	0.0252	0.0316
18	0.0456	0.0464	0.0478	0.0430	0.0486	0.0504	0.0264	0.0308	0.0374
20	0.0484	0.0504	0.0496	0.0408	0.0462	0.0490	0.0266	0.0288	0.0338
50	0.0470	0.0566	0.0576	0.0372	0.0500	0.0530	0.0156	0.0266	0.0370
90	0.0524	0.0536	0.0538	0.0388	0.0488	0.0510	0.0158	0.0370	0.0348
95	0.0420	0.0540	0.0538	0.0350	0.0620	0.0496	0.0182	0.0840	0.0358
99	0.0472	0.0660	0.0514	0.0422	0.1020	0.0496	0.0288	0.2550	0.0504

Table 2.2: Empirical sizes for Model I using linear trend adjusted statistics

			m=2.5			m=5			m=10
c	τ_τ	$\tau_{s,\tau}$	$\tau_{w,\tau}$	τ_τ	$\tau_{s,\tau}$	$\tau_{w,\tau}$	τ_τ	$\tau_{s,\tau}$	$\tau_{w,\tau}$
1	0.1058	0.0670	0.0500	0.2854	0.1148	0.0460	0.7224	0.3466	0.0424
2	0.1026	0.0714	0.0600	0.2388	0.0980	0.0554	0.6632	0.2206	0.0362
3	0.0760	0.0538	0.0516	0.1850	0.0728	0.0474	0.5580	0.1258	0.0304
4	0.0744	0.0582	0.0526	0.1442	0.0622	0.0464	0.4452	0.0784	0.0300
5	0.0628	0.0544	0.0538	0.1118	0.0546	0.0502	0.3142	0.0556	0.0294
6	0.0548	0.0498	0.0476	0.0812	0.0478	0.0450	0.1968	0.0404	0.0298
7	0.0552	0.0522	0.0524	0.0738	0.0490	0.0478	0.1164	0.0364	0.0328
8	0.0514	0.0522	0.0540	0.0564	0.0470	0.0446	0.0634	0.0260	0.0238
9	0.0484	0.0522	0.0520	0.0530	0.0542	0.0554	0.0460	0.0312	0.0328
10	0.0536	0.0556	0.0546	0.0496	0.0500	0.0504	0.0320	0.0264	0.0290
12	0.0510	0.0556	0.0532	0.0472	0.0484	0.0528	0.0234	0.0274	0.0314
14	0.0522	0.0552	0.0528	0.0454	0.0516	0.0534	0.0228	0.0302	0.0362
16	0.0436	0.0542	0.0540	0.0378	0.0494	0.0496	0.0232	0.0302	0.0376
18	0.0446	0.0480	0.0488	0.0396	0.0440	0.0446	0.0204	0.0284	0.0352
20	0.0468	0.0512	0.0516	0.0444	0.0476	0.0504	0.0250	0.0312	0.0374
50	0.0566	0.0608	0.0624	0.0492	0.0542	0.0564	0.0238	0.0298	0.0358
90	0.0512	0.0534	0.0514	0.0428	0.0484	0.0498	0.0198	0.0282	0.0302
95	0.0508	0.0562	0.0536	0.0418	0.0588	0.0474	0.0206	0.0586	0.0332
99	0.0512	0.0752	0.0564	0.0428	0.1204	0.0490	0.0306	0.3496	0.0404

Table 2.3: Empirical sizes for Model I using lagged first differences (Weighted Symmetric estimator)

c	m=2.5	m=5	m=10
1	0.042	0.044	0.043
2	0.038	0.047	0.045
3	0.041	0.050	0.045
4	0.044	0.057	0.054
5	0.047	0.056	0.052
6	0.047	0.059	0.039
7	0.038	0.052	0.042
8	0.043	0.051	0.040
9	0.046	0.050	0.048
10	0.043	0.050	0.046
12	0.042	0.055	0.039
14	0.043	0.058	0.045
16	0.047	0.049	0.039
18	0.043	0.046	0.044
20	0.045	0.054	0.045
50	0.056	0.054	0.028
90	0.046	0.035	0.026
95	0.044	0.046	0.032
99	0.042	0.047	0.037

Table 2.4: Empirical sizes for Model II using linear trend adjusted statistics

	m=0.5			m=1			m=2		
c	τ_τ	$\tau_{s,\tau}$	$\tau_{w,\tau}$	τ_τ	$\tau_{s,\tau}$	$\tau_{w,\tau}$	τ_τ	$\tau_{s,\tau}$	$\tau_{w,\tau}$
1	0.0462	0.0518	0.0520	0.0462	0.0518	0.0520	0.0462	0.0518	0.0520
2	0.0628	0.0620	0.0612	0.0674	0.0630	0.0604	0.0956	0.0696	0.0580
4	0.0634	0.0512	0.0504	0.1076	0.0504	0.0418	0.2810	0.0556	0.0214
6	0.0766	0.0526	0.0422	0.1638	0.0434	0.0256	0.5024	0.0172	0.0012
8	0.0874	0.0458	0.0390	0.2172	0.0262	0.0126	0.6878	0.0010	0.0002
10	0.0914	0.0392	0.0304	0.2624	0.0146	0.0074	0.8092	0.0002	0.0000
12	0.0978	0.0346	0.0268	0.2766	0.0044	0.0020	0.8646	0.0000	0.0000
14	0.0878	0.0262	0.0216	0.2648	0.0036	0.0016	0.8838	0.0000	0.0000
16	0.0774	0.0244	0.0178	0.2620	0.0012	0.0008	0.8900	0.0000	0.0000
18	0.0758	0.0214	0.0196	0.2260	0.0004	0.0004	0.8556	0.0000	0.0000
20	0.0702	0.0160	0.0130	0.1948	0.0004	0.0006	0.8110	0.0000	0.0000
22	0.0596	0.0106	0.0090	0.1532	0.0000	0.0000	0.7198	0.0000	0.0000
24	0.0522	0.0128	0.0112	0.1232	0.0004	0.0002	0.6076	0.0000	0.0000
26	0.0482	0.0104	0.0096	0.0876	0.0000	0.0000	0.4618	0.0000	0.0000
28	0.0350	0.0088	0.0078	0.0574	0.0000	0.0000	0.3090	0.0000	0.0000
30	0.0294	0.0076	0.0072	0.0352	0.0000	0.0000	0.1688	0.0000	0.0000
32	0.0246	0.0066	0.0068	0.0208	0.0000	0.0000	0.0810	0.0000	0.0000
34	0.0208	0.0054	0.0050	0.0126	0.0000	0.0000	0.0310	0.0000	0.0000
36	0.0194	0.0072	0.0064	0.0072	0.0000	0.0000	0.0114	0.0000	0.0000
38	0.0142	0.0048	0.0026	0.0052	0.0000	0.0000	0.0042	0.0000	0.0000
40	0.0158	0.0056	0.0050	0.0044	0.0000	0.0000	0.0020	0.0000	0.0000
50	0.0074	0.0032	0.0030	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
90	0.0244	0.0298	0.0242	0.0036	0.0086	0.0040	0.0000	0.0002	0.0000
95	0.0416	0.0512	0.0430	0.0228	0.0442	0.0256	0.0030	0.0188	0.0026
99	0.0494	0.0532	0.0500	0.0496	0.0554	0.0508	0.0478	0.0610	0.0504

Table 2.5: Empirical powers for Model I using the modified test

				n=	50				n=	100
ρ	λ=0	.02	.05	.25	.50	λ=0	.02	.05	.25	.50
CV*	-2.57	-2.89	-2.94	-3.26	-3.31	-2.55	-2.85	-2.95	-3.18	-3.26
1.00	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
	0.05†	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
0.95	0.12	0.10	0.10	0.08	0.09	0.26	0.17	0.15	0.13	0.14
	0.07	0.10	0.10	0.07	0.07	0.12	0.17	0.15	0.11	0.10
0.90	0.25	0.18	0.16	0.13	0.14	0.60	0.45	0.39	0.33	0.32
	0.12	0.18	0.16	0.10	0.09	0.31	0.45	0.39	0.25	0.22
0.80	0.58	0.45	0.42	0.32	0.32	0.98	0.94	0.91	0.85	0.83
	0.32	0.45	0.42	0.25	0.22	0.88	0.94	0.91	0.75	0.68
0.50		0.99	0.99	0.97	0.96		1.00	1.00	1.00	1.00
		0.99	0.99	0.94	0.90		1.00	1.00	1.00	1.00
CV‡	-2.57					-2.55				
1.00		0.10	0.11	0.19	0.23		0.11	0.12	0.20	0.23
0.95		0.21	0.21	0.29	0.35		0.32	0.33	0.41	0.48
0.90		0.33	0.33	0.42	0.49		0.66	0.66	0.74	0.77
0.80		0.66	0.67	0.73	0.76		0.99	0.99	0.99	0.99
0.50		1.00	1.00	1.00	1.00		1.00	1.00	1.00	1.00

*Critical values are from Fuller (1996) for λ=0 or from simulation otherwise.

†Underlined values are from the OLS estimator.

‡Critical values are from Fuller (1996).

Table 2.5: *continued*

ρ	λ=0	.02	.05	n= .25	250 .50
CV*	-2.54	-2.82	-2.87	-3.11	-3.20
1.00	0.05	0.05	0.05	0.05	0.05
	0.05†	0.05	0.05	0.05	0.05
0.95	0.78	0.61	0.58	0.49	0.46
	0.44	0.61	0.57	0.35	0.30
0.90	0.99	0.99	0.99	0.97	0.96
	0.97	0.99	0.99	0.90	0.86
0.80	1.00	1.00	1.00	1.00	1.00
	1.00	1.00	1.00	1.00	1.00
0.50		1.00	1.00	1.00	1.00
		1.00	1.00	1.00	1.00
CV‡	-2.54				
1.00		0.10	0.12	0.19	0.22
0.95		0.80	0.80	0.85	0.88
0.90		1.00	1.00	1.00	1.00
0.80		1.00	1.00	1.00	1.00
0.50		1.00	1.00	1.00	1.00

*Critical values are from Fuller (1996) for λ=0 or from simulation otherwise.
†Underlined values are from the OLS estimator.
‡Critical values are from Fuller (1996).

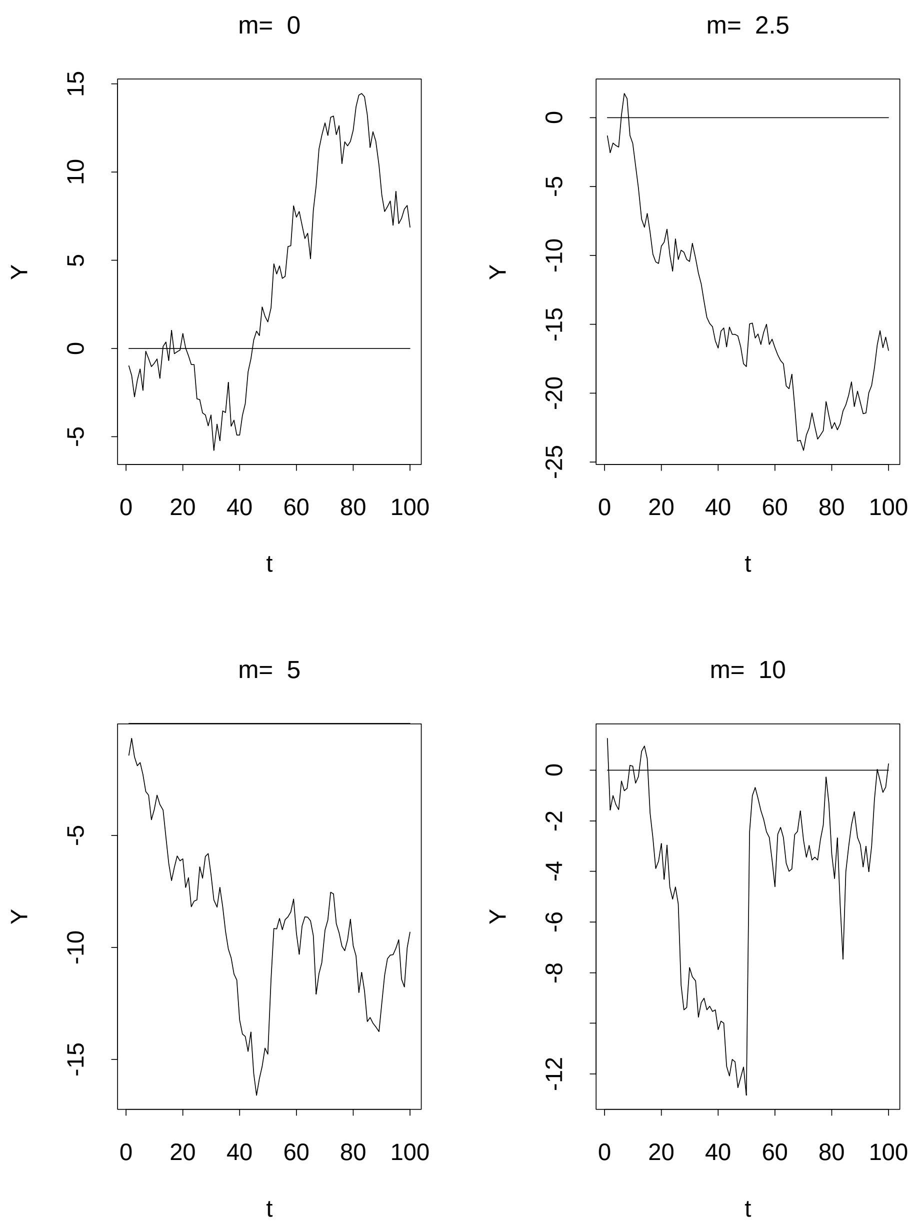

Figure 2.1: Random walk with a break in level (Model I, $c = 50$)

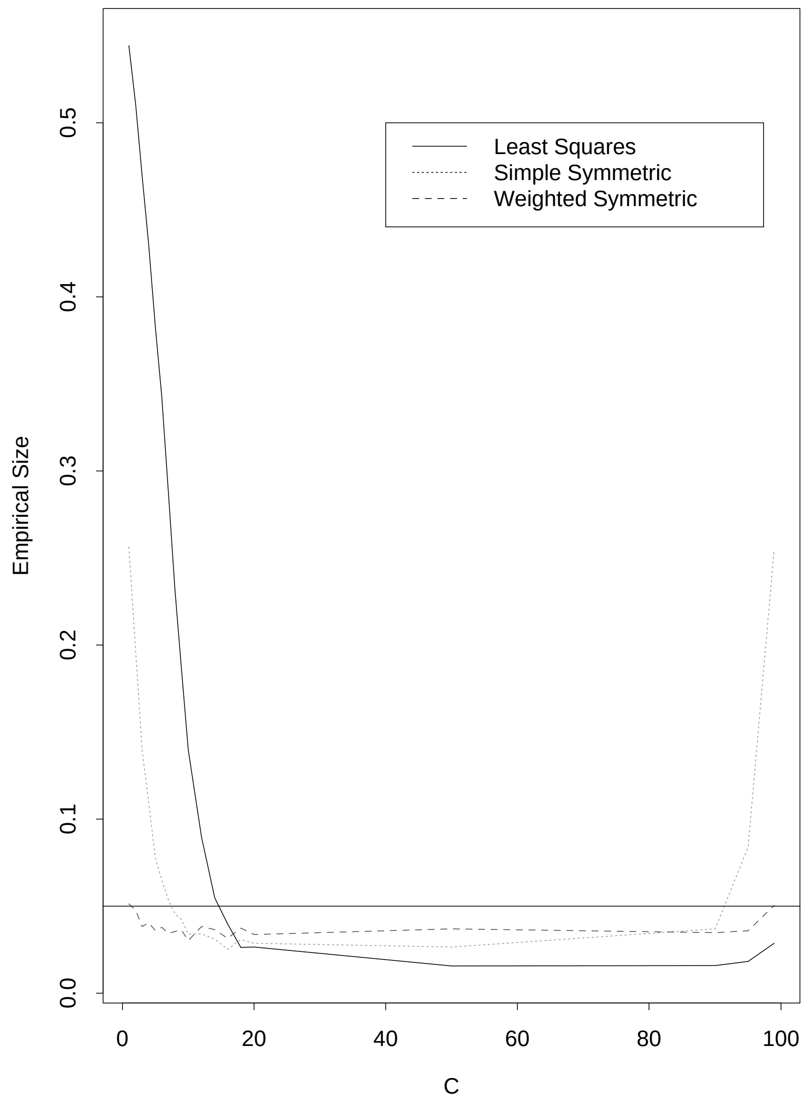

Figure 2.2: Empirical sizes for Model I with $m = 10$ using mean adjusted statistics

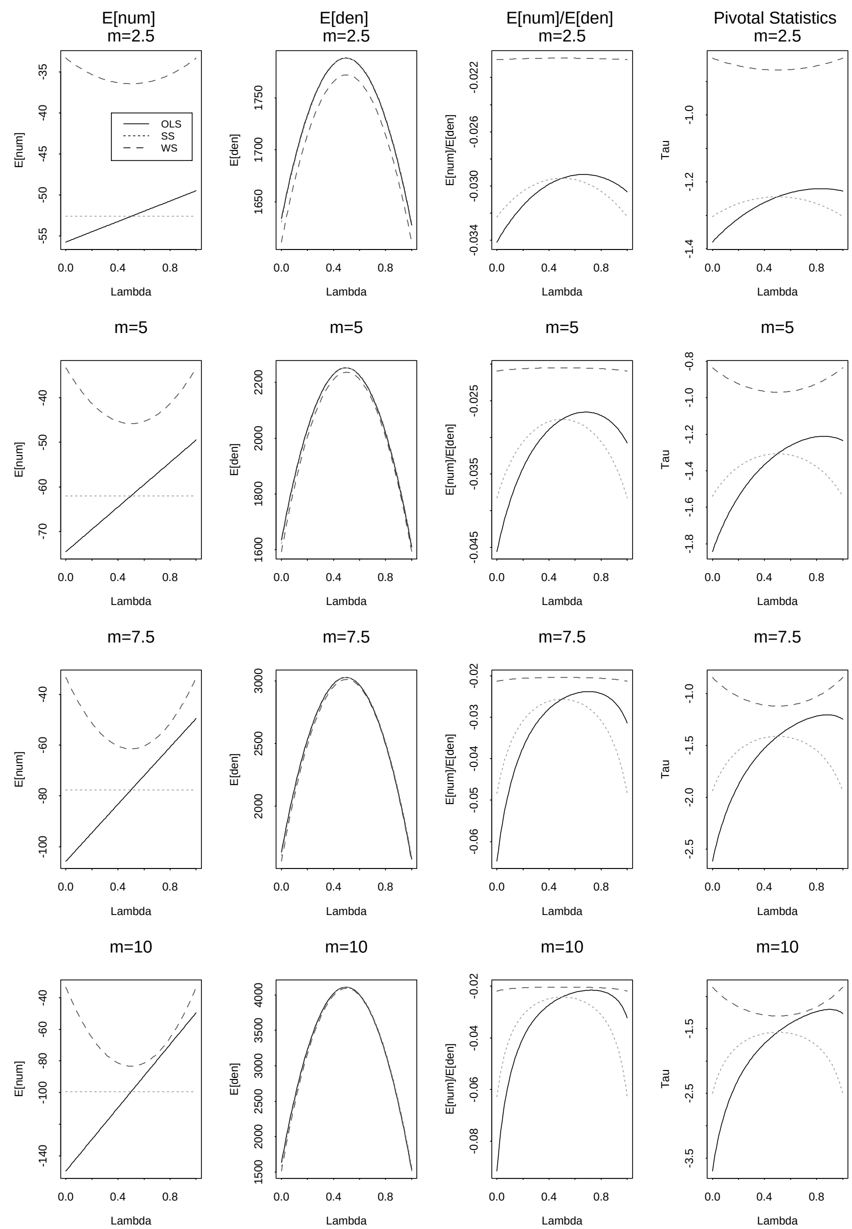

Figure 2.3: Expectations of quadratic forms, their ratios and pivotal statistics ($n = 100$)

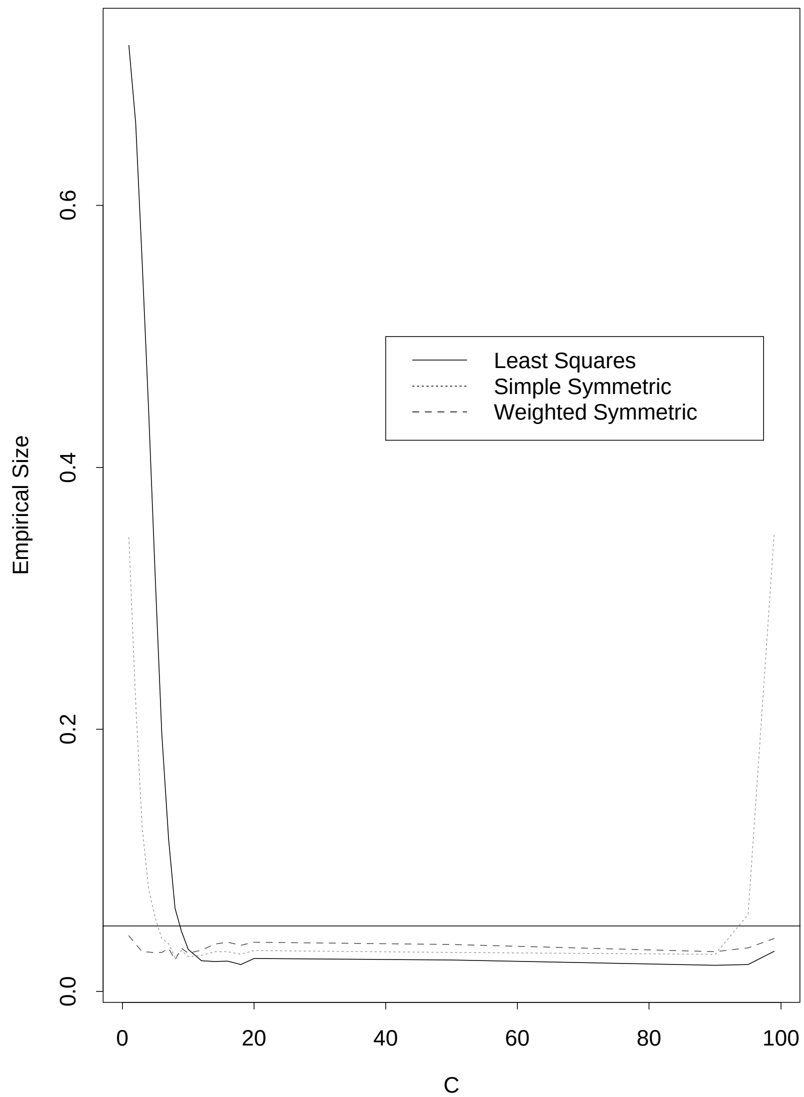

Figure 2.4: Empirical sizes for Model I with $m = 10$ using linear trend adjusted statistics

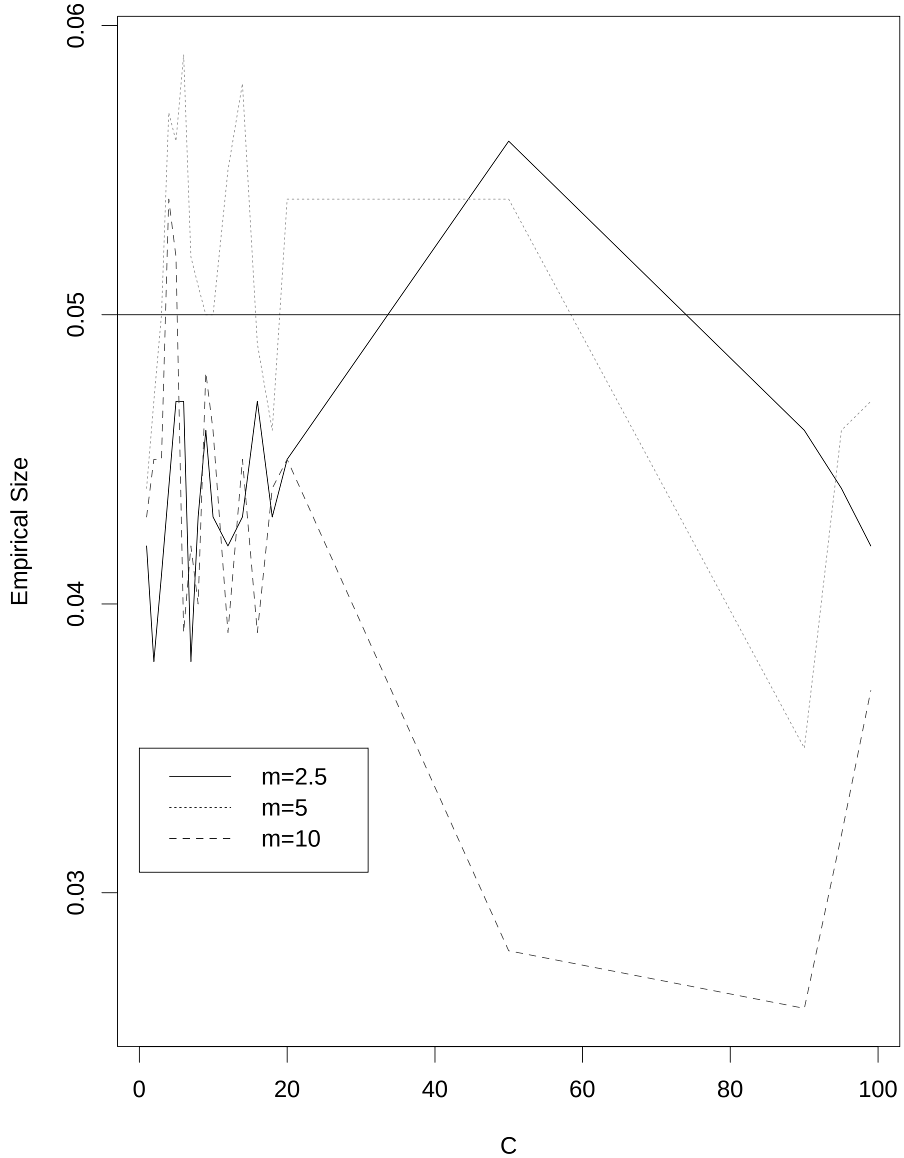

Figure 2.5: Empirical sizes for Model I using lagged first differences (Weighted Symmetric estimator)

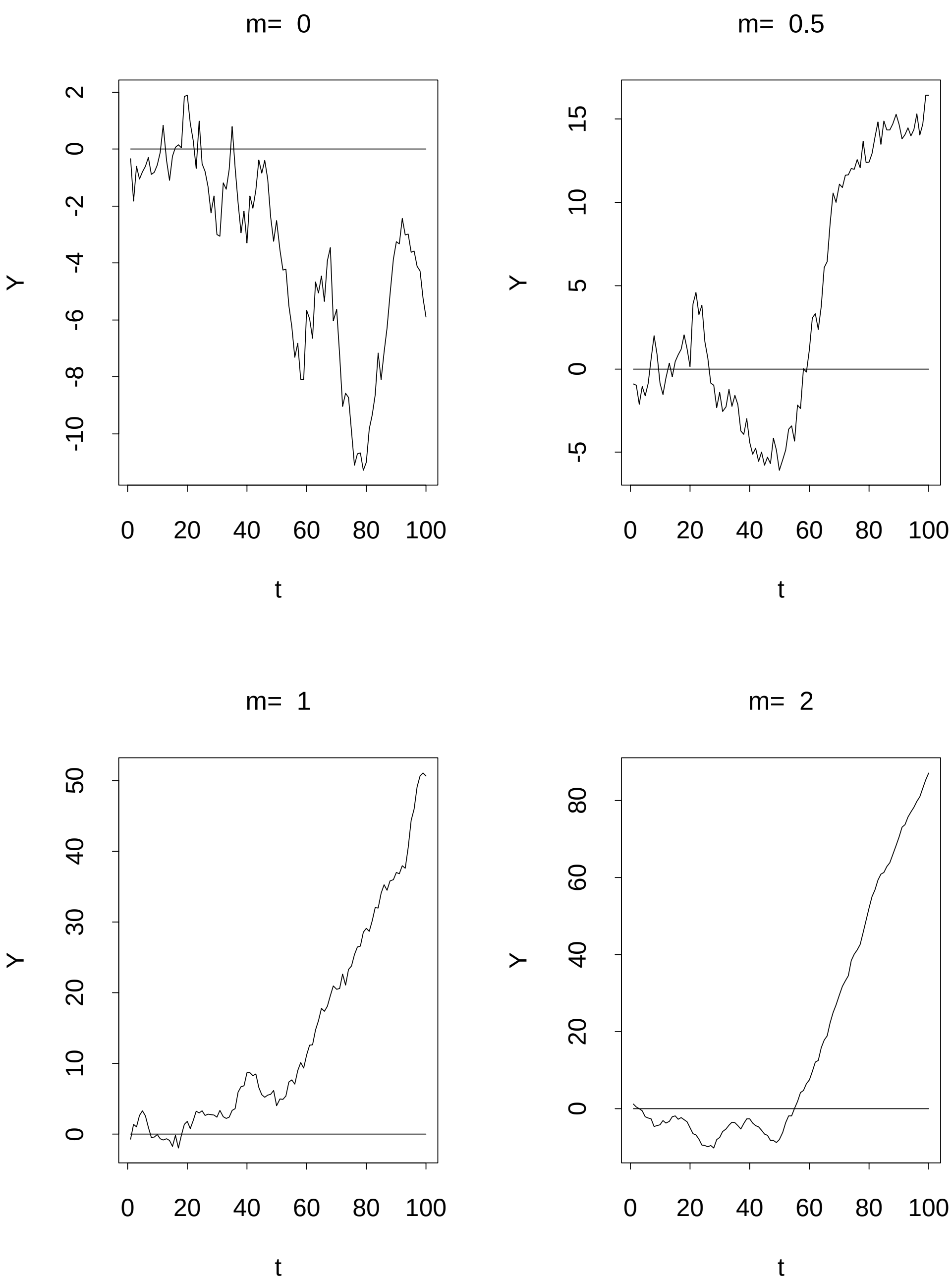

Figure 2.6: Random walk with a break in drift (Model II, $c = 50$)

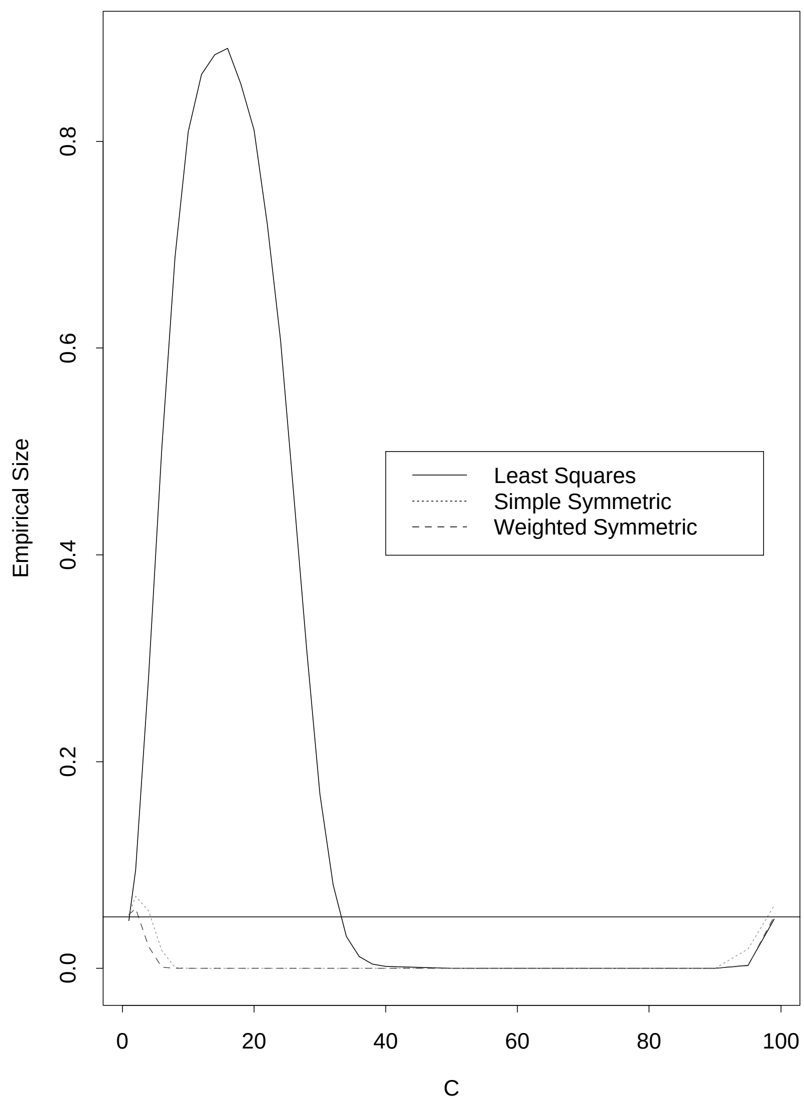

Figure 2.7: Empirical sizes for Model II with $m = 2$ using linear trend adjusted statistics

Chapter 3

Alternative Method for Testing the Unit Root Null Hypothesis in the Presence of a Break

3.1 Introduction

Since the pioneering work by Perron (1989), many researchers have been interested in testing for a unit root in time series with a trend-break. It is known that power decreases in finite samples as the trend-break becomes larger when the usual Dickey-Fuller (1979) test is used.

Perron (1989) suggested formal statistical tests of the null hypothesis of a unit root which can distinguish the unit root hypothesis from that of stationarity around a trend with a single trend-break (either in the intercept or the slope). His original approach assumed the trend-break is known a priori and treated as exogenous. Perron (1997) and Vogelsang and Perron (1998), on the other hand, considered unit root tests treating the date of a possible trend-break as unknown. They introduced various methods of choosing the break date. A variation of Perron's (1989) test in which the break point is estimated rather than fixed was also studied by Zivot and Andrews

(1992). They checked all possible break points and took the minimum t statistic.

In this Chapter, we present an alternative test statistic for testing the unit root hypothesis allowing for a possible trend break. Although our method assumes that the break time is unknown, it is ignored rather than estimated. Using the new test statistic, we perform a Monte Carlo simulation to obtain its empirical powers which are rather invariant to the break size.

Section 3.2 describes the idea of subgrouping and simulation results to find the optimal number of subgroups. As a result, we suggest new test statistics in section 3.3. Section 3.4 presents the data generating processes and some empirical power results from a Monte Carlo simulation. In section 3.5, we apply our test procedure to the real data analyzed originally by Nelson and Plosser (1982). We finally make concluding remarks in section 3.6.

3.2 Subgrouping of data

Our new test procedure is based on the idea of dividing the whole data into some subgroups of the same size. For each subgroup, a certain t type statistic is calculated. Then the minimum of all these statistics is defined as a new test statistic for the unit root null hypothesis in time series with a break.

Suppose we have stationary series around a broken trend. After dividing the data into subgroups, we expect to have a smaller value of the test statistic from a subgroup without a break than from another subgroup with a break. Therefore taking the minimum among all the statistics might give us reasonable power as we are using left tailed tests. This is a key motivation for our subgrouping idea.

The approach by Perron (1989) assumes the break point is known a priori. Using dummy variables, he combines the data before the break with the data after the break. We do not have to assume a known break point in our new procedure.

With the break point assumed unknown, Zivot and Andrews (1992) choose the break point that gives the least favorable result for the unit root null hypothesis. Then they take the t type statistic giving that break point. Our procedure is simpler because it does not consider estimating the unknown break point.

We perform some simulations to decide the optimal number of subgroups. The data generating process given by (3.1) in section 3.4 is used for various numbers of subgroups. We consider $n(\hat{\rho}-1)$ type statistics from the ordinary least squares (OLS) estimator, the simple symmetric (SS) estimator and the weighted symmetric (WS) estimator. The number of replications is 5,000 and the sample size per replication is $n = 100$. We consider the break point $c = 36$ (for OLS) or $c = 37$ (for SS and WS) and the break size $\theta = 1$, 3, 5 and 7. In Tables 3.1-3.3 and Figures 3.1-3.3, we show empirical size and power results from simulation.

Clearly the optimal number of subgroups is 2. We expect similar results for t type statistics. The next section describes our new test procedure which we call the "bisection" method.

3.3 Bisection method

From the previous work by Leybourne et al (1998) and Huh and Dickey (1999), we know that, in the presence of an early break, the conventional Dickey-Fuller (1979) test based on the least squares estimator can be subject to serious size distortion but tests based on the weighted symmetric estimator cure this problem. Therefore our bisection method in this Chapter is based on the WS estimator of ρ, $\tilde{\rho}_w$, and the associated pivotal statistic τ_w in the non-zero mean AR(1) process

$$Y_t = \mu + \rho Y_{t-1} + e_t.$$

Our new bisection test statistic is defined as

$$\tau_w^* \equiv \min(\tau_{w,1}, \tau_{w,2})$$

where $\tau_{w,k}$ is τ_w for subgroup k for $k = 1, 2$. Each subgroup is supposed to have the same number of observations. Note that we are dealing with the mean adjusted case. See Appendix B for the limiting distribution of τ_w^*. Figure 3.4 shows empirical distributions of τ_w^*, $\tau_{w,1}$, $\tau_{w,2}$ and the usual τ_w associated with the data generating process (3.1) in section 3.4. We assume $n = 100$, $c = 75$, $\theta = 5$ and $\rho = .7$. Notice that the empirical distributions of $\tau_{w,1}$ and τ_w^* are very close to each other since the break is in the second subgroup ($c = 75$). That does not necessarily mean that they have the same power for a certain ρ because the critical values are different as shown in Figure 3.4. In the figure, the vertical line on the left denotes the critical value for τ_w^*, so the power for τ_w^* is lower than that for $\tau_{w,1}$ as can be seen in Figure 3.5. This generally holds good for other values of ρ.

As to the linear trend adjusted case, we consider the WS estimator of ρ and the associated pivotal statistic $\tau_{w,\tau}$ in the model

$$Y_t = \mu + \beta t + \rho Y_{t-1} + e_t.$$

$\tau_{w,\tau}$ can be regarded as the linear trend adjusted version of τ_w. We then use $\tau_{w,\tau}^*$, the minimum of two $\tau_{w,\tau}$ statistics, as a new test statistic. See Huh and Dickey (1999) for more details about τ_w and $\tau_{w,\tau}$.

3.4 Empirical size and power results

In this section, we display some size and power simulation results using bisection test statistics.

We consider two data generating processes (DGPs) given by

$$Y_t = \theta\sigma I(t > c) + W_t, \quad W_t = \rho W_{t-1} + e_t, \qquad t = 1, 2, ..., n \tag{3.1}$$

$$Y_t = \gamma\sigma(t - c)I(t > c) + W_t, \quad W_t = \rho W_{t-1} + e_t, \quad t = 1, 2, ..., n \tag{3.2}$$

where the e_t are normal independent $(0,\sigma^2)$ random variables and c is the break point. We can assume $\sigma = 1$ since $\tilde{\rho}_w$ is the ratio of two quadratic forms so it is invariant to σ. Notice that DGP (3.1) corresponds to a break-in-level model with levels 0 and $\theta\sigma$ and DGP (3.2) to a break-in-slope model with slopes 0 and $\gamma\sigma$. Figures 3.6-3.9 present some typical time series data generated from DGPs (3.1) and (3.2).

We perform a Monte Carlo simulation for some values of θ, γ, c and ρ. The number of replications is 5,000 and the sample size per replication is $n = 100$. The critical values for τ_w^* and $\tau_{w,\tau}^*$ are also calculated by simulation.

3.4.1 Data with a break in level

3.4.1.1 Mean adjusted case

For DGP (3.1), we first use the mean adjusted statistics τ_w and τ_w^*. As might be expected from our Chapter 2 results, neither τ_w nor τ_w^* show size distortion for DGP (3.1).

If we use τ_w^* as a test statistic, the powers are in general higher than those for the usual τ_w when there is a break. Some exceptions are for the early break ($c = 1$) or the late break ($c = 99$). For given ρ, the powers for τ_w^* are relatively invariant whereas those for the usual τ_w decrease dramatically as θ becomes larger.

If there is no break ($\theta = 0$), on the other hand, the usual τ_w is more powerful than τ_w^*. This is so especially for large values of ρ. Notice that our power results for τ_w in Table 3.4 agree closely with those of Pantula et al (1994).

Table 3.4 and Figures 3.10-3.12 present the empirical size and power results for DGP (3.1) using τ_w and τ_w^*. As might be expected, τ_w outperforms τ_w^* for small breaks and τ_w^* outperforms τ_w for larger breaks. In all cases, τ_w^* is more uniform with respect to λ.

We compare the size and power properties of our test procedure with those of Vogelsang and Perron (1998). They consider a DGP

$$
\begin{aligned}
Y_t &= \theta I(t > T_b^c) + \gamma(t - T_b^c)I(t > T_b^c) + Z_t, \\
Z_t &= \rho Z_{t-1} + \alpha \Delta Z_{t-1} + e_t + \psi e_{t-1}, \qquad t = 1, \cdots, n
\end{aligned}
\tag{3.3}
$$

where $e_t \sim NI(0,1)$ and T_b^c stands for the true break date. With $\alpha = \psi = \gamma = 0$, DGP (3.3) is exactly the same as our DGP (3.1). They introduce 4 different test statistics ($T_b(t_{\hat{\alpha}})$, $T_b(|t_{\hat{\gamma}}|)$, $T_b(t_{\hat{\gamma}})$ and $T_b(F_{\hat{\theta},\hat{\gamma}})$) according to the methods of estimating the true break date. For their simulation, $n = 100$ and $T_b^c = 50$. When $\rho = .8$, our τ_w^* shows higher power than most of their statistics regardless of the break size θ. Although one statistic ($T_b(t_{\hat{\alpha}})$) with $\theta = 10$ has higher power than our τ_w^*, that can be attributed to its size distortion problem. In Table 3.5, we show empirical powers for DGP (3.1) using various statistics.

3.4.1.2 Linear trend adjusted case

We next use the linear trend adjusted statistics $\tau_{w,\tau}$ and $\tau_{w,\tau}^*$ for DGP (3.1). Both $\tau_{w,\tau}$ and $\tau_{w,\tau}^*$ retain empirical sizes ($\rho = 1$) close to the nominal 5% significance level.

When there is no break, $\tau_{w,\tau}$ shows higher power than $\tau_{w,\tau}^*$ for given ρ. Notice also that $\tau_{w,\tau}$ and $\tau_{w,\tau}^*$ generate lower power than τ_w and τ_w^* respectively.

When the break is fairly big ($\theta = 10$), the powers for $\tau_{w,\tau}^*$ are generally higher than those for $\tau_{w,\tau}$. As in the mean adjusted case, the powers for $\tau_{w,\tau}^*$ are relatively invariant compared with those for $\tau_{w,\tau}$.

The empirical size and power results using $\tau_{w,\tau}$ and $\tau_{w,\tau}^*$ are presented in Table 3.6 and Figures 3.13-3.15.

3.4.2 Data with a break in slope

For DGP (3.2) which represents a break-in-slope model, we adopt the linear trend adjusted statistics $\tau_{w,\tau}$ and $\tau_{w,\tau}^*$. Even though both $\tau_{w,\tau}$ and $\tau_{w,\tau}^*$ retain the nominal 5% significance level, $\tau_{w,\tau}$ becomes severely under-sized as γ grows larger.

As expected, $\tau_{w,\tau}^*$ gives smaller powers than $\tau_{w,\tau}$ if there is no break. As γ becomes larger, the powers for $\tau_{w,\tau}$ decrease dramatically except for the early or late breaks whereas $\tau_{w,\tau}^*$ maintains reasonable powers. $\tau_{w,\tau}^*$ is generally more powerful than $\tau_{w,\tau}$ for $\gamma > 0$.

The simulation results for DGP (3.2) using $\tau_{w,\tau}$ and $\tau^*_{w,\tau}$ are displayed in Table 3.7 and Figures 3.16-3.18.

To compare with the results of Vogelsang and Perron (1998), we consider DGP (3.3) with $\alpha = \psi = \theta = 0$ which is exactly the same as our DGP (3.2). When $\rho = .8$, the powers of our $\tau^*_{w,\tau}$ are as good as those of their statistics for some values of γ. Power comparisons between $\tau^*_{w,\tau}$ and statistics of Vogelsang and Perron (1998) are displayed in Table 3.8.

3.5 Empirical applications

We now apply our bisection method to the data set analyzed originally by Nelson and Plosser (1982). The data set consists of 14 major macroeconomic time series which include measures of output, spending, money, prices and interest rates. The data are annual, generally averages for the year, with starting dates from 1860 to 1909 and ending in 1970 in all cases. We analyze the natural logarithm of all the data except for the interest rate series, which is analyzed in levels form. Many researchers have referred to this data set. See Nelson and Plosser (1982) for more details about the data set.

In Table 3.9, we compare our unit root test results with those of Nelson and Plosser (1982), Perron (1989), Zivot and Andrews (1992) and Perron (1997). Numbers are the values of the test statistics.

Nelson and Plosser (1982) apply the usual Dickey-Fuller test with extra lags of the first differences of the data (augmented Dickey-Fuller test). Methods of Perron (1989) and Zivot and Andrews (1992) are briefly described in section 3.2. Perron (1997) is closely related to and complements Zivot and Andrews (1992) in that similar procedures are analyzed. See the original articles for more details.

Our method adopts the number of augmenting terms, k, producing the minimum test statistic for each series. The other methods also use various values of k according to some selection criteria. Nelson and Plosser (1982) use critical values from Fuller (1996) whereas the other methods, including ours, use their own critical values generated from simulation. Our bisection method using $\tau^*_{w,\tau}$ rejects the unit root null hypothesis at $\alpha = .05$ for the series "Real GNP", "Real per capita GNP", "Industrial production", "Unemployment rate", "Real wages" and "Money stock".

3.6 Summary

In this work, we suggest new test statistics τ^*_w and $\tau^*_{w,\tau}$ for testing the unit root null hypothesis in the presence of a trend break. These statistics are based on the idea of dividing the data into subgroups of the same size. The optimal number of subgroups turns out to be 2.

Our bisection method can be used without assuming a known break point unlike Perron's (1989) original approach. It is simpler than the methods of Zivot and Andrews (1992) and Perron (1997). When there is a trend-break and the break size is fairly big, the empirical powers of the new statistics are in general higher than the usual τ_w and $\tau_{w,\tau}$ respectively. Based on simulation studies, the power properties of our test procedure are as good as those of Vogelsang and Perron's (1998).

We also apply our procedure to the well-known Nelson and Plosser (1982) data set and compare the results with those of other researchers.

Table 3.1: Empirical size and power for DGP (3.1) using various subgroups (OLS, $n = 100$, $c = 36$)

No. of subgroups	1	2	5	10	20
$\theta = 1$, $\rho = .1$	1.0000	1.0000	0.9994	0.6938	0.0994
.2	1.0000	1.0000	0.9936	0.5424	0.0884
.3	1.0000	1.0000	0.9600	0.4026	0.0826
.4	1.0000	1.0000	0.8554	0.2948	0.0752
.5	1.0000	0.9984	0.6788	0.2096	0.0682
.6	1.0000	0.9780	0.4740	0.1434	0.0634
.7	0.9974	0.8388	0.2898	0.0974	0.0598
.8	0.9132	0.5042	0.1626	0.0744	0.0544
.9	0.4288	0.1870	0.0962	0.0528	0.0524
1	0.0474	0.0470	0.0530	0.0504	0.0488
$\theta = 3$, $\rho = .1$	1.0000	1.0000	0.9986	0.6764	0.0946
.2	1.0000	1.0000	0.9888	0.5246	0.0848
.3	0.9996	1.0000	0.9416	0.3856	0.0798
.4	0.9976	0.9992	0.8126	0.2806	0.0728
.5	0.9806	0.9852	0.6260	0.1980	0.0658
.6	0.9144	0.9052	0.4298	0.1342	0.0606
.7	0.7458	0.6786	0.2564	0.0930	0.0576
.8	0.4824	0.3650	0.1472	0.0710	0.0530
.9	0.2382	0.1464	0.0886	0.0476	0.0504
1	0.0460	0.0478	0.0514	0.0498	0.0476

Table 3.1: *continued*

No. of subgroups	1	2	5	10	20
$\theta = 5,\ \rho = .1$	0.7650	1.0000	0.9986	0.6760	0.0946
.2	0.5990	1.0000	0.9886	0.5240	0.0844
.3	0.4212	1.0000	0.9408	0.3846	0.0792
.4	0.2654	0.9990	0.8104	0.2792	0.0726
.5	0.1612	0.9818	0.6224	0.1966	0.0656
.6	0.0976	0.8920	0.4230	0.1330	0.0604
.7	0.0712	0.6440	0.2482	0.0908	0.0576
.8	0.0626	0.3116	0.1394	0.0694	0.0530
.9	0.0686	0.1114	0.0808	0.0466	0.0504
1	0.0386	0.0418	0.0464	0.0470	0.0476
$\theta = 7,\ \rho = .1$	0.0064	1.0000	0.9986	0.6760	0.0946
.2	0.0018	1.0000	0.9886	0.5240	0.0900
.3	0.0004	1.0000	0.9408	0.3846	0.0768
.4	0	0.9990	0.8104	0.2792	0.0724
.5	0	0.9818	0.6224	0.1966	0.0708
.6	0	0.8920	0.4230	0.1330	0.0678
.7	0.0004	0.6424	0.2480	0.0908	0.0632
.8	0.0008	0.3060	0.1382	0.0692	0.0504
.9	0.0128	0.0978	0.0790	0.0464	0.0546
1	0.0334	0.0358	0.0442	0.0468	0.0520

Table 3.2: Empirical size and power for DGP (3.1) using various subgroups (SS, $n = 100$, $c = 37$)

No. of subgroups	1	2	5	10	20
$\theta = 1$, $\rho = .1$	1.0000	1.0000	0.9968	0.6278	0.1916
.2	1.0000	1.0000	0.9806	0.4670	0.1602
.3	1.0000	1.0000	0.9178	0.3584	0.1266
.4	1.0000	0.9998	0.7978	0.2696	0.1144
.5	1.0000	0.9992	0.5980	0.1874	0.1058
.6	1.0000	0.9754	0.3856	0.1336	0.0802
.7	0.9982	0.8314	0.2360	0.1024	0.0766
.8	0.9384	0.4970	0.1326	0.0730	0.0698
.9	0.4960	0.1856	0.0800	0.0644	0.0542
1	0.0536	0.0470	0.0512	0.0482	0.0494
$\theta = 3$, $\rho = .1$	1.0000	1.0000	0.9934	0.6062	0.1872
.2	1.0000	1.0000	0.9684	0.4490	0.1562
.3	0.9998	1.0000	0.8896	0.3418	0.1242
.4	0.9980	0.9990	0.7462	0.2516	0.1106
.5	0.9866	0.9886	0.5396	0.1736	0.1010
.6	0.9324	0.9036	0.3392	0.1258	0.0778
.7	0.7828	0.6590	0.2152	0.0946	0.0732
.8	0.5340	0.3600	0.1242	0.0682	0.0682
.9	0.2770	0.1548	0.0734	0.0600	0.0522
1	0.0524	0.0436	0.0472	0.0470	0.0478

Table 3.2: *continued*

No. of subgroups	1	2	5	10	20
$\theta = 5,\ \rho = .1$	0.8080	1.0000	0.9932	0.6060	0.1854
.2	0.6590	1.0000	0.9682	0.4482	0.1544
.3	0.4718	1.0000	0.8888	0.3416	0.1224
.4	0.3176	0.9986	0.7420	0.2510	0.1092
.5	0.1868	0.9848	0.5340	0.1724	0.1002
.6	0.1244	0.8860	0.3294	0.1246	0.0770
.7	0.0832	0.6200	0.2018	0.0930	0.0722
.8	0.0680	0.3078	0.1142	0.0666	0.0672
.9	0.0804	0.1162	0.0672	0.0586	0.0506
1	0.0436	0.0406	0.0448	0.0444	0.0474
$\theta = 7,\ \rho = .1$	0.0110	1.0000	0.9932	0.6060	0.1852
.2	0.0024	1.0000	0.9682	0.4482	0.1544
.3	0.0014	1.0000	0.8888	0.3416	0.1224
.4	0	0.9986	0.7420	0.2510	0.1092
.5	0	0.9848	0.5338	0.1724	0.1002
.6	0.0002	0.8860	0.3292	0.1246	0.0768
.7	0	0.6184	0.2004	0.0930	0.0722
.8	0.0014	0.3012	0.1132	0.0666	0.0672
.9	0.0154	0.1008	0.0654	0.0584	0.0504
1	0.0394	0.0334	0.0428	0.0440	0.0474

Table 3.3: Empirical size and power for DGP (3.1) using various subgroups (WS, $n = 100$, $c = 37$)

No. of subgroups	1	2	5	10	20
$\theta = 1,\ \rho = .1$	1.0000	1.0000	0.9988	0.5346	0.1500
.2	1.0000	1.0000	0.9842	0.4218	0.1246
.3	1.0000	1.0000	0.9312	0.3052	0.1012
.4	1.0000	0.9998	0.8060	0.2262	0.0958
.5	1.0000	0.9992	0.6092	0.1666	0.0904
.6	1.0000	0.9844	0.4166	0.1218	0.0698
.7	0.9992	0.8750	0.2438	0.0932	0.0634
.8	0.9592	0.5470	0.1486	0.0708	0.0600
.9	0.5518	0.1942	0.0738	0.0560	0.0498
1	0.0526	0.0488	0.0514	0.0484	0.0484
$\theta = 3,\ \rho = .1$	1.0000	1.0000	0.9966	0.5144	0.1456
.2	1.0000	1.0000	0.9748	0.4044	0.1204
.3	0.9996	1.0000	0.9050	0.2880	0.0996
.4	0.9994	0.9992	0.7600	0.2122	0.0928
.5	0.9924	0.9914	0.5554	0.1562	0.0866
.6	0.9576	0.9318	0.3722	0.1148	0.0678
.7	0.8338	0.7214	0.2154	0.0882	0.0622
.8	0.5912	0.4088	0.1344	0.0658	0.0590
.9	0.3070	0.1544	0.0688	0.0522	0.0468
1	0.0538	0.0440	0.0476	0.0472	0.0474

Table 3.3: *continued*

No. of subgroups	1	2	5	10	20
$\theta = 5,\ \rho = .1$	0.8564	1.0000	0.9966	0.5142	0.1432
.2	0.7314	1.0000	0.9746	0.4042	0.1182
.3	0.5670	0.9998	0.9014	0.2872	0.0964
.4	0.4028	0.9978	0.7582	0.2118	0.0904
.5	0.2524	0.9874	0.5490	0.1556	0.0844
.6	0.1656	0.9174	0.3654	0.1144	0.0664
.7	0.1088	0.6748	0.2084	0.0872	0.0610
.8	0.0846	0.3446	0.1248	0.0640	0.0574
.9	0.0876	0.1172	0.0630	0.0512	0.0462
1	0.0452	0.0410	0.0452	0.0446	0.0466
$\theta = 7,\ \rho = .1$	0.0232	1.0000	0.9966	0.5142	0.1428
.2	0.0064	1.0000	0.9746	0.4042	0.1178
.3	0.0020	0.9998	0.9014	0.2872	0.0962
.4	0.0004	0.9978	0.7582	0.2118	0.0906
.5	0	0.9874	0.5490	0.1556	0.0842
.6	0.0002	0.9172	0.3648	0.1144	0.0662
.7	0.0002	0.6736	0.2076	0.0872	0.0604
.8	0.0016	0.3366	0.1240	0.0640	0.0570
.9	0.0142	0.1024	0.0612	0.0512	0.0460
1	0.0412	0.0360	0.0428	0.0444	0.0464

Table 3.4: Empirical size and power for DGP (3.1) using τ_w and τ_w^*

		$\theta = 0$		$\theta = 5$		$\theta = 10$	
ρ	c	τ_w	τ_w^*	τ_w	τ_w^*	τ_w	τ_w^*
0.5	1	1.0000	1.0000	0.9998	0.9968	0.9810	0.9900
0.5	20	1.0000	1.0000	0.6532	0.9878	0.0000	0.9884
0.5	50	1.0000	1.0000	0.1898	1.0000	0.0000	0.9996
0.5	80	1.0000	0.9996	0.6576	0.9884	0.0000	0.9902
0.5	99	1.0000	0.9996	1.0000	0.9970	0.9774	0.9940
0.8	1	0.9834	0.6006	0.8782	0.4746	0.5392	0.3858
0.8	20	0.9824	0.5890	0.2132	0.3476	0.0000	0.3416
0.8	50	0.9854	<u>0.5994</u>	0.0780	<u>0.5918</u>	0.0000	<u>0.5960</u>
0.8	80	0.9840	0.6020	0.2594	0.4016	0.0000	0.4056
0.8	99	0.9816	0.5918	0.8532	0.5094	0.5614	0.4784
0.9	1	0.6090	0.2256	0.4218	0.1690	0.1904	0.1362
0.9	20	0.6236	0.2244	0.1214	0.1210	0.0006	0.0952
0.9	50	0.6220	0.2316	0.0912	0.2332	0.0002	0.2234
0.9	80	0.6294	0.2188	0.1746	0.1496	0.0056	0.1380
0.9	99	0.6256	0.2258	0.4632	0.2020	0.2708	0.1790
1	1	0.0588	0.0516	0.0504	0.0494	0.0480	0.0502
1	20	0.0538	0.0524	0.0470	0.0398	0.0358	0.0298
1	50	0.0516	0.0496	0.0518	0.0484	0.0328	0.0498
1	80	0.0508	0.0466	0.0496	0.0434	0.0404	0.0310
1	99	0.0584	0.0488	0.0520	0.0518	0.0472	0.0488

Table 3.5: Empirical size and power for DGP (3.1) using various statistics ($n = 100$, $c = 50$)

ρ			1			0.8
θ	0	5	10	0	5	10
$T_b(t_{\hat{\alpha}})^\dagger$	0.040	0.108	0.507	0.295	0.435	0.861
$T_b(\lvert t_{\hat{\gamma}}\rvert)^\dagger$	0.049	0.048	0.032	0.301	0.098	0.042
$T_b(t_{\hat{\gamma}})^\dagger$	0.050	0.050	0.040	0.350	0.133	0.055
$T_b(F_{\hat{\theta},\hat{\gamma}})^\dagger$	0.052	0.050	0.020	0.339	0.194	0.163
τ_w^*	<u>0.050</u>	<u>0.048</u>	<u>0.050</u>	<u>0.599</u>	<u>0.592</u>	<u>0.596</u>

† Values are from Vogelsang and Perron (1998)

Table 3.6: Empirical size and power for DGP (3.1) using $\tau_{w,\tau}$ and $\tau^*_{w,\tau}$

		$\theta = 0$		$\theta = 5$		$\theta = 10$	
ρ	c	$\tau_{w,\tau}$	$\tau^*_{w,\tau}$	$\tau_{w,\tau}$	$\tau^*_{w,\tau}$	$\tau_{w,\tau}$	$\tau^*_{w,\tau}$
0.5	1	1.0000	0.9880	0.9986	0.9498	0.9270	0.9088
0.5	20	1.0000	0.9848	0.8418	0.8906	0.0018	0.8802
0.5	50	1.0000	0.9850	0.9720	0.9850	0.0992	0.9854
0.5	80	1.0000	0.9844	0.8406	0.9044	0.0012	0.8854
0.5	99	1.0000	0.9892	0.9984	0.9516	0.9258	0.9170
0.8	1	0.8292	0.2970	0.6226	0.2494	0.3082	0.1912
0.8	20	0.8386	0.2954	0.2842	0.2016	0.0050	0.1486
0.8	50	0.8346	0.3040	0.4354	0.3068	0.0484	0.3016
0.8	80	0.8312	0.2986	0.2946	0.2152	0.0060	0.1848
0.8	99	0.8238	0.2938	0.6160	0.2542	0.3454	0.2372
0.9	1	0.3066	0.1084	0.2150	0.0936	0.0982	0.0816
0.9	20	0.2968	0.1126	0.1298	0.0932	0.0162	0.0594
0.9	50	0.2974	0.1152	0.1874	0.1130	0.0528	0.1090
0.9	80	0.2906	0.1108	0.1426	0.0936	0.0232	0.0654
0.9	99	0.2982	0.1082	0.2306	0.1048	0.1318	0.0902
1	1	0.0574	0.0558	0.0544	0.0524	0.0426	0.0458
1	20	0.0620	0.0496	0.0496	0.0470	0.0342	0.0316
1	50	0.0540	0.0498	0.0528	0.0466	0.0342	0.0480
1	80	0.0498	0.0502	0.0536	0.0448	0.0362	0.0314
1	99	0.0494	0.0476	0.0456	0.0512	0.0430	0.0494

Table 3.7: Empirical size and power for DGP (3.2) using $\tau_{w,\tau}$ and $\tau^*_{w,\tau}$

		$\gamma=0$		$\gamma=1$		$\gamma=2$	
ρ	c	$\tau_{w,\tau}$	$\tau^*_{w,\tau}$	$\tau_{w,\tau}$	$\tau^*_{w,\tau}$	$\tau_{w,\tau}$	$\tau^*_{w,\tau}$
0.5	1	1.0000	0.9858	1.0000	0.9872	1.0000	0.9860
0.5	20	1.0000	0.9862	0.0000	0.8884	0.0000	0.8804
0.5	50	1.0000	0.9860	0.0000	0.9852	0.0000	0.9892
0.5	80	1.0000	0.9862	0.0000	0.8848	0.0000	0.8852
0.5	99	1.0000	0.9868	1.0000	0.9860	1.0000	0.9822
0.8	1	0.8296	0.3098	0.8172	0.2974	0.8322	0.3084
0.8	20	0.8170	0.2934	0.0000	0.1544	0.0000	0.1572
0.8	50	0.8256	<u>0.3060</u>	0.0000	<u>0.2980</u>	0.0000	<u>0.3020</u>
0.8	80	0.8330	0.3126	0.0000	0.1704	0.0000	0.1672
0.8	99	0.8170	0.3016	0.8138	0.2940	0.7774	0.3064
0.9	1	0.2962	0.1128	0.2994	0.1006	0.3002	0.1088
0.9	20	0.3050	0.1138	0.0000	0.0524	0.0000	0.0538
0.9	50	0.2968	0.1106	0.0000	0.1118	0.0000	0.1206
0.9	80	0.2974	0.1078	0.0000	0.0624	0.0000	0.0582
0.9	99	0.3026	0.1130	0.2894	0.1114	0.0552	0.1180
1	1	0.0572	0.0510	0.0556	0.0568	0.0566	0.0470
1	20	0.0526	0.0498	0.0000	0.0302	0.0000	0.0282
1	50	0.0552	0.0504	0.0000	0.0496	0.0000	0.0516
1	80	0.0622	0.0514	0.0000	0.0260	0.0000	0.0270
1	99	0.0574	0.0482	0.0522	0.0520	0.0552	0.0524

Table 3.8: Empirical size and power for DGP (3.2) using various statistics ($n = 100$, $c = 50$)

ρ			1			0.8
γ	0	1	2	0	1	2
$T_b(t_{\hat{\alpha}})^\dagger$	0.040	0.044	0.076	0.295	0.236	0.386
$T_b(\|t_{\hat{\gamma}}\|)^\dagger$	0.049	0.036	0.032	0.301	0.239	0.230
$T_b(t_{\hat{\gamma}})^\dagger$	0.050	0.072	0.061	0.350	0.376	0.371
$T_b(F_{\hat{\theta},\hat{\gamma}})^\dagger$	0.052	0.027	0.021	0.339	0.193	0.184
$\tau^*_{w,\tau}$	<u>0.050</u>	<u>0.050</u>	<u>0.052</u>	<u>0.306</u>	<u>0.298</u>	<u>0.302</u>

†Values are from Vogelsang and Perron (1998)

Table 3.9: Test results for the unit root null hypothesis

Series	n	N&P[‡]	P1[§]	Z&A[¶]	P2[‖]	Bisection[**]
Real GNP	62	-2.99	-5.03[*]	-5.58[*]	-5.93[*]	-4.09[†]
Nominal GNP	62	-2.32	-5.42[*]	-5.82[*]	-8.16[*]	-2.28
Real per capita GNP	62	-3.04	-4.09[†]	-4.61	-4.81	-3.71[†]
Industrial production	111	-2.53	-5.47[*]	-5.95[*]	-6.01[*]	-5.72[†]
Employment	81	-2.66	-4.51[*]	-4.95[†]	-5.14[†]	-3.46
Unemployment rate	81	-3.55[†]	N/A	N/A	N/A	-4.06[†]
GNP deflator	82	-2.52	-4.04[†]	-4.12	-4.14	-2.67
Consumer prices	111	-1.97	-1.28	-2.76	-3.09	-3.18
Wages	71	-2.09	-5.41[*]	-5.30[†]	-5.41[†]	-2.93
Real wages	71	-3.04	-4.28[†]	-4.74	-5.41	-4.35[†]
Money stock	82	-3.08	-4.29[†]	-4.34	-4.69	-4.48[†]
Velocity	102	-1.66	-1.66	-3.39	-2.81	-2.44
Interest rate	71	0.686	-0.45	-0.98	-1.35	-0.80
Common stock prices	100	-2.05	-4.87[†]	-5.61[*]	-5.50[†]	-3.34

[*] statistical significance at the 1% level
[†] statistical significance at the 5% level
[‡] Nelson and Plosser (1982)
[§] Perron (1989)
[¶] Zivot and Andrews (1992)
[‖] Perron (1997)
[**] critical value from simulation : -3.62 for $n = 100$

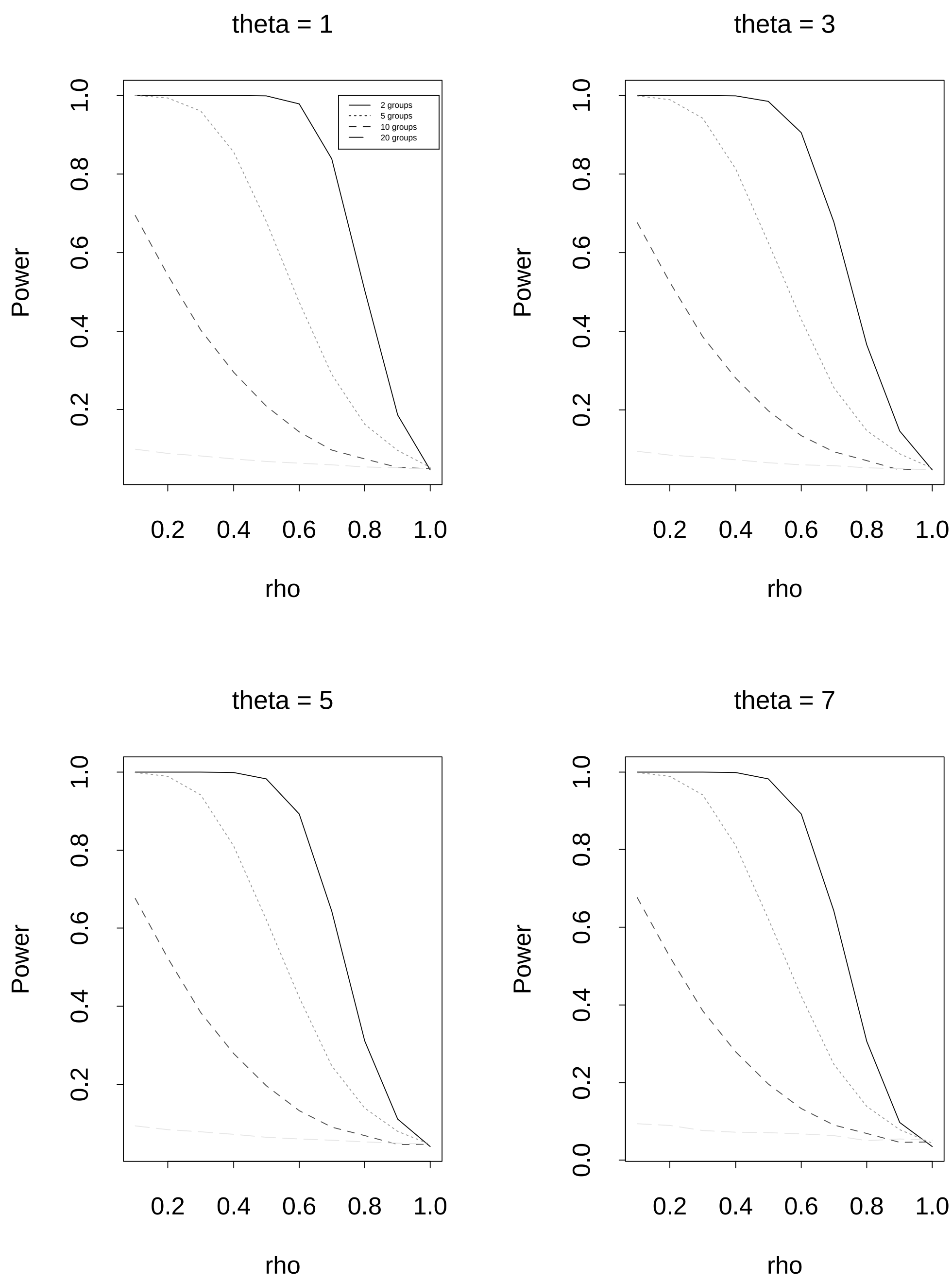

Figure 3.1: Empirical size and power for DGP (3.1) using various subgroups (OLS, $n = 100$, $c = 36$)

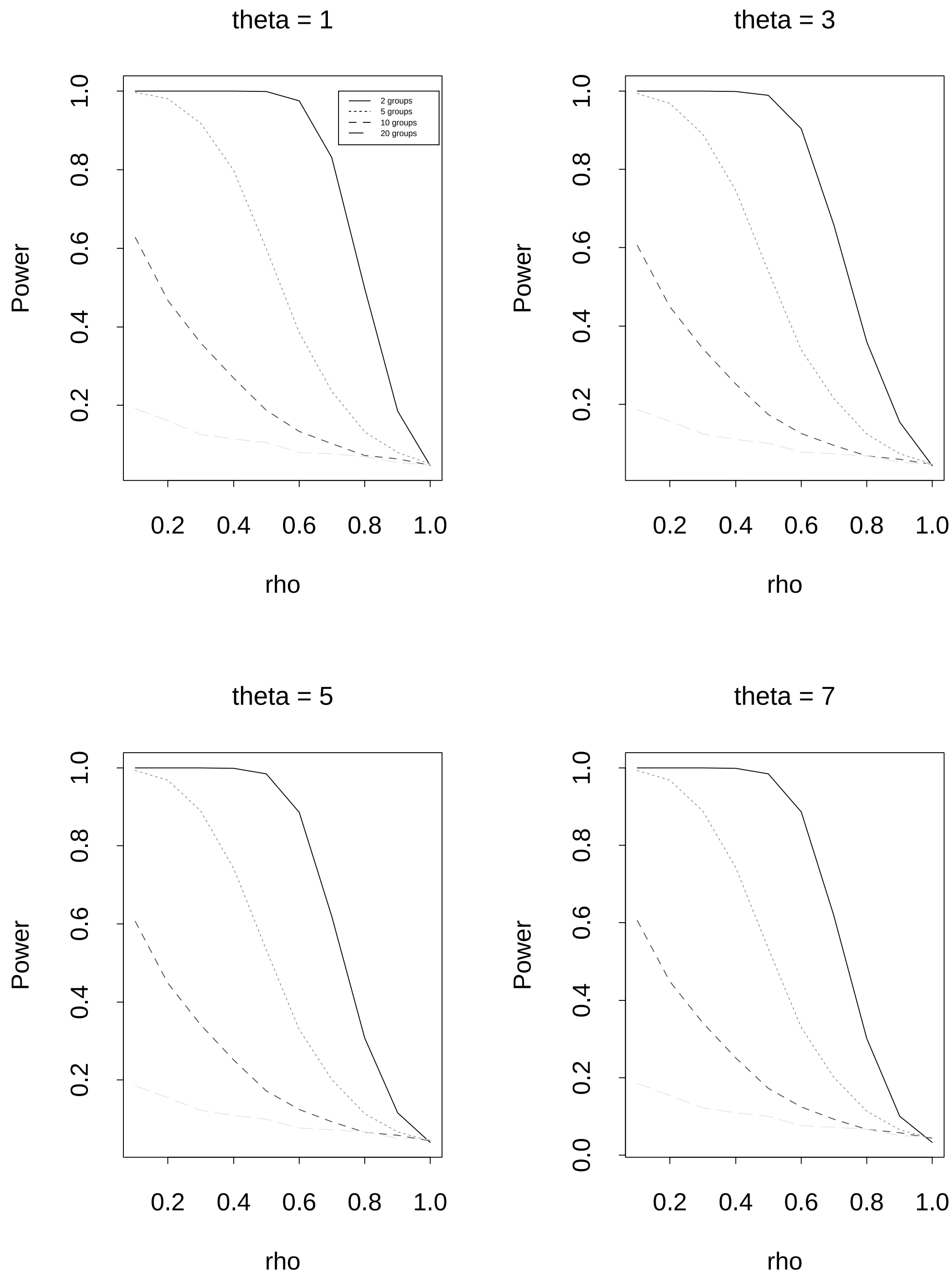

Figure 3.2: Empirical size and power for DGP (3.1) using various subgroups (SS, $n = 100$, $c = 37$)

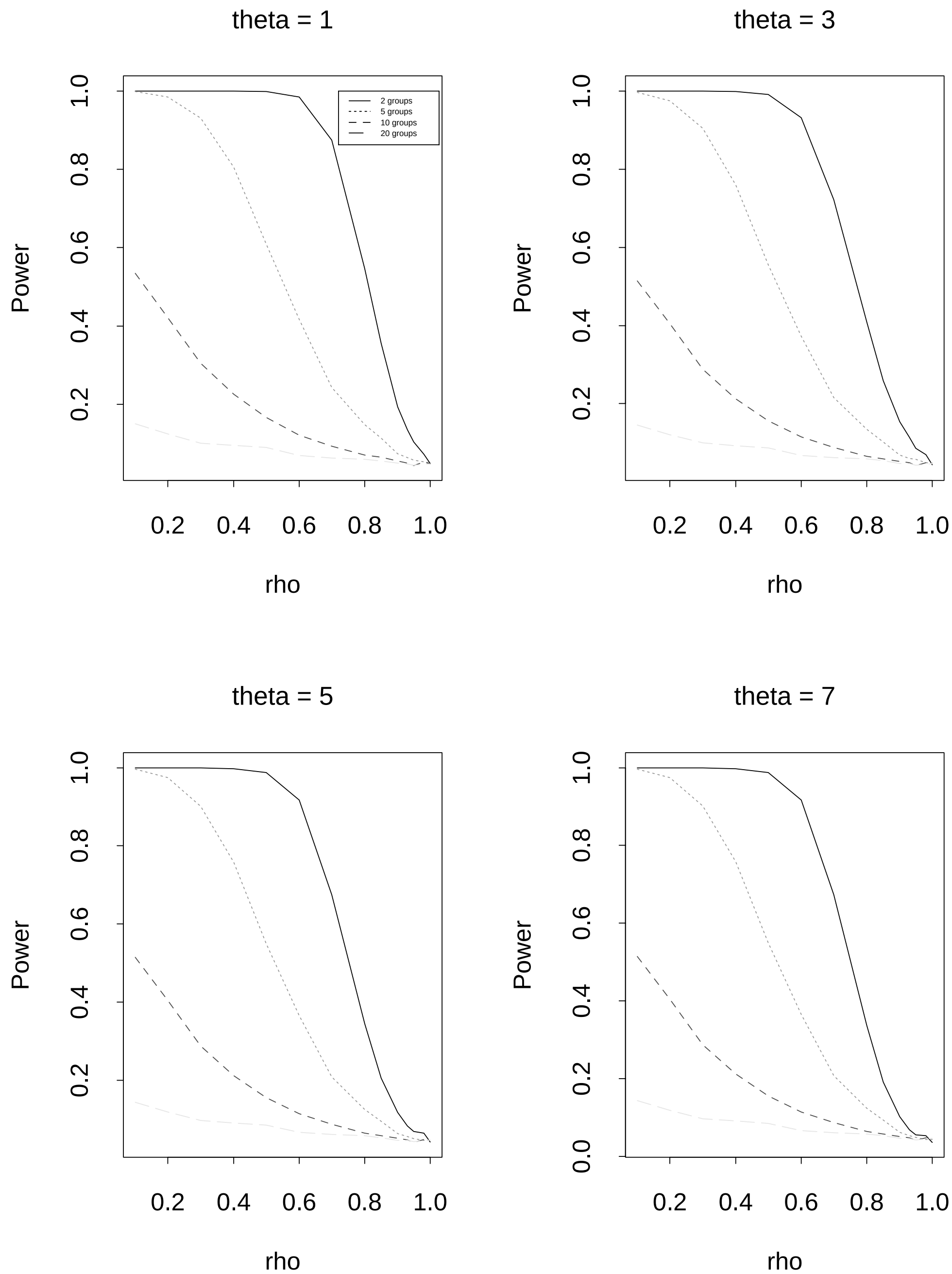

Figure 3.3: Empirical size and power for DGP (3.1) using various subgroups (WS, $n = 100$, $c = 37$)

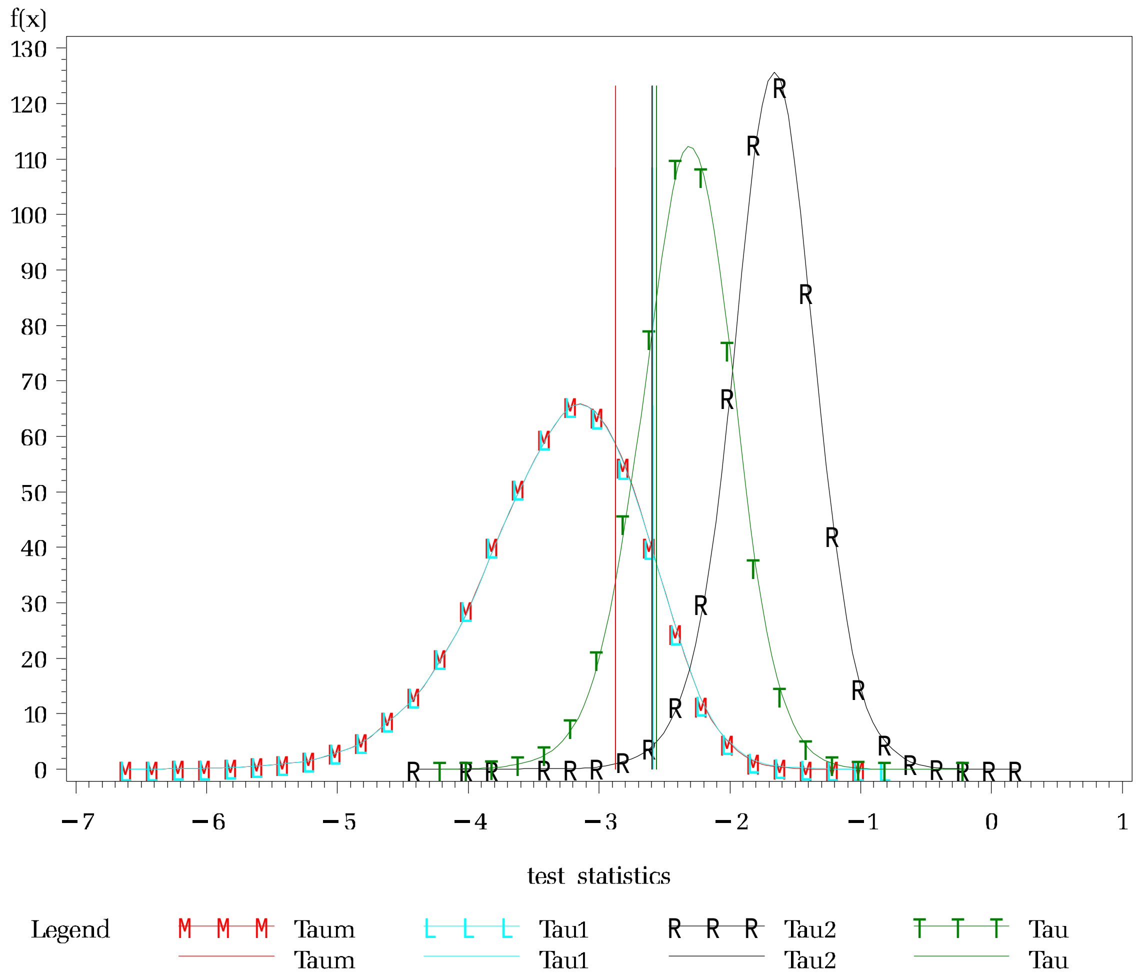

Figure 3.4: Empirical distributions of test statistics (WS, $n = 100$, $c = 75$, $\theta = 5$ and $\rho = .7$; T for τ_w, M for τ_w^*, L for $\tau_{w,1}$ and R for $\tau_{w,2}$; M overlays L.)

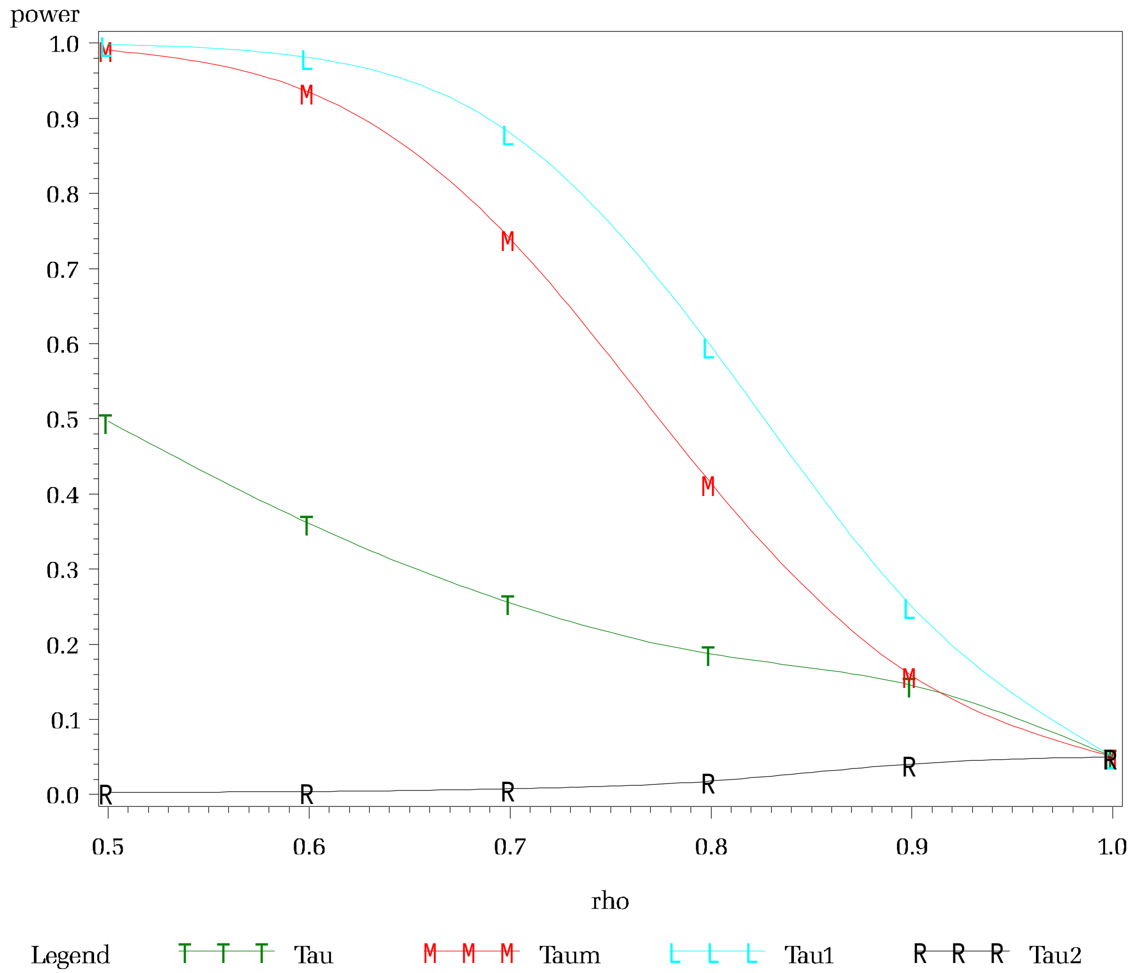

Figure 3.5: Empirical size and power of test statistics (WS, $n = 100$, $c = 75$ and $\theta = 5$; T for τ_w, M for τ_w^*, L for $\tau_{w,1}$ and R for $\tau_{w,2}$)

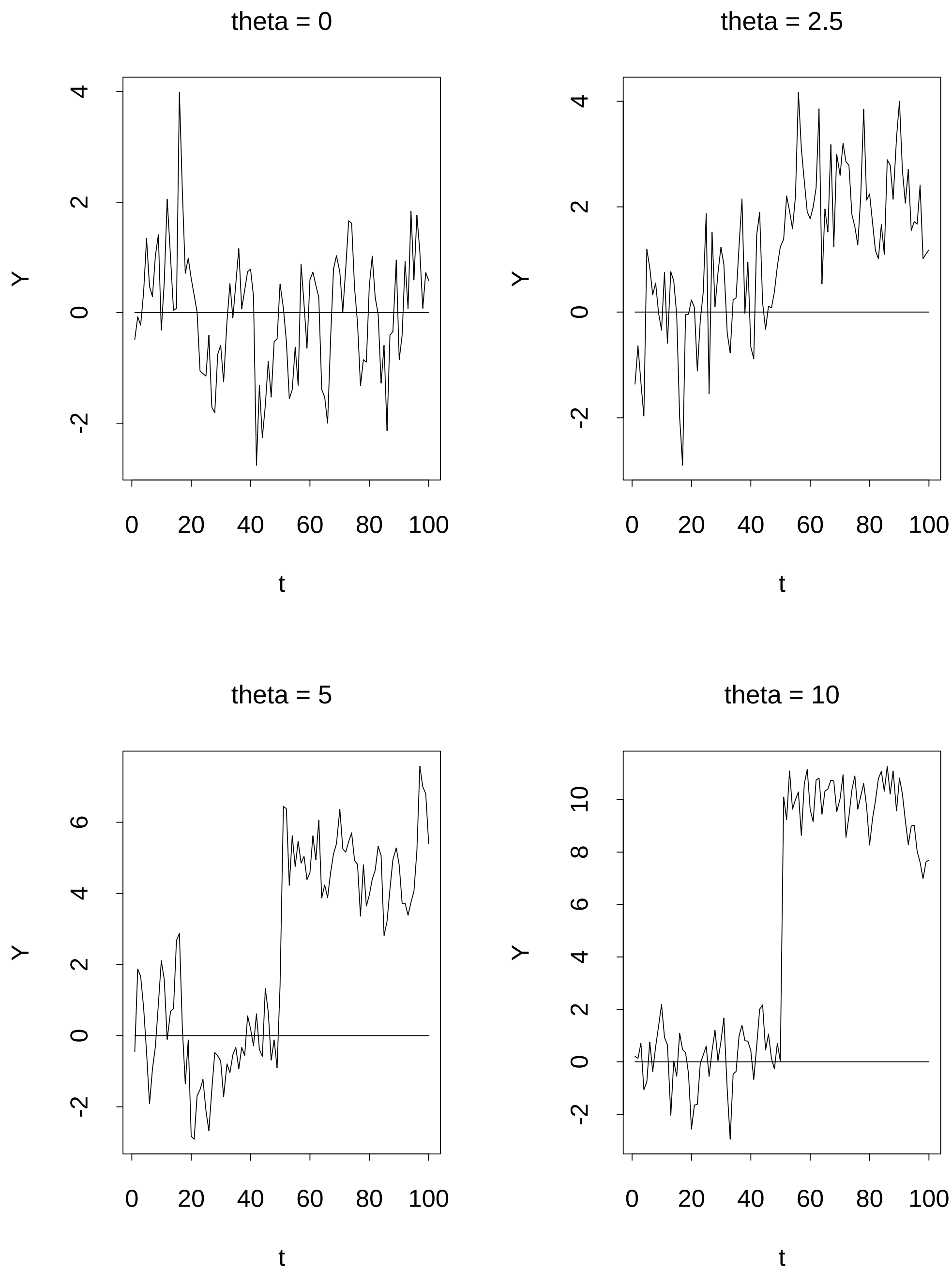

Figure 3.6: Data from a break-in-level model (DGP (3.1), $\rho = 0.5$, $c = 50$)

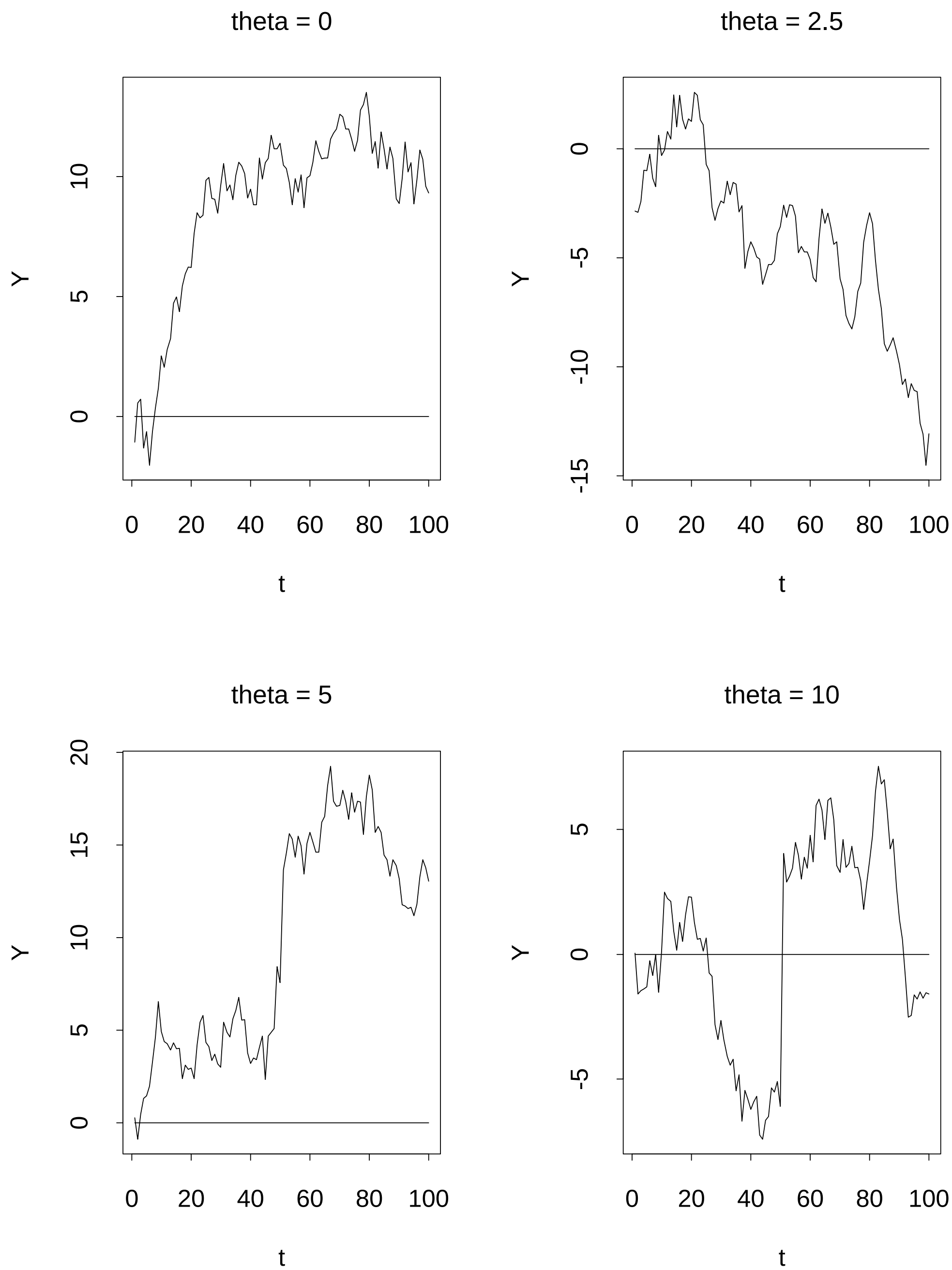

Figure 3.7: Data from a break-in-level model (DGP (3.1), $\rho = 1$, $c = 50$)

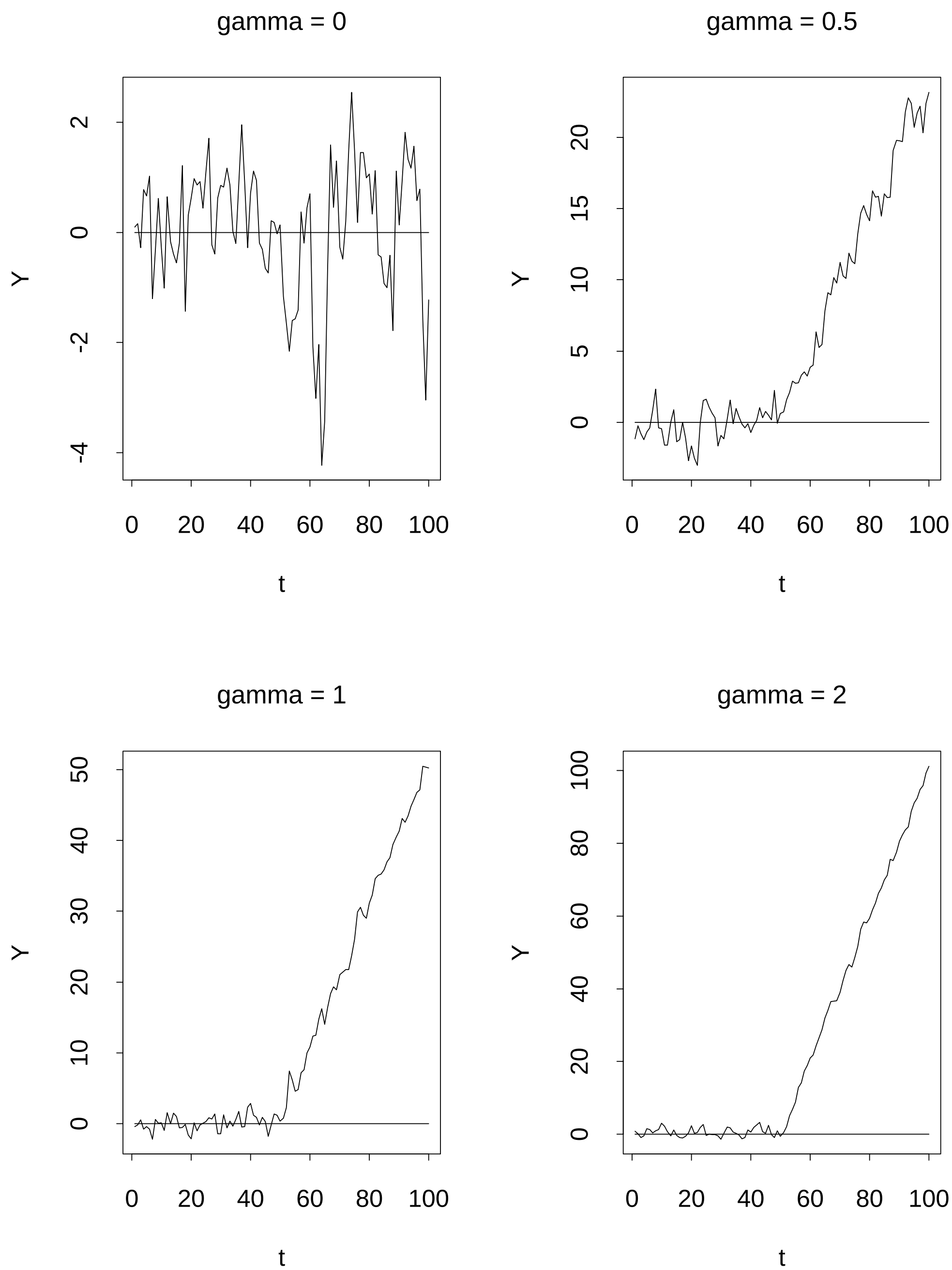

Figure 3.8: Data from a break-in-slope model (DGP (3.2), $\rho = 0.5$, $c = 50$)

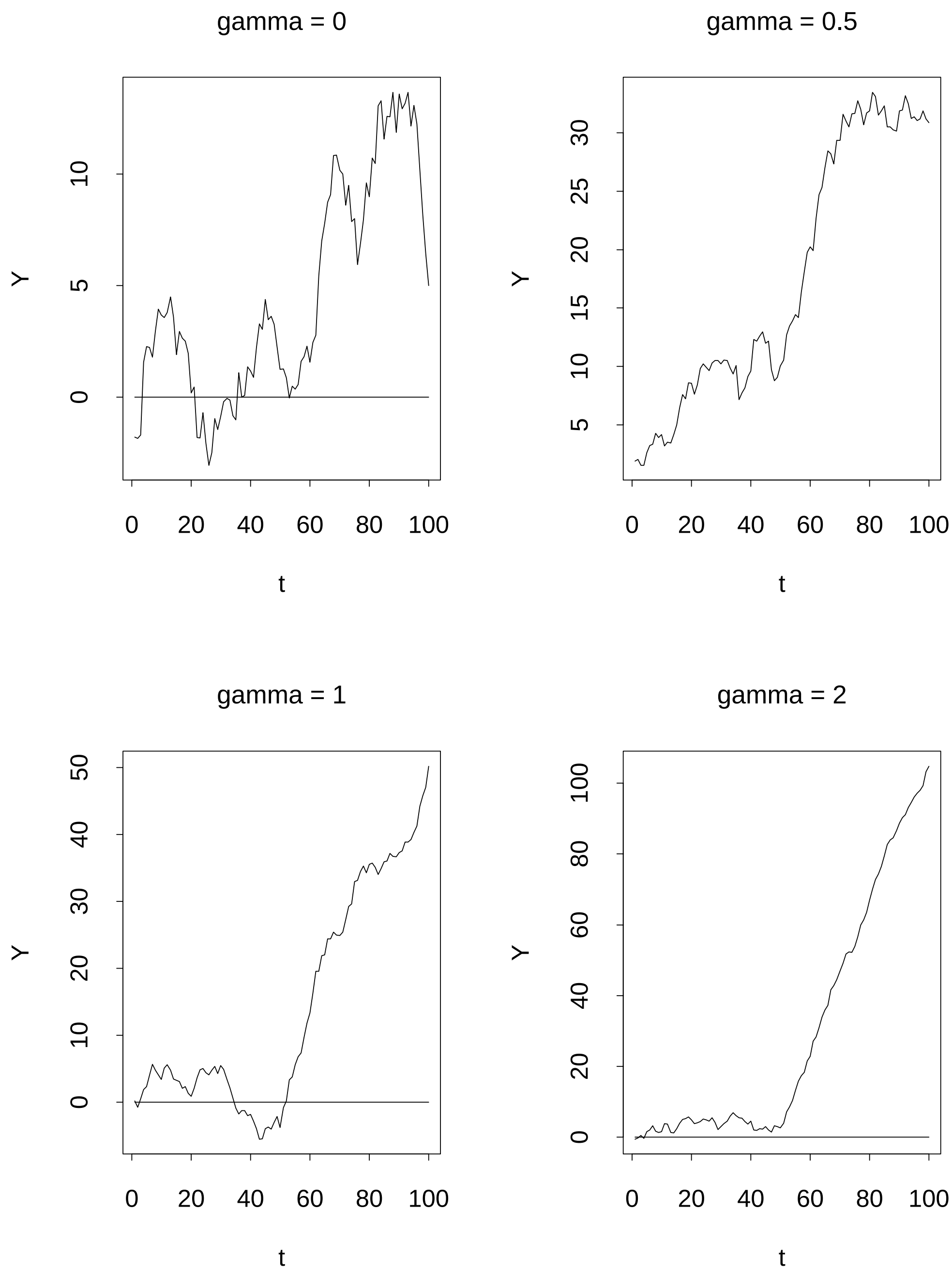

Figure 3.9: Data from a break-in-slope model (DGP (3.2), $\rho = 1$, $c = 50$)

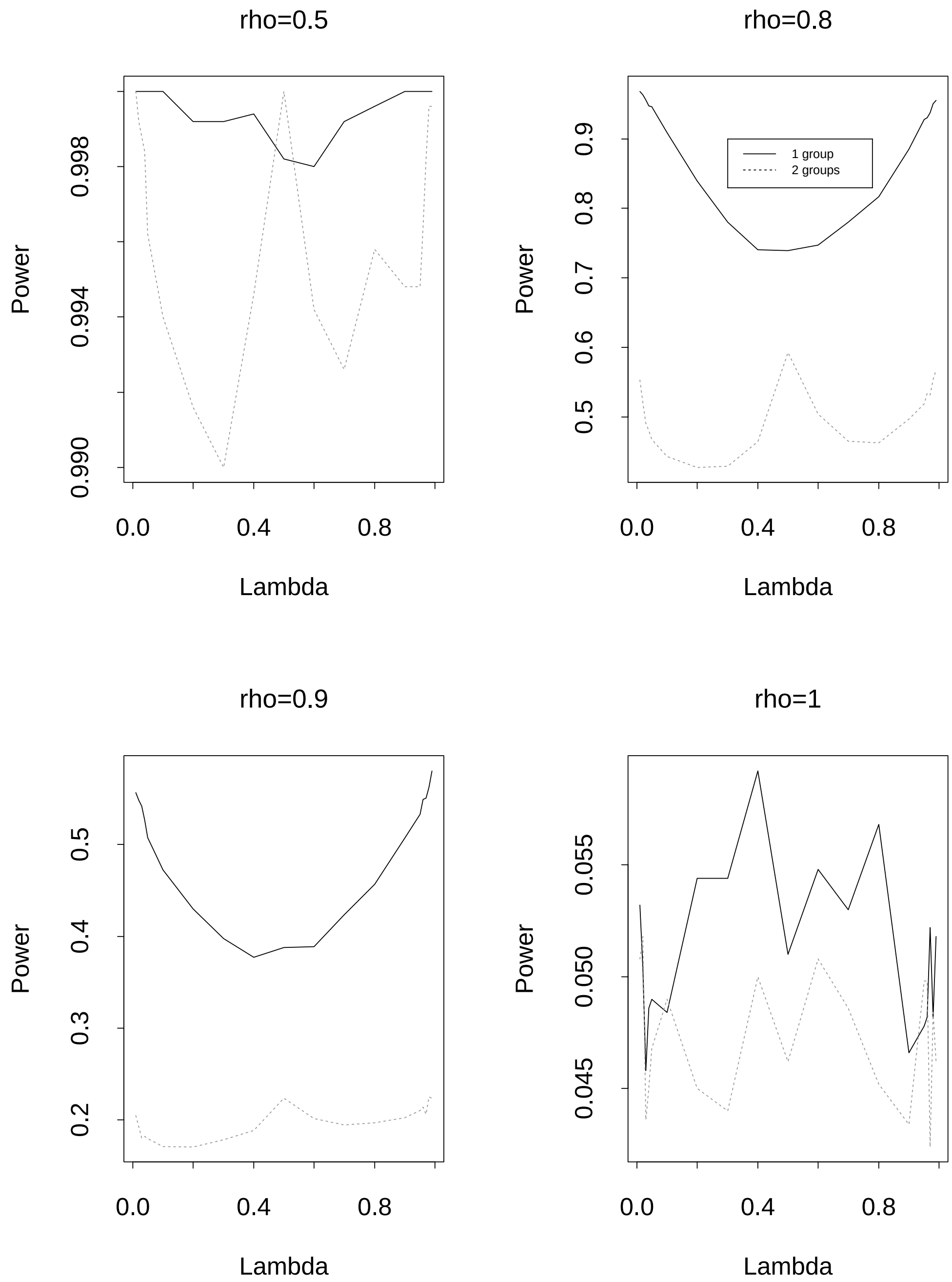

Figure 3.10: Empirical size and power for DGP (3.1) using τ_w (1 group) and τ_w^* (2 groups) ($\theta = 2.5$, $\lambda = c/n$)

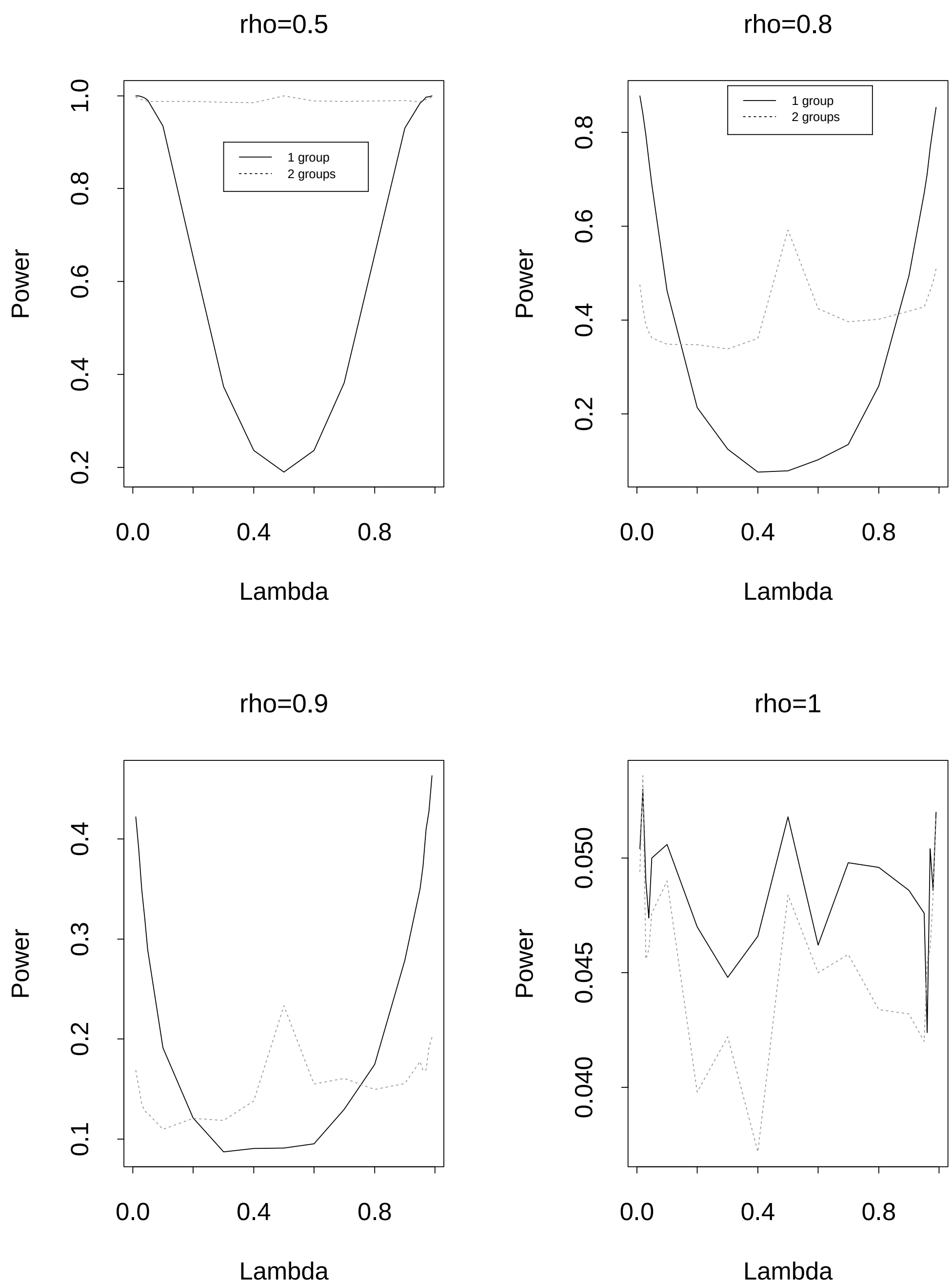

Figure 3.11: Empirical size and power for DGP (3.1) using τ_w (1 group) and τ_w^* (2 groups) ($\theta = 5$, $\lambda = c/n$)

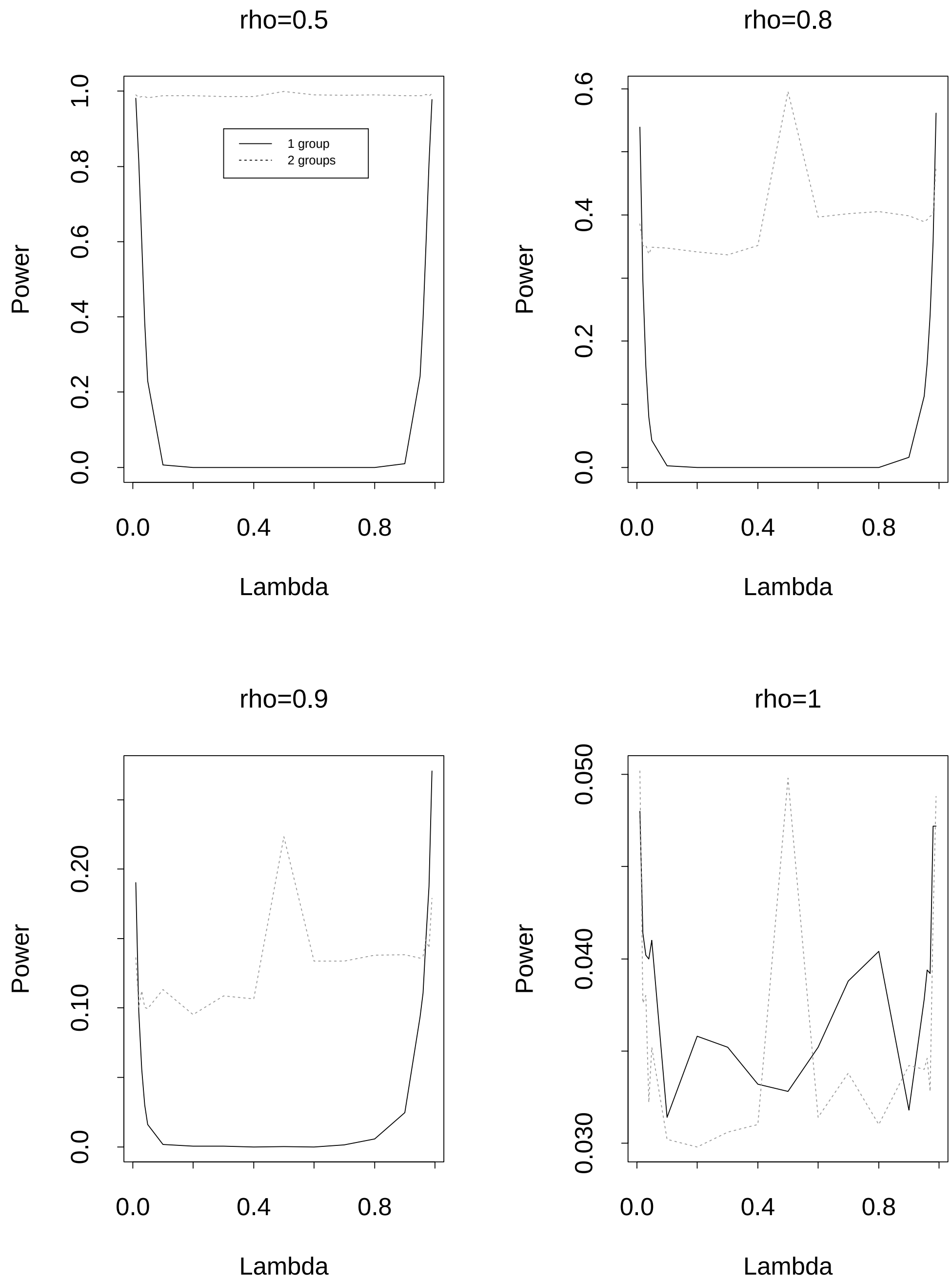

Figure 3.12: Empirical size and power for DGP (3.1) using τ_w (1 group) and τ_w^* (2 groups) ($\theta = 10$, $\lambda = c/n$)

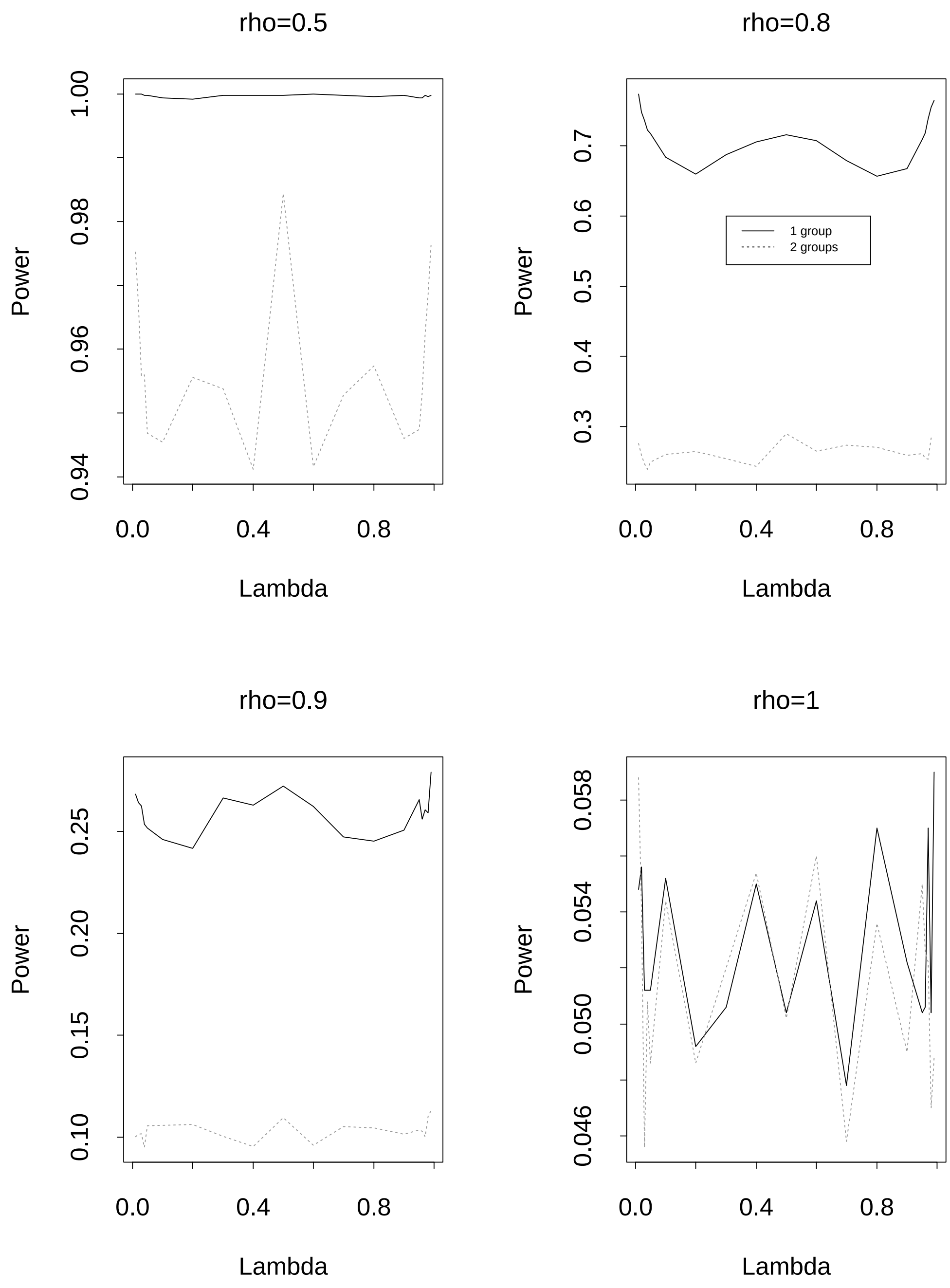

Figure 3.13: Empirical size and power for DGP (3.1) using $\tau_{w,\tau}$ (1 group) and $\tau^*_{w,\tau}$ (2 groups) ($\theta = 2.5$, $\lambda = c/n$)

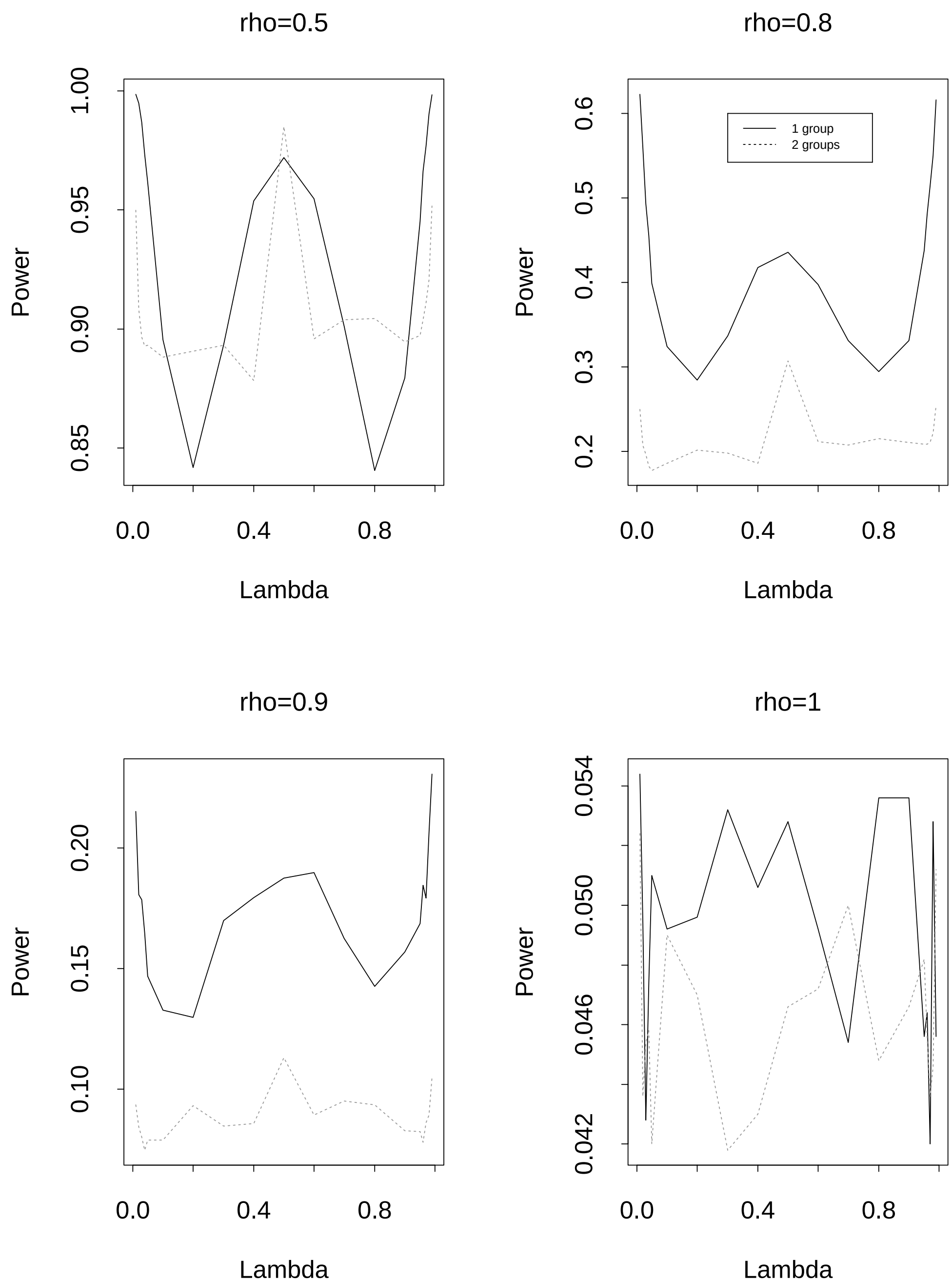

Figure 3.14: Empirical size and power for DGP (3.1) using $\tau_{w,\tau}$ (1 group) and $\tau^*_{w,\tau}$ (2 groups) ($\theta = 5$, $\lambda = c/n$)

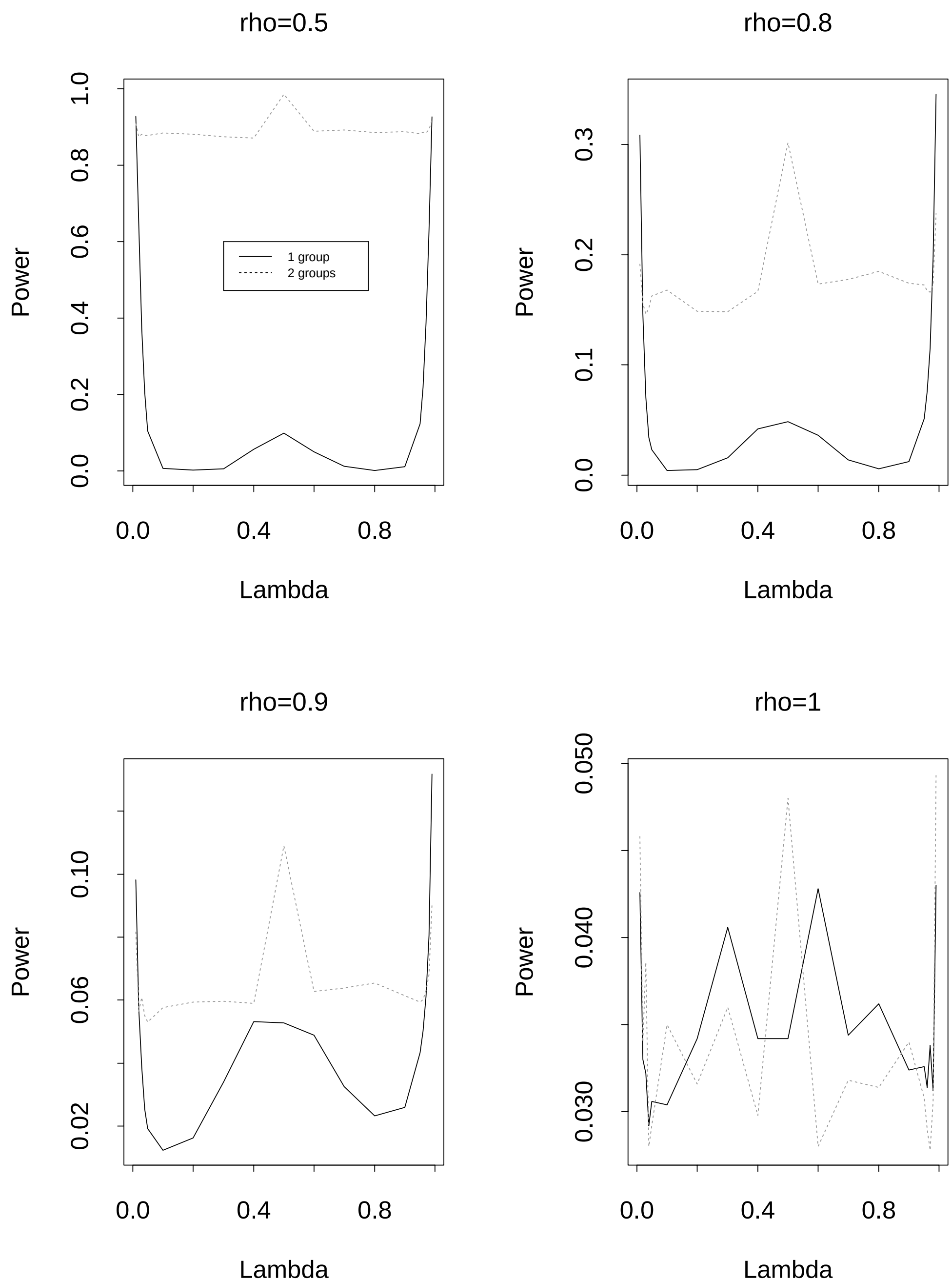

Figure 3.15: Empirical size and power for DGP (3.1) using $\tau_{w,\tau}$ (1 group) and $\tau^*_{w,\tau}$ (2 groups) ($\theta = 10$, $\lambda = c/n$)

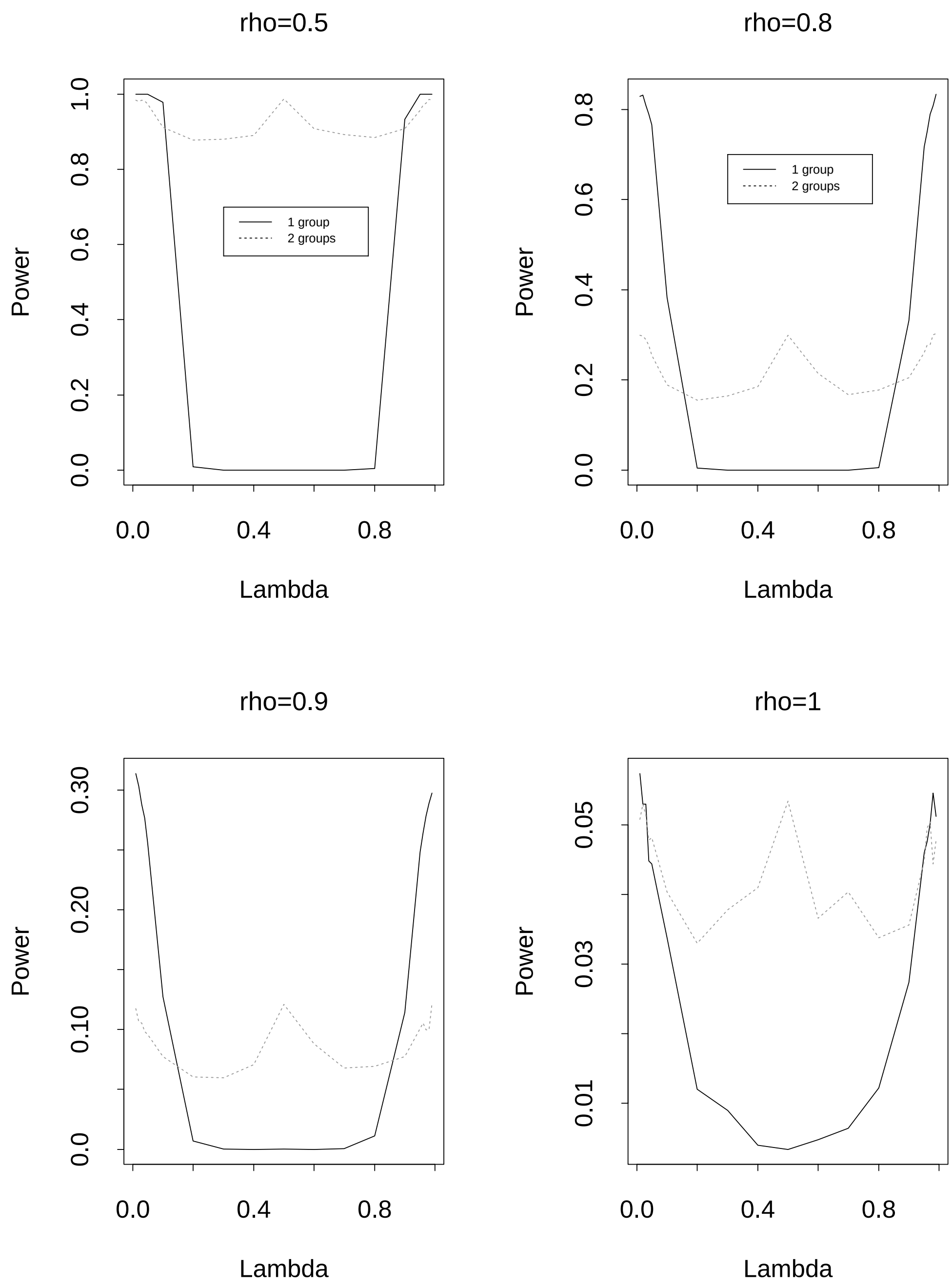

Figure 3.16: Empirical size and power for DGP (3.2) using $\tau_{w,\tau}$ (1 group) and $\tau^*_{w,\tau}$ (2 groups) ($\gamma = 0.5$, $\lambda = c/n$)

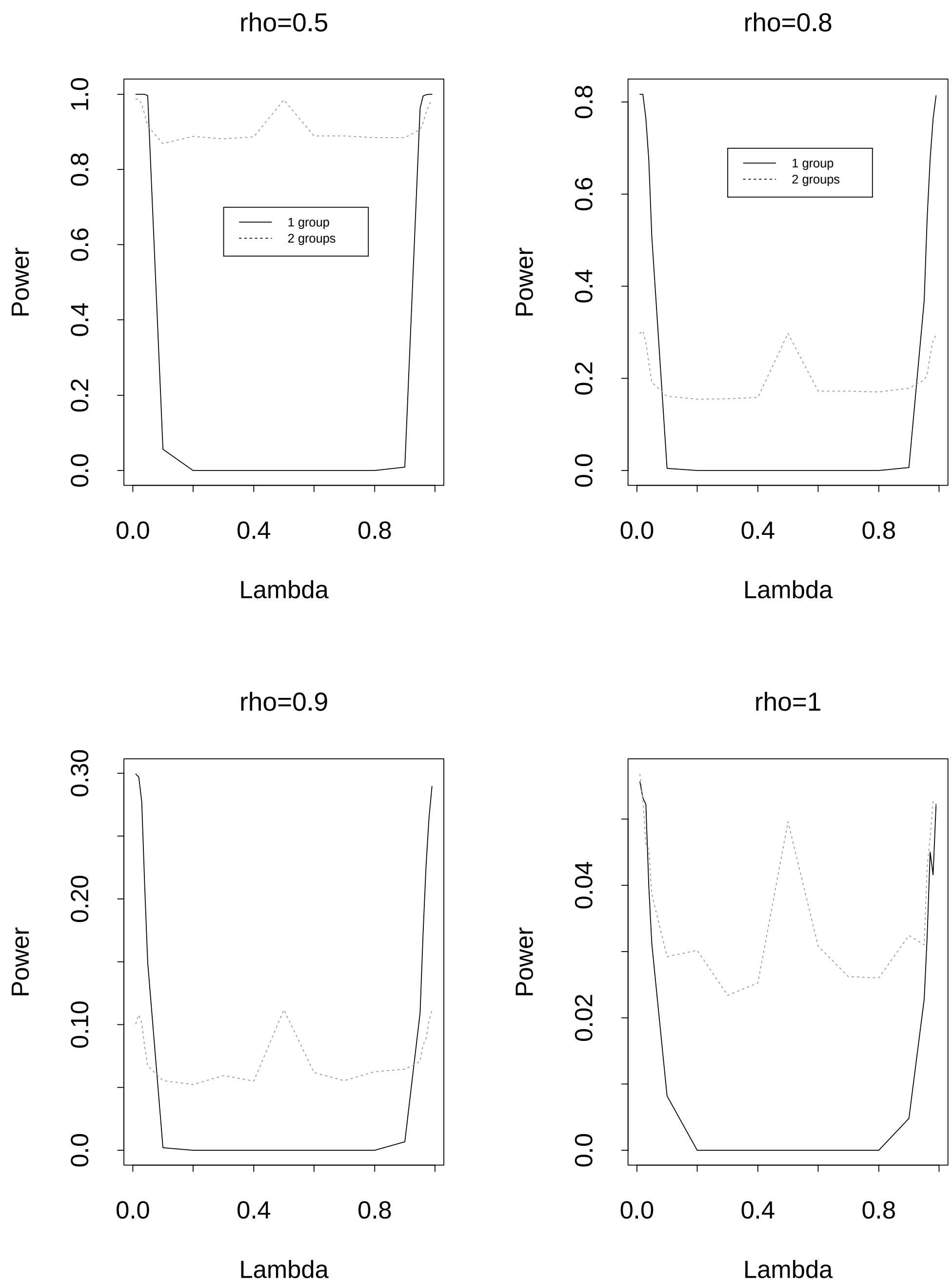

Figure 3.17: Empirical size and power for DGP (3.2) using $\tau_{w,\tau}$ (1 group) and $\tau^*_{w,\tau}$ (2 groups) ($\gamma = 1$, $\lambda = c/n$)

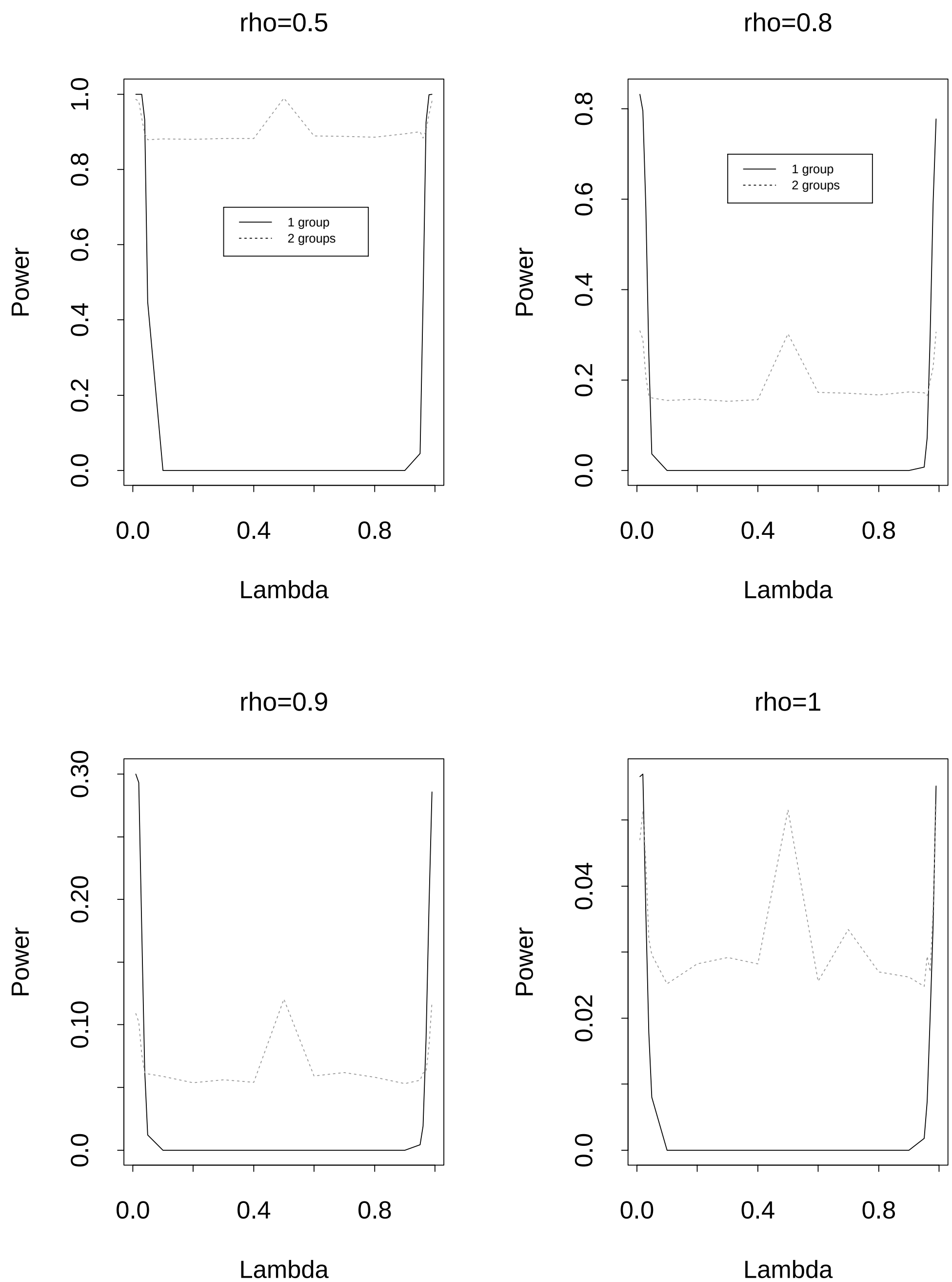

Figure 3.18: Empirical size and power for DGP (3.2) using $\tau_{w,\tau}$ (1 group) and $\tau^*_{w,\tau}$ (2 groups) ($\gamma = 2$, $\lambda = c/n$)

Chapter 4

Temporal Analysis of Hydrologic Variability in Stream Flow Data

4.1 Introduction

In recent decades, stream management policies have been established to examine extensive environmental degradation and loss of biological diversity in stream systems (Poff et al, 1997). Investigating the causes of this degradation, Poff et al (1997) reported that the quantity and timing of stream flow are critical determinants of water supply, water quality and the ecological integrity of stream systems.

Researchers have recognized five components of the flow regime as regulating ecological processes in stream ecosystems: the magnitude, frequency, duration, timing and rate of change of hydrologic conditions (Poff and Ward, 1989; Richter et al, 1996; Walker et al, 1995). These components can also be used to chracterize some specific hydrologic variables that are critical to the physical, chemical and biological conditions of stream systems (Poff et al, 1997). Variability and predictability of hydrologic chracteristics seem to be particularly important to biological conditions in streams (Poff and Allan, 1995; Clausen and Biggs; 1997).

Since 1889, the U.S. Geological Survey (USGS) has collected stream flow data needed to understand and monitor the national water resources. The USGS Streamgaging Program includes 7,292 gaging stations with varying lengths of years of stream flow data. Even though there have been many studies of hydrologic time series, to our knowledge, no empirical study has evaluated temporal properties of the particular hydrologic variability chracteristics we study here, namely, the number and duration of low and high flows.

In this Chapter, we perform a time series analysis for these hydrologic variables of interest to USGS scientists. We particularly focus on looking for the required sample size to detect a hypothetical future level shift of a given magnitude. We also cluster the stations by this sample size result to view the physiographic effect.

In section 4.2, we describe the stream flow data that we are interested in and show some preliminary results for those data. In section 4.3, we present the problem of finding the sample size needed to detect a certain level shift. We apply three different approaches to solve this problem. Section 4.4 contains the estimates and standard errors of the sample sizes for our data. Section 4.5 deals with the clustering of the stations according to various clustering methods and criteria. Finally we make some summarizing comments in section 4.6.

4.2 Preliminary results

4.2.1 Description of data

Based on preliminary analyses, we focus this study on the statistical properties of the number and duration of low and high flows as target variables. Notice that our findings in this work are based on these variables of "exceedence" and that other hydrologic variables may well yield different results.

We need formal definitions of the 4 hydrologic variables NL, DL, NH and DH of interest. NL and NH are numerical counts of the number of low and high flow pulses per year, respectively. A pulse is defined as a period when the water conditions are less than the low pulse threshold or greater than the high pulse threshold. DL and DH are the average durations, in days, of low and high flow pulses for the year, respectively. The low flow and high flow thresholds are defined as the 25th and 75th percentiles, respectively, of the flow distribution.

These measures are usually calculated by a software package called Indicators of Hydrologic Alteration (IHA) using records of daily mean discharge (cubic feet per second). In this study, instead, we use the statistical package SAS* because it is commonly available and has many useful features for statistical analysis. Using SAS, we can produce exactly the same values of the variables as IHA.

We opted to make some minor changes that we felt improved the variable definitions. When there is no low or high pulse that stretches across the boundary between years, the values of our four study variables are exactly the same in SAS and IHA. When there is a low or high pulse crossing a year boundary, the values of NL and NH are still the same, but those of DL and DH differ between our SAS program and IHA. While calculating DL and DH, IHA considers the portion of the duration in the previous year but ignores that in the current year. Our SAS program takes into account the complete duration both in the previous and in the current year. We could match IHA exactly, but prefer our modified statistic.

The stream flow data to be analyzed are taken from 50 stations in North Carolina, South Carolina and Virginia. Representing Coastal, Piedmont and Mountainous physiographic regions, these sites are relatively unaffected by artificial diversions or other human influences. The original data of daily mean discharges have quite a few seeming outliers but investigation revealed them to be valid numbers. Each station has a minimum of 14 years of continuous data where a hydrologic year begins in October and ends in the next September. The starting point of recording in the first year and ending point in the last year may differ for each station, but that does not affect the validity of the results. Our data for this study, therefore, consist of

*SAS is the registered trademark of SAS Institute, Cary, NC.

200 (50 stations $\times$ 4 variables) time series of different numbers of observations. In Table 4.1, we show some descriptive information for the stations of interest. Notice the wide range of drainage areas varying from 24.1 to 341.

4.2.2 Model fitting

Our first task is to fit the appropriate Autoregressive Moving Average (ARMA) model to each of our 200 series. By checking the autocorrelation of the residuals, we conclude that the first order autoregressive AR(1) model fits most series with no evidence of lack of fit (Brocklebank and Dickey, 1986). For the few series, shown in Table 4.2, that suggest higher order models, the AR(1) still captures almost all of the correlation structure so we use the AR(1) model throughout.

It is of importance to check whether the model is stationary or nonstationary. The detection of level shifts in nonstationary series is considerably more complicated than in stationary series. The AR(1) model is stationary if $|\rho| < 1$ where ρ is the lag 1 autoregressive coefficient. We can show that all of our data are stationary through the Dickey-Fuller unit root test (Dickey and Fuller, 1979).

Using SAS PROC AUTOREG, we also investigate whether there is a linear trend in each series. For most cases, the linear trend is insignificant. A few series (17 of our 200) in Table 4.2 showing a significant linear trend might need closer examination. Figure 4.1 displays the series of 4 variables in station 01632000 as an example. The horizontal line in each plot is the series mean.

In each series, the estimate of ρ as well as the sample mean and sample standard deviation are calculated. These statistics will be used later in section 4.5 for clustering the stations. Estimates of ρ for each station and target variable are displayed in Table 4.3. The distribution of those statistics across the stations is presented in Figure 4.2. Notice again that each station has a different number of observations.

4.3 Sample size for detecting the level shift

In the previous section, we indicated that our stream flow data are specifically chosen not to have interventions and so it is not surprising that they appear stationary. They are baseline data. Based on these data, we are particularly interested in pursuing the following research question : "How many samples would we need to detect a certain mean change if one occurs?". Therefore, for given values of δ and β, we are going to find the necessary sample size to detect a level shift of size δ with probability $1-\beta$. This is the problem of finding the sample size with which the hypothesis test of

$$H_0 : \delta = 0 \tag{4.1}$$

has power $1-\beta$ at significance level α.

We follow three alternative approaches to solve this problem. In section 4.3.1, a simple method using ordinary least squares (OLS) theory is presented. This method assumes a known error variance σ^2. In section 4.3.2, we pursue an approach using generalized least squares (GLS) without assuming σ^2 is known. Section 4.3.3 illustrates a frequency domain method for which identification of the autocorrelation structure is not needed.

4.3.1 Method I : Ordinary least squares

We consider a stationary AR(1) model which has a break of size δ in level. We assume the break occurs exactly at the midpoint of data. We can formally represent this model as

$$Y_t = \beta_0 + \delta X_t + U_t, \qquad U_t = \rho U_{t-1} + e_t, \qquad |\rho| < 1, \qquad t = 1, \cdots, 2n \tag{4.2}$$

where $X_t = (0, \cdots, 0, 1, \cdots, 1)'$, $e_t \sim NI(0, \sigma^2)$, $U_1 \sim N(0, \frac{\sigma^2}{1-\rho^2})$ and U_1 and e_t's are independent.

The OLS estimator of δ is simply the difference of two means such that

$$\hat{\delta} = \bar{Y}_2 - \bar{Y}_1$$

where

$$\bar{Y}_1 = \frac{1}{n}\sum_{t=1}^{n} Y_t$$

and

$$\bar{Y}_2 = \frac{1}{n}\sum_{t=n+1}^{2n} Y_t.$$

A proper test statistic for testing $H_0 : \delta = 0$ is given as

$$\frac{\bar{Y}_2 - \bar{Y}_1}{\sqrt{\text{Var}(\bar{Y}_2 - \bar{Y}_1)}}$$

Now we are looking for the sample size that satisfies

$$P(\frac{\bar{Y}_2 - \bar{Y}_1}{\sqrt{\text{Var}(\bar{Y}_2 - \bar{Y}_1)}} > Z_\alpha) = 1 - \beta \tag{4.3}$$

at significance level, say, $\alpha = .05$ and type II error probability, say, $\beta = .2$.

First of all, we need to find out what $\text{Var}(\bar{Y}_2 - \bar{Y}_1)$ looks like for Model (4.2). By introducing some matrices and using $\gamma(h)$ to denote the covariance between U_t and U_{t+h}, $\text{Var}(\bar{Y}_2 - \bar{Y}_1)$ can be written as

$$\text{Var}(\bar{Y}_2 - \bar{Y}_1) = \text{Var}(\bar{U}_2 - \bar{U}_1) = \boldsymbol{c}^0\boldsymbol{\Gamma}\boldsymbol{c}$$

where

$$\boldsymbol{c}^0 = (\frac{1}{n}, \cdots, \frac{1}{n}, -\frac{1}{n}, \cdots, -\frac{1}{n})$$

and

$$\boldsymbol{\Gamma} = \begin{pmatrix} \gamma(0) & \gamma(1) & \cdots & \gamma(n) & \cdots & \gamma(2n-1) \\ \gamma(1) & \gamma(0) & \cdots & \gamma(n-1) & \cdots & \gamma(2n-2) \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \gamma(n) & \gamma(n-1) & \cdots & \gamma(0) & \cdots & \gamma(n-1) \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \gamma(2n-1) & \gamma(2n-2) & \cdots & \gamma(n-1) & \cdots & \gamma(0) \end{pmatrix}.$$

After some algebra, $\mathrm{Var}(\bar{Y}_2 - \bar{Y}_1)$ can be written as

$$\begin{aligned} & \mathrm{Var}(\bar{Y}_2 - \bar{Y}_1) \\ = \ & \frac{2\gamma(0)}{n^2(1-\rho)^2}[(1-\rho^2)n - 3\rho + 4\rho^{n+1} - \rho^{2n+1}]. \end{aligned}$$

From basic statistical theory, we know that the power under $H_a : \delta > 0$ for the one-sided test is

$$\begin{aligned} 1-\beta &= P(\frac{\bar{Y}_2 - \bar{Y}_1}{\sqrt{\mathrm{Var}(\bar{Y}_2 - \bar{Y}_1)}} > Z_\alpha) \\ &= P(\frac{\bar{Y}_2 - \bar{Y}_1 - \delta}{\sqrt{\mathrm{Var}(\bar{Y}_2 - \bar{Y}_1)}} > Z_\alpha - \frac{\delta}{\sqrt{\mathrm{Var}(\bar{Y}_2 - \bar{Y}_1)}}) \\ &= P(Z > Z_\alpha - \frac{\delta}{\sqrt{\mathrm{Var}(\bar{Y}_2 - \bar{Y}_1)}}). \end{aligned}$$

where Z is a standard normal random variable. We assume $\gamma(0)$ and ρ are known and so is $\mathrm{Var}(\bar{Y}_2 - \bar{Y}_1)$. Therefore to find the sample size that satisfies (4.3), we need to solve the following equation for n,

$$Z_\alpha - \frac{\delta}{\sqrt{\mathrm{Var}(\bar{Y}_2 - \bar{Y}_1)}} = Z_{1-\beta} = -Z_\beta$$

which becomes

$$\mathrm{Var}(\bar{Y}_2 - \bar{Y}_1) = (\frac{\delta}{Z_\alpha + Z_\beta})^2.$$

For a two-sided test with $H_a : \delta \neq 0$, n can be obtained from

$$\mathrm{Var}(\bar{Y}_2 - \bar{Y}_1) = (\frac{\delta}{Z_{\alpha/2} + Z_\beta})^2$$

, i.e.,

$$\frac{2\gamma(0)}{n^2(1-\rho)^2}[(1-\rho^2)n - 3\rho + 4\rho^{n+1} - \rho^{2n+1}] = (\frac{\delta}{Z_{\alpha/2} + Z_\beta})^2 \equiv V_0 \qquad (4.4)$$

for $\delta \neq 0$.

Suppose δ is given as a multiple of σ, i.e., $\delta = k\sigma$. Since $\gamma(0) = \sigma^2/(1-\rho^2)$ for the AR(1) model, (4.4) becomes

$$\frac{2\sigma^2}{n^2(1-\rho)^2(1-\rho^2)}[(1-\rho^2)n - 3\rho + 4\rho^{n+1} - \rho^{2n+1}] = (\frac{k\sigma}{Z_{\alpha/2}+Z_\beta})^2.$$

Cancelling σ^2 on both sides, we have

$$\frac{2}{n^2(1-\rho)^2(1-\rho^2)}[(1-\rho^2)n - 3\rho + 4\rho^{n+1} - \rho^{2n+1}] = (\frac{k}{Z_{\alpha/2}+Z_\beta})^2 \equiv V_1. \quad (4.5)$$

In practice, we need to use the sample estimates $\hat{\gamma}(0)$ and $\hat{\rho}$ for the parameters $\gamma(0)$ and ρ. We can numerically solve the equations (4.4) or (4.5) for n using a search technique. In Table 4.4, we display the values of n calculated from (4.5) for several values of β, k and ρ.

One way to obtain an initial value for n is to use the value that results from setting $\rho = 0$. If $\rho = 0$, we obtain

$$n = \frac{2}{V_1} \quad (4.6)$$

in (4.5).

Another way is just to ignore the higher order terms in ρ, the terms ρ^{n+1} and ρ^{2n+1} being very small for reasonable n and $|\rho| < 1$. Then (4.5) reduces to a quadratic equation for n, namely

$$V_1(1-\rho)^2(1-\rho^2)n^2 - 2(1-\rho^2)n + 6\rho = 0. \quad (4.7)$$

Solving for n, we have

$$n = \frac{(1+\rho) + \sqrt{(1+\rho)^2 - 6V_1\rho(1-\rho^2)}}{V_1(1-\rho)(1-\rho^2)}. \quad (4.8)$$

Clearly, (4.8) should be a much better starting value than (4.6) and, for ρ not too close to 1, should provide a good approximation to n without further adjustment. That is, we can get an approximate n without any search algorithm.

4.3.2 Method II : Generalized least squares

In this section, we show that the noncentrality parameter of the test statistic for testing $H_0 : \delta = 0$ is a function of the sample size n. We already know that the power depends on the noncentrality parameter. Therefore we conclude that the sample size n can be determined once the desired value of the power is given.

Model (4.2) does not satisfy OLS assumptions because the U_t's are autocorrelated. Therefore we use GLS implemented through the so-called "Cochrane-Orcutt" transformation (Dinardo et al, 1996).

From Model (4.2), we have

$$\begin{aligned} Y_t &= \beta_0 + \delta X_t + U_t && (4.9) \\ \text{and} \qquad \rho Y_{t-1} &= \rho\beta_0 + \rho\delta X_{t-1} + \rho U_{t-1}. && (4.10) \end{aligned}$$

Subtracting (4.10) from (4.9), we get

$$\begin{aligned} Y_t - \rho Y_{t-1} &= (1-\rho)\beta_0 + \delta(X_t - \rho X_{t-1}) + U_t - \rho U_{t-1} \\ &= (1-\rho)\beta_0 + \delta(X_t - \rho X_{t-1}) + e_t \qquad \text{for } t \geq 2. \end{aligned}$$

For $t = 1$, Y_1 can be written as

$$Y_1 = \beta_0 + \delta X_1 + U_1 = \beta_0 + U_1$$

since $X_1 = 0$. Multiplying by $\sqrt{1-\rho^2}$ on both sides yields

$$\sqrt{1-\rho^2}Y_1 = \sqrt{1-\rho^2}\beta_0 + \sqrt{1-\rho^2}U_1.$$

Therefore our transformed model in matrix form is

$$\begin{pmatrix} \sqrt{1-\rho^2}Y_1 \\ Y_2-\rho Y_1 \\ \vdots \\ Y_n-\rho Y_{n-1} \\ Y_{n+1}-\rho Y_n \\ Y_{n+2}-\rho Y_{n+1} \\ \vdots \\ Y_{2n}-\rho Y_{2n-1} \end{pmatrix} = \begin{pmatrix} \sqrt{1-\rho^2}\beta_0 \\ (1-\rho)\beta_0 \\ \vdots \\ (1-\rho)\beta_0 \\ (1-\rho)\beta_0+\delta \\ (1-\rho)\beta_0+(1-\rho)\delta \\ \vdots \\ (1-\rho)\beta_0+(1-\rho)\delta \end{pmatrix} + \begin{pmatrix} \sqrt{1-\rho^2}U_1 \\ e_2 \\ \vdots \\ e_n \\ e_{n+1} \\ e_{n+2} \\ \vdots \\ e_{2n} \end{pmatrix}$$

$$= \begin{pmatrix} \sqrt{1-\rho^2} & 0 \\ 1-\rho & 0 \\ \vdots & \vdots \\ 1-\rho & 0 \\ 1-\rho & 1 \\ 1-\rho & 1-\rho \\ \vdots & \vdots \\ 1-\rho & 1-\rho \end{pmatrix} \begin{pmatrix} \beta_0 \\ \delta \end{pmatrix} + \boldsymbol{\epsilon}$$

This model is of the form

$$\boldsymbol{Z} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon} \tag{4.11}$$

where $\boldsymbol{\epsilon} \sim N(\mathbf{0}, \boldsymbol{I}\sigma^2)$ since $\sqrt{1-\rho^2}U_1 \sim N(0,\sigma^2)$ and U_1 is independent of the e_t's.

Now Model (4.11) satisfies the Gauss-Markov assumption. We are going to test the hypothesis (4.1) and find the sample size that gives power 1-β at significance level α. Notice here that the GLS estimate $\hat{\delta}_G$ is not exactly the same as the OLS estimate $\bar{Y}_2 - \bar{Y}_1$ but is asymptotically equivalent for $|\rho| < 1$.

4.3.2.1 Noncentrality parameter

In the theory of testing the general linear hypothesis, the sum of squares for the linear hypothesis $H_0 : \boldsymbol{K}^0\boldsymbol{\beta} = \boldsymbol{m}$ is $Q = (\boldsymbol{K}^0\hat{\boldsymbol{\beta}} - \boldsymbol{m})'[\boldsymbol{K}^0(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{K}]^{-1}(\boldsymbol{K}^0\hat{\boldsymbol{\beta}} - \boldsymbol{m})$. Furthermore, a proper test statistic for $H_0 : \boldsymbol{K}^0\boldsymbol{\beta} = \boldsymbol{m}$ is

$$F = \frac{Q/r(\boldsymbol{K}^0)}{\text{MSE}}$$

where $r()$ stands for the rank of a matrix and MSE stands for the error mean square (Rawlings et al, 1998). Notice that this test statistic F is obtained without assuming σ^2 is known.

Applying the above to our case, we have

$$\boldsymbol{K}^0 = (0 \quad 1) \qquad , \qquad \boldsymbol{\beta} = \begin{pmatrix} \beta_0 \\ \delta \end{pmatrix}$$

and

$$\boldsymbol{m} = 0$$

so $r(\boldsymbol{K}^0) = 1$ and $Q = \hat{\delta}_G^2/c_{11}$ where c_{11} is the (2,2)th element of $(\boldsymbol{X}'\boldsymbol{X})^{-1}$.

The test statistic for the hypothesis $H_0 : \delta = 0$ becomes

$$F = \frac{\hat{\delta}_G^2/c_{11}}{\text{MSE}}.$$

This has a noncentral F distribution with 1 and $2n - 2$ degrees of freedom and noncentrality parameter

$$\lambda = \frac{\delta^2}{2c_{11}\sigma^2} = \frac{k^2}{2c_{11}}$$

where $\delta = k\sigma$. By some tedious algebra, we can show that

$$c_{11} = \frac{2}{(1-\rho)^2 n - \rho(\rho - 3)}$$

and so

$$\lambda = \frac{k^2\{(1-\rho)^2 n - \rho(\rho - 3)\}}{4}.$$

Under (4.1), our test statistic F is distributed as the central F distribution with 1 and $2n-2$ degrees of freedom. Using F_α, the critical value from the central F distribution at level α, the power for a certain δ is the probability that the noncentral F exceeds F_α. Having found λ that delivers the desired power, we then compute the n that gives this λ.

4.3.2.2 Comparison of theoretical and empirical powers

We showed that our test statistic F for testing the hypothesis $H_0 : \delta = 0$ is distributed according to a noncentral F distribution with 1 and $2n-2$ df and noncentrality parameter λ. For specific values of $\delta = k\sigma$, we can find the power since it is the probability that the noncentral F will exceed the critical value from the central F. The power is a function of λ and thus, of k, n and ρ. See "power1" in Table 4.5.

As a check on the theory, we have a Monte Carlo simulation to produce empirical powers.

Two kinds of empirical powers, denoted as power2 and power3 appear in Table 4.5. For the former, we use the values of ρ as if they are known. Our simulations involve 10,000 replicates at each (k, n, ρ) combination. The Monte Carlo standard error is no more than $\sqrt{(.5)(.5)/10000} = 0.005$.

For the latter, we generate 10,000 AR(1) series, and hence 10,000 $\hat{\rho}$'s, for each ρ. Then we calculate the theoretical power substituting $\hat{\rho}$ for ρ, and obtain the average of them. This is to examine how much variation we introduce when $\hat{\rho}$ is used instead of the unknown ρ.

Table 4.5 shows theoretical and two empirical powers for values of n, ρ and $k = 1$. As expected, for any given value of ρ, power becomes higher as n becomes larger. For fixed n, power becomes lower as ρ gets larger. Notice that the empirical results are generally in close agreement with the theoretical ones. However, power3 is slightly higher than power1 for $\rho > 0$. This is because of the downward bias in $\hat{\rho}$ which can be confirmed by examining its empirical distribution.

In practice, if a 95% confidence interval (CI) for ρ is given as $\hat{\rho}_L \le \rho \le \hat{\rho}_U$, then the corresponding 95% CI for the power can be obtained for a fixed n using power1 in Table 4.5. Furthermore, an interval for the sample size n, $\hat{n}_L \le n \le \hat{n}_U$, can also be obtained for a fixed power. For example, suppose a 95% CI for ρ is $(.3, .5)$. Then for $n = 20$, a 95% CI for the power is $(.406, .612)$ and, for a desired power of 0.5, a 95% CI for n is available by an interpolation.

Figure 4.3 displays the theoretical powers for some values of n, k and ρ.

4.3.3 Method III : Frequency domain method

In this section, we are going to use spectral or "frequency domain" methods. To obtain the sample size satisfying (4.3), $\text{Var}(\bar{Y}_2 - \bar{Y}_1)$ from spectral theory is needed.

We first find the limiting behavior of $n\text{Var}(\bar{Y}_2 - \bar{Y}_1)$. We know that

$$n\text{Var}(\bar{Y}_2 - \bar{Y}_1) = n\text{Var}(\bar{Y}_1) + n\text{Var}(\bar{Y}_2) - 2n\text{Cov}(\bar{Y}_1, \bar{Y}_2).$$

From Theorem 6.1.2, Fuller (1996, p310),

$$\begin{aligned} \lim_{n\to\infty} n\text{Var}(\bar{Y}_1) &= 2\pi f(0) \\ \text{and } \lim_{n\to\infty} n\text{Var}(\bar{Y}_2) &= 2\pi f(0) \end{aligned} \tag{4.12}$$

where $f(0)$ is the spectral density of Y_t evaluated at zero.

Now,

$$\begin{aligned} & n\text{Cov}(\bar{Y}_1, \bar{Y}_2) \\ = \; & n\text{Cov}\left(\frac{1}{n}\sum_{t=1}^{n} Y_t, \frac{1}{n}\sum_{t=n+1}^{2n} Y_t\right) \\ = \; & \frac{1}{n}\text{Cov}\left(\sum_{t=1}^{n} Y_t, \sum_{t=n+1}^{2n} Y_t\right) \\ = \; & \frac{1}{n}\{\gamma(1) + \cdots + (n-1)\gamma(n-1) + n\gamma(n) + (n-1)\gamma(n+1) + \cdots + \gamma(2n-1)\}. \end{aligned}$$

For stationary ARMA processes, we can show that

$$|n\text{Cov}(\bar{Y}_1, \bar{Y}_2)| \le \frac{M}{n}\sum_{i=1}^{2n-1} \varepsilon^i \tag{4.13}$$

where M is a finite constant and ε is a constant in (0,1). See Fuller (1996, p106). Hence,

$$n\text{Cov}(\bar{Y}_1, \bar{Y}_2) \longrightarrow 0 \qquad \text{as} \qquad n \longrightarrow \infty. \tag{4.14}$$

By (4.12) and (4.14), we obtain

$$\lim_{n\to\infty} n\text{Var}(\bar{Y}_2 - \bar{Y}_1) = 2 \cdot 2\pi f(0) = 4\pi f(0).$$

For reasonably large n, $\text{Var}(\bar{Y}_2 - \bar{Y}_1)$ can be approximated as

$$\text{Var}(\bar{Y}_2 - \bar{Y}_1) \doteq \frac{4\pi f(0)}{n}.$$

Similar reasoning to (4.4) leads us to find the sample size n satisfying

$$\frac{4\pi f(0)}{n} = (\frac{\delta}{Z_{\alpha/2} + Z_\beta})^2 = V_0.$$

Therefore the required sample size using spectral theory is

$$n = \frac{4\pi f(0)}{V_0}. \tag{4.15}$$

This method requires no knowledge of the autocorrelation structure.

In practice, the spectral density $f(\omega)$ is estimated by the smoothed periodogram. Using SAS PROC SPECTRA, we can obtain spectral density estimates. Since our data do not show strong autocorrelations, we apply a long weight sequence.

If the series is white noise, then $f(\omega) = \frac{\sigma^2}{2\pi}$ and so (4.15) becomes

$$n = \frac{2\sigma^2}{V_0} = \frac{2\sigma^2(Z_{\alpha/2} + Z_\beta)^2}{\delta^2}.$$

This formula for n is exactly the same as the one for two independent samples in elementary statistics.

4.4 Results for stream flow data

4.4.1 Sample size estimates

Our original goal was to seek the necessary sample size to detect a shift δ for a certain power $1-\beta$ at significance level α.

Method I tells us that the sample size n can be obtained by using the equation (4.8) when $\delta = k\sigma$. From Method II, we know that the power depends upon the values of k, n and ρ. This means that we can obtain the sample size n once the values of k, ρ and the power are given. That is, we can find the required sample size n to detect a level shift of size $\delta = k\sigma$ with a certain power. When adopting either method, we need to use the estimate $\hat{\rho}$ as if it is the fixed parameter ρ. In Method III, we can calculate the sample size n for given δ using (4.15).

Tables 4.6 and 4.7 present the sample size n needed to detect 1σ level shift with $\beta = .2$ using Method I and II, respectively. Sample sizes calculated using Method I, where $k = 1$ and $\beta = .2$, ranged from 9 years for NH to 368 years for NL (Table 4.6). Mean sample sizes are 29.7, 27.6, 21.2 and 18.9 years for NL, DL, NH and DH respectively. Mean values are 28.6 years for measures of low-flow variability (NL and DL combined) and 20.0 years for measures of high-flow variability (NH and DH combined).

From Method II, the necessary sample size n to detect a level shift of size $\delta = k\sigma$ where $k = 1$ and $\beta = .2$ ranged from 8 years for NH to 334 years for NL (Table 4.7). Mean sample sizes are 29.7, 27.9, 21.9 and 19.7 years for NL, DL, NH and DH respectively. Mean values are 28.8 years for measures of low-flow variability (NL and DL combined) and 20.8 years for measures of high-flow variability (NH and DH combined). In Table 4.8, the sample sizes for $\delta = 3$ by Method III are displayed.

In our data, we have 50 $\hat{\rho}$'s from 50 stations for each hydrologic variable: NL, DL, NH and DH. For each variable, we now consider using $\hat{\rho}_0$, the weighted average of 50 $\hat{\rho}$'s, as a common estimate for all 50 stations. The number of observations in each station is the weighting factor and so

$$\hat{\rho}_0 = \frac{\sum_{j=1}^{50} T_j \hat{\rho}_j}{\sum_{j=1}^{50} T_j}$$

where T_j and $\hat{\rho}_j$ are the number of observations and the autocorrelation estimate for station j. This would be justified if the $\hat{\rho}$'s were estimating a common ρ_0.

Suppose we are going to test $H_0 : \rho_j = \rho_0$ where ρ_j is the lag 1 autocorrelation for station j and ρ_0 is a fixed value of ρ. In the AR(1) model, we know that if $|\rho| < 1$,

$$\sqrt{T}(\hat{\rho} - \rho) \xrightarrow{L} N(0, 1 - \rho^2) \qquad \text{as} \qquad T \to \infty \tag{4.16}$$

or

$$\hat{\rho} \sim AN(\rho, \frac{1 - \rho^2}{T})$$

where T is the number of observations in a series. Therefore, under $H_0 : \rho_j = \rho_0$ for $j = 1, \cdots, 50$, we know that

$$\frac{\hat{\rho}_j - \rho_0}{\sqrt{(1 - \rho_0^2)/T_j}} \tag{4.17}$$

is approximately distributed as $N(0, 1)$ and so

$$\sum_{j=1}^{50} \frac{(\hat{\rho}_j - \rho_0)^2}{(1 - \rho_0^2)/T_j}$$

is approximately distributed as χ^2_{50}. Using a weighted average $\hat{\rho}_0$ instead of ρ_0 in (4.17), we test to see if

$$\hat{Z}_j \equiv \frac{\hat{\rho}_j - \hat{\rho}_0}{\sqrt{(1 - \hat{\rho}_0^2)/T_j}}$$

is $N(0, 1)$ for $j = 1, \cdots, 50$ or

$$\hat{H} \equiv \sum_{j=1}^{50} \frac{(\hat{\rho}_j - \hat{\rho}_0)^2}{(1 - \hat{\rho}_0^2)/T_j}$$

is distributed as χ^2_{49} where the degree of freedom is adjusted due to the constraint $\sum_{j=1}^{50} T_j(\hat{\rho}_j - \hat{\rho}_0) = 0$. Using the Kolmogorov test in SAS/INSIGHT with $\hat{\rho}_0$ in place of ρ_0, we do not reject the hypothesis that these $\hat{Z}_j$'s are $N(0,1)$ and thus do not encounter an argument against using the common $\hat{\rho}_0$. The test statistic $\hat{H}$ for each target variable also finds that ρ_j is not significantly different from $\hat{\rho}_0$ (p-values are 0.63, 0.29, 0.05 and 0.81 for NL, DL, NH and DH respectively).

The weighted averages of $\hat{\rho}$'s, with T_j's as weights, for 4 target variables are $\hat{\rho}_0 = \{0.14902, 0.15083, 0.055525, 0.043875\}$. If we use them in the equation (4.8), Method I gives a set of sample sizes {22, 22, 18, 18} for $k = 1$. By Method II for $k = 1$, we obtain {23, 23, 19, 19} as a set of sample sizes. Notice that if we just ignore ρ (as if $\rho = 0$) then the necessary sample sizes are 16 and 17 in Method I and II, respectively. As we already know, this argument does not apply to Method III which is robust to model specification.

Since GLS, used in Method II, is a better approach than OLS used in Method I, we would expect sample sizes obtained from Method I to be bigger than those obtained from Method II. Results in Tables 4.6 and 4.7, however, show that this advantage of Method II is somewhat reduced by the additional variability due to admitting that σ^2 is unknown. If we also assume unknown σ^2 in Method I, and use the t distribution instead of the Normal distribution, we have slightly different results as shown in Table 4.9. Under the same assumption about σ^2, the superiority of GLS to OLS is more clearly seen by comparing numbers in Tables 4.7 and 4.9. The only substantial differences are in stations with large ρ values (n_2 of Station 02015700, n_1 and n_2 of Station 2197300) as might be expected.

To compare the time domain method (TDM, Method I) with the frequency domain method (FDM, Method III), we present plots of sample sizes (Figures 4.4 - 4.5) for $\delta = 3$. For the frequency domain method, triangular weights and flat weights are applied. In those plots, the reference line $y = x$ is drawn to highlight possible outliers. As a whole, the sample sizes from the two methods seem to match well.

To check for linear relationships among sample size estimates in Table 4.6, we calculate Pearson correlation coefficients for each pair of variables and display them in the correlation matrix in Table 4.10. We notice that n_1 and n_2, n_1 and n_4, and n_2 and n_4 are significantly correlated.

When some possible outliers in the sample size estimates are discarded, however, results are somewhat different as can be seen in Table 4.11. Without 3 outlying stations (02015700, 02131309 and 02197300), we conclude that only n_1 (NL) and n_2 (DL), and n_1 and n_4 (DH), are highly correlated.

Principal component analysis can be used to further explore the correlation structure of the 4 variables (n_1, n_2, n_3 and n_4) in Table 4.6. Using SAS PROC PRINCOMP with 3 outlying stations excluded again, we obtain eigenvectors as

$$\begin{pmatrix} 0.655154 & -.203874 & -.054787 & -.725401 \\ 0.539751 & -.495899 & 0.314705 & 0.603087 \\ 0.262411 & 0.697406 & 0.666841 & -.009371 \\ 0.458892 & 0.475546 & -.673262 & 0.331651 \end{pmatrix}.$$

The 4 principal component vectors called P1, P2, P3 and P4 are then calculated as in Table 4.12. The proportion of the total sum of squares accounted for by the first 2 principal components is about 70.4%. The first 3 principal components account for about 89.3% of the total sum of squares of the 4 original variables. The plot of the first 2 principal components, Figure 4.6, shows that 4 stations (22, 29, 41 and 45) in the lower right corner are somewhat different from the other stations. Serial numbers in Table 4.1 are used in the plot to identify the stations. That is, 22 stands for Station 02065500, 29 for Station 02091700, 41 for Station 02143000 and 45 for Station 02153780. Among the 4 stations, Station 02153780 (45) and Station 02065500 (22) differ from each other primarily in the first principal component.

4.4.2 Measure of accuracy

To gain further insight into the accuracy of the sample sizes obtained above, we use a standard linearization technique.

Ignoring more terms for simplicity in (4.5), we have

$$n = \frac{2}{V_1(1-\rho)^2} \equiv h(\rho).$$

By Taylor series approximation,

$$\hat{n} = h(\hat{\rho}) = h(\rho) + h^{'}(\rho)(\hat{\rho} - \rho) + R$$

where R is an appropriate remainder term and

$$h^{'}(\rho) = \frac{4}{V_1(1-\rho)^3}.$$

Assuming R is ignorable,

$$\hat{n} - n = h^{'}(\rho)(\hat{\rho} - \rho).$$

Therefore, by (4.16), we have

$$\text{Var}(\hat{n}) = [h^{'}(\rho)]^2 \text{Var}(\hat{\rho}) \doteq [\frac{4}{V_1(1-\rho)^3}]^2 \frac{1-\rho^2}{T}$$

and so

$$\text{s.e.}(\hat{n}) = \frac{4}{V_1(1-\rho)^3} \cdot \frac{\sqrt{1-\rho^2}}{\sqrt{T}} = \frac{4\sqrt{1-\rho^2}(Z_{\alpha/2} + Z_\beta)^2}{(1-\rho)^3 k^2} \cdot \frac{1}{\sqrt{T}}.$$

That is, the standard error of $\hat{n}$ is inversely proportional to the square root of T for given k and ρ. We can calculate the standard error of $\hat{n}$ for each station and target variable. In Table 4.6, it ranges between 1.28 (n_3 of Station 2129590) and 418.66 (n_1 of Station 2197300). For $\alpha = .05$, $\beta = .2$ and the maximum $\hat{\rho}_0$ (=.151), s.e.$(\hat{n}) = 50.715/(\sqrt{T}k^2)$.

Furthermore, after some algebra, we can show that

$$\frac{\text{s.e.}(\hat{n})}{\hat{n}} = \frac{(1+\rho)}{(1+\rho) + \sqrt{(1+\rho)^2 - 6V_1\rho(1-\rho^2)}} \cdot \frac{4\sqrt{1-\rho^2}}{(1-\rho)\sqrt{T}}.$$

s.e.$(\hat{n})/\hat{n}$ varies between 0.157 (n_3 of Station 2129590) and 1.139 (n_1 of Station 2197300) for sample size estimates in Table 4.6. That means we can expect the standard errors of $\hat{n}$ lie in between 15.7% and 113.9% of $\hat{n}$.

In Method III, we have

$$n = \frac{4\pi f(0)}{V_0} \equiv g[f(0)].$$

Again by Taylor series expansion,

$$\hat{n} = g[\hat{f}(0)] = g[f(0)] + g^{'}[f(0)][\hat{f}(0) - f(0)] + R$$

where

$$g^{'}[f(0)] = \frac{4\pi}{V_0}.$$

By Corollary 7.2.2, Fuller (1996, p373),

$$\text{Var}(\hat{f}(0)) = \frac{2}{2d_n + 1} f^2(0) + o(d_n^{-1})$$

where d_n is an increasing sequence of positive integers used as weights in computing $\hat{f}(0)$ and satisfying $\lim_{n\to\infty} d_n = \infty$ and $\lim_{n\to\infty} \frac{d_n}{n} = 0$.

Assuming R is ignorable,

$$\hat{n} - n = g^{'}[f(0)][\hat{f}(0) - f(0)]$$

and so

$$\text{Var}(\hat{n}) = \{g^{'}[f(0)]\}^2 \text{Var}[\hat{f}(0)] \doteq (\frac{4\pi}{V_0})^2 \frac{2}{2d_n + 1} f^2(0).$$

Therefore s.e.$(\hat{n})$ can be written as

$$\text{s.e.}(\hat{n}) = \frac{4\pi}{V_0}\sqrt{\frac{2}{2d_n+1}} f(0) = \frac{4\pi(Z_{\alpha/2} + Z_\beta)^2}{\delta^2}\sqrt{\frac{2}{2d_n+1}} f(0) = \sqrt{\frac{2}{2d_n+1}}\hat{n}.$$

Approximate 95% confidence interval endpoints are thus given as

$$\hat{n} \pm 1.96 \times \text{s.e.}(\hat{n}) = \hat{n}(1 \pm 1.96\sqrt{\frac{2}{2d_n+1}}).$$

4.5 Cluster analysis

In this section, we consider geographical clustering of the stations using a statistical tool called cluster analysis. Cluster analysis is a set of statistical techniques used to find groups of observations that are statistically similar. Observations in a cluster are similar and observations not in the same cluster are dissimilar.

SAS offers more than 10 different clustering methods including single linkage, complete linkage, average linkage and centroid. All methods are based on the usual agglomerative hierarchical clustering sequence. Each observation begins in a cluster by itself. The two closest clusters are merged to form a new cluster that replaces the two old clusters. Merging of the two closest clusters is repeated until all observations lie in the last big cluster. The various clustering methods differ in how the distance between two clusters is computed. For more details about cluster analysis, refer to SAS Institute Inc. (1989).

We can also use several criteria on which to base the clustering. One of them is the set of descriptive statistics {sample mean, sample standard deviation, $\hat{\rho}$} mentioned in section 4.2. Another is the set of sample sizes $\{n_1, n_2, n_3, n_4\}$ in section 4.4. We need to be cautious in interpreting clustering results since the standard errors of sample size estimates are all different among the stations as shown in section 4.4.2.

Since clustering continues until only one cluster is left, we need to stop with an appropriate number of clusters. The number of clusters can be determined by the so-called Cubic Clustering Criterion (CCC). A reasonable rule for using the CCC is that clusterings corresponding to peaks with values greater than 2 are candidates for the appropriate number of clusters. There may be several peaks, and interpretability should always play a role in the judgement about the appropriate number of clusters.

Figures 4.7 (from Table 4.13) and 4.8 are the maps of clusters based on the descriptive statistics and the sample sizes, respectively. We use the number of clusters that maximizes CCC. Figure 4.9 displays another clustering result by the sample sizes. This time we pick the number of clusters so that the largest cluster contains no more than 25 stations. Notice that the clustering results can be different according to the clustering methods, criteria or how we apply the CCC. We conclude that our clusters do not correspond to any specific physiographic segregation. See SAS Institute Inc. (1994) for more details about SAS graphics.

To further examine the effect of some physiographic variables on our clustering results, we regress each sample size $\{n_1, n_2, n_3, n_4\}$ on elevation and drainage area using SAS PROC REG. For various values of δ or k, the p-values for the model F-test vary from .08 to .78 and show no significance at $\alpha = .05$. We thus have no evidence that these particular physiographic variables, elevation and drainage area, have influence on the sample sizes for these particular rivers. For reference, we include some 3 dimensional plots of elevation, drainage area and the sample size from SAS PROC G3D (Figures 4.10 - 4.11).

4.6 Summary

In this paper, we perform a temporal analysis of some hydrologic variables in stream flow data. The data are taken from 50 gaging stations in three states (NC, SC and VA) and we are interested in 4 hydrologic variables: number and duration of low and high flows (NL, DL, NH and DH).

Results of a preliminary analysis are as follows. The AR(1) model fits well for most series. All the series are verified to be stationary by the Dickey-Fuller test. Linear trends do not appear to exist in most cases.

The problem of finding the sample size needed to detect a level shift is thoroughly discussed through 3 different approaches. Some comments to compare the 3 approaches are included. We also mention a measure of accuracy for the sample size estimates.

The average sample sizes needed to detect 1σ level shift are 29.7, 27.6, 21.2 and 18.9 years for NL, DL, NH and DH, respectively, based on Method I (Table 4.6).

Those from Method II (Table 4.7) are 29.7, 27.9, 21.9 and 19.7 years for NL, DL, NH and DH respectively. From the tables, the mean sample size estimates determined by ordinary least squares were 28.6 years for measures of low-flow variability (NL and DL combined) and 20.0 years for measures of high-flow variability (NH and DH combined). These means become 28.8 and 20.8 when generalized least squares is used.

We have a much wider range of sample sizes for NL and DL than for NH and DH. To be consistent with the historic data, more years are needed to characterize low flow variability compared to high flow variability for these sites.

A cluster analysis is performed to obtain clusters based on the sample sizes. The resulting maps show clusters that do not seem to match any particular physiographic segregation. This is consistent with results from a regression of the sample size on some readily available physiographic variables.

Table 4.1: Descriptive information for the stations of interest

No	ID	State	RY*	Begin	End	25p†	75p‡	E§	DA¶
1	01632000	VA	71	Apr-25	Mar-95	15	178	1051.80	210.0
2	01634500	VA	58	Oct-37	Jul-95	17	97	647.09	103.0
3	01638480	VA	25	Oct-70	Jul-95	18	100	249.15	89.6
4	01643700	VA	29	Oct-65	Jul-95	22	158	329.80	123.0
5	01644000	VA	66	Apr-30	Jul-95	54	359	248.93	332.0
6	01646000	VA	61	Apr-35	Oct-94	22	61	151.30	57.9
7	01663500	VA	52	Aug-42	Oct-92	97	390	288.30	287.0
8	01665500	VA	53	Oct-42	Dec-94	45	178	439.44	114.0
9	01671100	VA	34	Oct-61	Oct-94	19	104	132.30	107.0
10	02015700	VA	35	Aug-60	Sep-94	45	158	1610.14	110.0
11	02017500	VA	69	Oct-26	Jul-95	21	146	1254.30	104.0
12	02020500	VA	57	Oct-38	Oct-94	18	170	1384.84	144.0
13	02027800	VA	36	Aug-60	Oct-94	62	191	444.39	147.0
14	02030500	VA	69	Apr-26	Sep-95	71	215	238.78	226.0
15	02041000	VA	49	Oct-46	Aug-95	36	145	177.20	158.0
16	02042500	VA	54	Apr-42	Aug-95	61	346	6.07	252.0
17	02044500	VA	44	Oct-50	Aug-95	90	324	184.88	309.0
18	02046000	VA	49	Oct-46	Aug-95	17	115	129.94	112.0
19	02052500	VA	42	Oct-53	Aug-95	6.8	64	152.59	65.2
20	02053800	VA	35	Oct-60	Jul-95	41	127	1361.87	110.0
21	02061500	VA	59	Apr-37	Jul-95	130	366	544.02	320.0
22	02065500	VA	49	Oct-46	Aug-95	40	103	370.19	98.0
23	02069700	VA	33	Oct-62	Jul-95	70	148	871.60	84.6
24	02070000	VA	66	Oct-28	Aug-95	70	136	730.94	108.0
25	02082770	NC	32	Aug-63	Sep-94	45	165	130.00	166.0

*Record Years
†25th Percentile (low flow threshold)
‡75th Percentile (high flow threshold)
§Elevation (ft)
¶Drainage Area (mi^2)

Table 4.1: *continued*

No	ID	State	RY*	Begin	End	25p†	75p‡	E§	DA¶
26	02082950	NC	34	Oct-59	Sep-94	32	166	116.44	177.0
27	02083800	NC	38	Dec-56	Sep-94	12	88	30.00	78.1
28	02088470	NC	27	Aug-64	Oct-90	34	220	129.24	191.0
29	02091700	NC	31	Oct-56	Sep-87	10	122	30.00	93.3
30	02092000	NC	39	Feb-50	Sep-88	24	224	-2.07	182.0
31	02092500	NC	42	Jan-51	Sep-94	24	233	19.15	168.0
32	02106000	NC	42	Feb-50	Sep-91	28	152	80.52	92.8
33	02112120	NC	29	Apr-64	Sep-94	105	209	964.85	128.0
34	02112360	NC	28	Apr-64	Sep-94	76	143	927.12	78.8
35	02113850	NC	28	Apr-64	Sep-94	168	344	880.97	231.0
36	02118500	NC	42	Jan-51	Sep-94	104	226	734.78	155.0
37	02129590	SC	16	Oct-79	Sep-94	8.8	36	100.00	26.4
38	02131150	SC	26	Nov-66	Sep-92	5.1	32	75.00	27.4
39	02131309	SC	19	Aug-76	Sep-94	4.3	33	302.68	24.3
40	02135300	SC	27	Jul-68	Sep-94	36	136	164.53	96.0
41	02143000	NC	58	Aug-25	Sep-94	62	140	890.99	83.2
42	02143040	NC	30	Oct-61	Sep-94	22	50	1103.00	25.7
43	02149000	NC	43	Jan-51	Sep-94	76	153	815.40	79.0
44	02152100	NC	35	Mar-59	Sep-94	47	97	890.00	60.5
45	02153780	SC	14	Oct-80	Sep-94	7.2	22	565.00	24.1
46	02157000	SC	38	Oct-50	Sep-88	32	67	680.00	44.4
47	02175500	SC	43	Feb-51	Sep-94	157	431	64.35	341.0
48	02176500	SC	43	Feb-51	Sep-94	17	239	50.30	203.0
49	02197300	SC	29	Jun-66	Sep-94	88	114	165.00	98.7
50	02197400	SC	20	Mar-74	Sep-94	45	105	117.00	59.3

*Record Years
†25th Percentile (low flow threshold)
‡75th Percentile (high flow threshold)
§Elevation (ft)
¶Drainage Area (mi^2)

Table 4.2: Series suggesting higher order model and/or linear trend

Station	Target variables suggesting higher order model
01644000	NH : ARMA(1,1)
01646000	NH : AR(2)
01671100	DL : AR(5)
02015700	DL : AR(2)
02020500	NL : AR(4)
02030500	NH : AR(2)
02069700	DH : AR(3)
02070000	NL : ARMA(4,1), NH : ARMA(1,1)
02113850	NL : AR(5)
02143000	NL : AR(2), DL : AR(4)
02197400	DL : MA(3)
Station	**Target variables suggesting linear trend (slope; p-value)**
01634500	DH (0.075; 0.0223)
01638480	NL (0.277; 0.0043)
01643700	DL (0.910; 0.0016)
01644000	DH (0.056; 0.0186)
01646000	NH (0.127; 0.0040)
02017500	NH (-0.039; 0.0191), DH (0.065; 0.0031)
02041000	DL (0.254; 0.0016)
02046000	DL (0.186; 0.0445)
02053800	NL (-0.172; 0.0177)
02061500	NH (-0.081; 0.0379)
02069700	DH (0.183; 0.0270)
02091700	DL (0.400; 0.0107), NH (-0.173; 0.0403)
02092000	NL (-0.085; 0.0441)
02113850	DL (0.231; 0.0072)
02153780	DL (-0.474; 0.0434)

Table 4.3: Estimates of the lag 1 autoregressive coefficient ρ for each station and target variable

Station	NL	DL	NH	DH	Station	NL	DL	NH	DH
01632000	0.17	0.21	-0.22	-0.08	02082950	0.08	0.21	-0.15	-0.08
01634500	0.04	0.21	0.03	0.03	02083800	-0.02	0.08	0.11	-0.11
01638480	-0.03	0.10	0.16	0.13	02088470	-0.07	-0.39	0.01	-0.22
01643700	-0.19	0.01	0.06	-0.05	02091700	0.26	0.51	0.19	-0.16
01644000	0.16	0.02	0.15	-0.02	02092000	0.08	0.41	0.11	0.16
01646000	0.22	-0.01	0.40	0.09	02092500	0.01	0.14	0.19	-0.18
01663500	0.00	0.22	0.13	0.15	02106000	0.04	0.07	0.37	-0.12
01665500	0.01	0.07	0.05	0.12	02112120	0.27	0.24	0.17	0.23
01671100	0.26	0.16	0.08	0.19	02112360	0.33	0.16	0.22	0.30
02015700	0.23	0.71	0.16	0.26	02113850	0.11	-0.11	0.14	0.24
02017500	0.16	0.14	-0.28	0.33	02118500	0.25	0.24	0.10	0.17
02020500	-0.04	0.11	-0.10	-0.04	02129590	0.12	-0.36	-0.67	-0.19
02027800	0.29	0.30	-0.08	-0.01	02131150	0.07	-0.10	-0.40	0.08
02030500	0.32	0.21	0.10	0.23	02131309	-0.03	-0.38	0.59	-0.29
02041000	0.19	0.04	0.31	0.05	02135300	0.18	0.07	-0.32	-0.31
02042500	0.11	0.16	0.19	-0.01	02143000	0.41	0.41	0.10	0.01
02044500	0.21	-0.04	0.08	-0.06	02143040	-0.08	0.09	-0.02	-0.14
02046000	-0.10	0.04	0.05	-0.09	02149000	-0.10	0.16	0.08	0.06
02052500	-0.06	0.09	-0.10	-0.06	02152100	0.20	0.06	-0.14	-0.08
02053800	-0.05	0.09	0.09	0.17	02153780	0.12	0.41	-0.58	-0.25
02061500	0.18	0.18	-0.03	-0.04	02157000	0.28	0.17	0.37	0.21
02065500	0.50	0.45	0.06	0.21	02175500	0.30	0.12	0.14	0.08
02069700	0.25	0.35	0.24	0.24	02176500	0.18	0.07	-0.18	-0.05
02070000	0.33	0.14	-0.02	-0.03	02197300	0.80	0.58	0.12	0.39
02082770	-0.06	-0.03	0.04	0.15	02197400	0.11	0.11	0.30	0.08

Table 4.4: Sample sizes to detect a level shift of size $\delta = k\sigma$ for various values of β and ρ (Method I)

k	0.5	1.0	1.5	2.0	2.5
$\beta = 0.1,\ \rho = 0.1$	104	26	12	7	4
0.2	131	33	14	8	5
0.3	171	42	19	10	6
0.4	233	57	25	13	8
0.5	335	83	36	19	12
0.6	523	129	56	30	18
0.7	930	230	100	54	33
0.8	2095	519	227	125	77
0.9	8392	2088	920	511	322
$\beta = 0.2,\ \rho = 0.1$	78	20	9	5	3
0.2	98	24	11	6	4
0.3	128	32	14	7	4
0.4	173	43	18	10	6
0.5	250	61	26	14	8
0.6	390	96	41	22	13
0.7	694	171	74	40	23
0.8	1564	386	168	91	56
0.9	6265	1556	684	378	237
$\beta = 0.3,\ \rho = 0.1$	61	15	7	4	3
0.2	77	19	8	5	3
0.3	100	25	11	6	3
0.4	136	33	14	7	4
0.5	196	48	20	10	5
0.6	306	75	32	16	9
0.7	545	133	57	30	17
0.8	1228	302	131	70	42
0.9	4924	1221	535	294	183

Table 4.5: Theoretical and empirical powers for some values of ρ, n and $k = 1$ using Method II

ρ	n	power1*	power2†	power3‡	ρ	n	power1	power2	power3
-0.7	10	0.92890	0.9294	0.9207	0.1	10	0.49149	0.4913	0.5086
-0.7	20	0.99920	0.9989	0.9985	0.1	20	0.79904	0.8001	0.8028
-0.7	30	0.99999	1.0000	0.9999	0.1	30	0.93160	0.9294	0.9237
-0.5	10	0.86108	0.8603	0.8528	0.3	10	0.35953	0.3608	0.3842
-0.5	20	0.99491	0.9959	0.9920	0.3	20	0.61212	0.6125	0.6399
-0.5	30	0.99988	0.9999	0.9995	0.3	30	0.78185	0.7845	0.7947
-0.3	10	0.76038	0.7583	0.7570	0.5	10	0.25418	0.2512	0.2798
-0.3	20	0.97648	0.9764	0.9698	0.5	20	0.40649	0.4055	0.4539
-0.3	30	0.99837	0.9980	0.9963	0.5	30	0.53866	0.5445	0.5905
-0.1	10	0.63201	0.6275	0.6382	0.7	10	0.18565	0.1849	0.2055
-0.1	20	0.92038	0.9157	0.9128	0.7	20	0.24663	0.2494	0.2924
-0.1	30	0.98648	0.9872	0.9798	0.7	30	0.30317	0.3048	0.3700
0.0	10	0.56201	0.5612	0.5738	0.9	10	0.15690	0.1560	0.1656
0.0	20	0.86895	0.8700	0.8651	0.9	20	0.16923	0.1737	0.1933
0.0	30	0.96771	0.9673	0.9589	0.9	30	0.17738	0.1804	0.2166

*Theoretical power

†Empirical power using ρ

‡Empirical power using $\hat{\rho}$ generated from ρ

Table 4.6: Sample sizes for detecting the level shift ($k = 1$, $\beta = .2$, Method I)

Station*	n_1†	n_2‡	n_3§	n_4¶	Station	n_1	n_2	n_3	n_4
1632000	23	25	12	14	2082950	19	25	13	14
1634500	18	25	17	17	2083800	16	19	20	14
1638480	15	19	22	21	2088470	15	10	16	12
1643700	12	17	18	15	2091700	28	64	24	13
1644000	22	17	22	16	2092000	19	44	20	22
1646000	26	16	43	19	2092500	16	21	24	12
1663500	16	26	21	22	2106000	17	18	39	13
1665500	16	18	18	20	2112120	29	27	23	26
1671100	29	22	19	24	2112360	34	22	26	31
2015700	26	185	22	28	2113850	20	13	21	27
2017500	22	21	11	34	2118500	27	27	19	23
2020500	15	20	14	15	2129590	21	10	9	12
2027800	31	31	14	16	2131150	18	14	10	19
2030500	33	25	19	27	2131309	15	10	93	11
2041000	24	17	33	18	2135300	23	19	11	11
2042500	20	22	24	16	2143000	44	43	19	16
2044500	25	15	19	15	2143040	14	19	16	13
2046000	14	17	18	14	2149000	14	22	19	18
2052500	15	19	14	15	2152100	24	18	13	14
2053800	15	19	19	23	2153780	20	44	9	11
2061500	23	23	15	15	2157000	30	23	39	25
2065500	60	50	18	25	2175500	31	20	21	19
2069700	27	37	27	27	2176500	24	18	12	15
2070000	35	21	16	15	2197300	368	87	21	41
2082770	15	15	17	22	2197400	20	20	31	19

*If $\rho = 0$, 16 is the required sample size.
†Series 1 : number of low pulses(NL)
‡Series 2 : duration of low pulses(DL)
§Series 3 : number of high pulses(NH)
¶Series 4 : duration of high pulses(DH)

Table 4.7: Sample sizes for detecting the level shift ($k = 1$, $\beta = .2$, Method II)

Station*	n_1 †	n_2 ‡	n_3 §	n_4 ¶	Station	n_1	n_2	n_3	n_4
01632000	24	26	13	15	02082950	20	26	14	15
01634500	19	26	18	18	02083800	17	20	21	14
01638480	16	20	23	22	02088470	15	10	17	12
01643700	13	18	19	16	02091700	29	62	25	14
01644000	23	18	23	17	02092000	20	43	21	23
01646000	27	17	43	20	02092500	17	22	24	13
01663500	17	26	22	23	02106000	18	19	39	14
01665500	17	19	19	21	02112120	30	28	23	27
01671100	29	23	20	25	02112360	34	23	26	32
02015700	27	171	23	29	02113850	21	14	22	28
02017500	23	22	12	35	02118500	28	28	20	24
02020500	16	21	15	16	02129590	21	11	8	13
02027800	31	32	15	17	02131150	19	15	10	20
02030500	33	26	20	27	02131309	16	10	89	11
02041000	25	18	33	19	02135300	24	20	11	11
02042500	21	23	25	17	02143000	43	43	20	17
02044500	26	16	20	16	02143040	15	20	17	14
02046000	15	18	19	15	02149000	15	23	20	19
02052500	16	20	15	16	02152100	25	19	14	15
02053800	16	20	20	24	02153780	21	44	9	12
02061500	24	24	16	16	02157000	30	24	39	26
02065500	59	49	19	26	02175500	32	21	22	20
02069700	28	37	28	27	02176500	24	19	13	16
02070000	35	22	17	16	02197300	334	82	22	41
02082770	16	16	18	23	02197400	21	21	32	20

*If $\rho = 0$, 17 is the required sample size.
†Series 1 : number of low pulses(NL)
‡Series 2 : duration of low pulses(DL)
§Series 3 : number of high pulses(NH)
¶Series 4 : duration of high pulses(DH)

Table 4.8: Sample sizes for detecting the level shift ($\delta = 3$, $\beta = .2$, Method III)

Station	n_1*	n_2†	n_3‡	n_4§	Station	n_1	n_2	n_3	n_4
1632000	12	520	14	18	2082950	32	32	32	12
1634500	32	58	26	28	2083800	18	240	34	22
1638480	22	56	22	22	2088470	18	42	12	22
1643700	24	144	32	24	2091700	12	86	32	14
1644000	32	154	38	20	2092000	14	204	18	26
1646000	66	46	82	16	2092500	10	120	14	24
1663500	24	128	32	24	2106000	14	48	18	72
1665500	20	214	20	24	2112120	82	30	52	16
1671100	20	180	22	36	2112360	68	22	48	28
2015700	16	64	26	18	2113850	74	16	72	18
2017500	24	154	10	32	2118500	88	64	48	14
2020500	10	362	16	14	2129590	12	32	26	42
2027800	32	60	32	70	2131150	26	60	24	38
2030500	48	58	52	22	2131309	22	54	32	28
2041000	30	98	34	42	2135300	12	168	20	40
2042500	10	142	12	82	2143000	270	36	26	18
2044500	34	130	32	40	2143040	38	22	26	12
2046000	18	134	36	46	2149000	80	70	36	24
2052500	14	130	28	16	2152100	76	18	30	16
2053800	28	30	16	18	2153780	22	22	14	68
2061500	58	112	40	24	2157000	50	32	44	10
2065500	60	44	42	28	2175500	16	72	22	116
2069700	62	26	40	28	2176500	18	258	18	56
2070000	90	36	46	128	2197300	230	30	62	14
2082770	30	18	30	12	2197400	28	92	80	26

*Series 1 : number of low pulses(NL)
†Series 2 : duration of low pulses(DL)
‡Series 3 : number of high pulses(NH)
§Series 4 : duration of high pulses(DH)

Table 4.9: Sample sizes for detecting the level shift ($k = 1$, $\beta = .2$, Method I using t distribution)

Station*	n_1†	n_2‡	n_3§	n_4¶	Station	n_1	n_2	n_3	n_4
1632000	24	26	13	15	2082950	20	26	14	15
1634500	19	26	18	18	2083800	17	20	21	15
1638480	16	20	23	22	2088470	16	11	17	13
1643700	13	18	19	16	2091700	29	65	25	14
1644000	23	18	23	17	2092000	20	45	21	23
1646000	27	17	44	20	2092500	17	22	25	13
1663500	17	27	22	23	2106000	18	19	40	14
1665500	17	19	19	21	2112120	30	28	24	27
1671100	30	23	20	25	2112360	35	23	27	32
2015700	27	186	23	29	2113850	21	14	22	28
2017500	23	22	12	35	2118500	28	28	20	24
2020500	16	21	15	16	2129590	22	11	10	13
2027800	32	32	15	17	2131150	19	15	11	20
2030500	34	26	20	28	2131309	16	11	94	12
2041000	25	18	34	19	2135300	24	20	11	12
2042500	21	23	25	17	2143000	45	44	20	17
2044500	26	16	20	16	2143040	15	20	17	14
2046000	15	18	19	15	2149000	15	23	20	19
2052500	16	20	15	16	2152100	25	19	14	15
2053800	16	20	20	24	2153780	21	45	10	12
2061500	24	24	16	16	2157000	31	24	40	26
2065500	61	51	19	26	2175500	32	21	23	20
2069700	29	38	28	28	2176500	25	19	13	16
2070000	36	22	17	16	2197300	369	88	22	42
2082770	16	16	18	23	2197400	21	21	32	20

*If $\rho = 0$, 17 is the required sample size.
†Series 1 : number of low pulses(NL)
‡Series 2 : duration of low pulses(DL)
§Series 3 : number of high pulses(NH)
¶Series 4 : duration of high pulses(DH)

Table 4.10: Correlation matrix for the 4 variables n_1, n_2, n_3 and n_4 in Table 4.6

	n_1	n_2	n_3	n_4
n_1	1.00000*	0.36155	-0.00987	0.53652
	(0.0000)†	(0.0099)	(0.9458)	(0.0001)
n_2	0.36155	1.00000	-0.04960	0.37902
	(0.0099)	(0.0000)	(0.7323)	(0.0066)
n_3	-0.00987	-0.04960	1.00000	-0.01729
	(0.9458)	(0.7323)	(0.0000)	(0.9051)
n_4	0.53652	0.37902	-0.01729	1.00000
	(0.0001)	(0.0066)	(0.9051)	(0.0000)

*Pearson correlation coefficient

†p-value for testing if the correlation coefficient is 0

Table 4.11: Correlation matrix for the 4 variables n_1, n_2, n_3 and n_4 in Table 4.6 (outlying stations excluded)

	n_1	n_2	n_3	n_4
n_1	1.00000 *	0.51043	0.10561	0.32329
	(0.0000)†	(0.0002)	(0.4799)	(0.0267)
n_2	0.51043	1.00000	0.00585	0.07863
	(0.0002)	(0.0000)	(0.9688)	(0.5993)
n_3	0.10561	0.00585	1.00000	0.23648
	(0.4799)	(0.9688)	(0.0000)	(0.1095)
n_4	0.32329	0.07863	0.23648	1.00000
	(0.0267)	(0.5993)	(0.1095)	(0.0000)

*Pearson correlation coefficient

†p-value for testing if the correlation coefficient is 0

Table 4.12: Principal component vectors for the 4 variables n_1, n_2, n_3 and n_4 in Table 4.6 (outlying stations excluded)

Station	P1*	P2†	P3‡	P4§	Station	P1	P2	P3	P4
1632000	38.14	-2.06	5.18	2.92	2082950	35.78	-0.55	6.07	5.82
1634500	37.55	3.87	6.77	7.5	2083800	32.41	7.92	9.01	4.31
1638480	35.49	12.85	5.69	7.34	2088470	24.93	8.85	4.92	-1.02
1643700	28.64	8.81	6.6	6.35	2091700	65.15	-14.53	25.86	22.37
1644000	36.7	10.04	8.04	-0.61	2092000	51.54	-1.28	11.33	19.86
1646000	45.67	25.79	19.49	-3.31	2092500	33.62	8.77	13.66	4.81
1663500	40.12	8.95	6.5	11.17	2106000	37.05	20.99	21.99	2.47
1665500	34.1	9.88	3.33	5.71	2112120	51.54	9.1	4.74	3.65
1671100	46.87	7.84	1.85	0.01	2112360	55.2	15.03	1.53	-1.36
2017500	44.24	8.94	-10.15	7.88	2113850	38.02	16.96	-1.18	2.09
2020500	31.18	3.92	4.71	6.02	2118500	47.8	5.29	4.2	4.15
2027800	48.06	-4.32	6.62	1.38	2129590	27.02	2.74	-0.08	-5.31
2030500	52.49	6.97	0.55	-0.08	2131150	30.69	5.4	-2.7	1.59
2041000	41.82	18.25	13.92	-1.5	2135300	33.26	-1.21	4.65	-1.68
2042500	38.62	9.36	11.06	3.84	2143000	64.36	-9.43	13.02	-0.86
2044500	36.34	7.85	5.92	-4.29	2143040	29.59	5.06	7.13	5.46
2046000	29.5	7.93	7.16	4.57	2149000	34.29	8.05	6.71	8.9
2052500	30.64	4.42	4.39	5.42	2152100	35.28	1.9	3.59	-2.03
2053800	35.62	11.71	2.34	8.03	2153780	44.26	-14.39	11.35	15.59
2061500	38.3	1.5	5.88	2.02	2157000	53.78	21.57	14.77	0.03
2065500	82.49	-12.59	7.62	-5.25	2175500	45.33	7.44	5.81	-4.32
2069700	57.14	7.82	9.99	11.43	2176500	35.47	1.68	2.25	-1.69
2070000	45.35	0.74	5.26	-7.9	2197400	40.75	16.66	13.08	3.56
2082770	32.48	11.82	0.42	5.3					

*The first principal component
†The second principal component
‡The third principal component
§The fourth principal component

Table 4.13: Clustering of the stations by the descriptive statistics (complete linkage)

Cluster	1	2	3	4	5	6
Station	02030500	01634500	01665500	01632000	02027800	01646000
	02065500	01638480	01671100	02020500	02042500	02113850
	02069700	01643700	02017500		02088470	02143000
	02070000	01644000	02083800		02092500	02197300
	02082950	01663500	02092000		02106000	
	02112120	02015700	02135300		02129590	
	02112360	02041000	02176500		02131150	
	02118500	02044500			02175500	
	02143040	02046000				
	02149000	02052500				
	02152100	02053800				
	02153780	02061500				
	02157000	02082770				
		02091700				
		02131309				
		02197400				

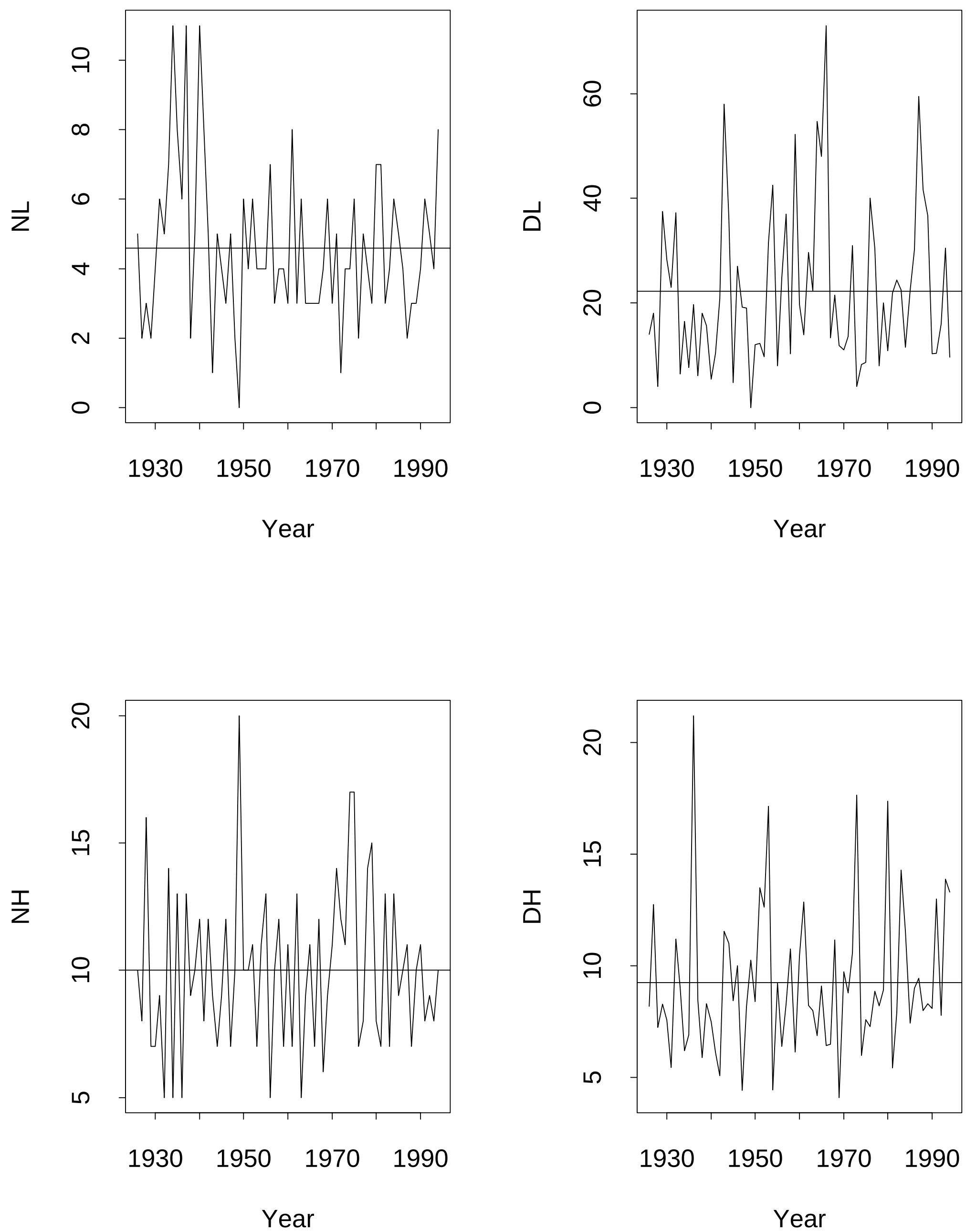

Figure 4.1: Time series data for station 01632000

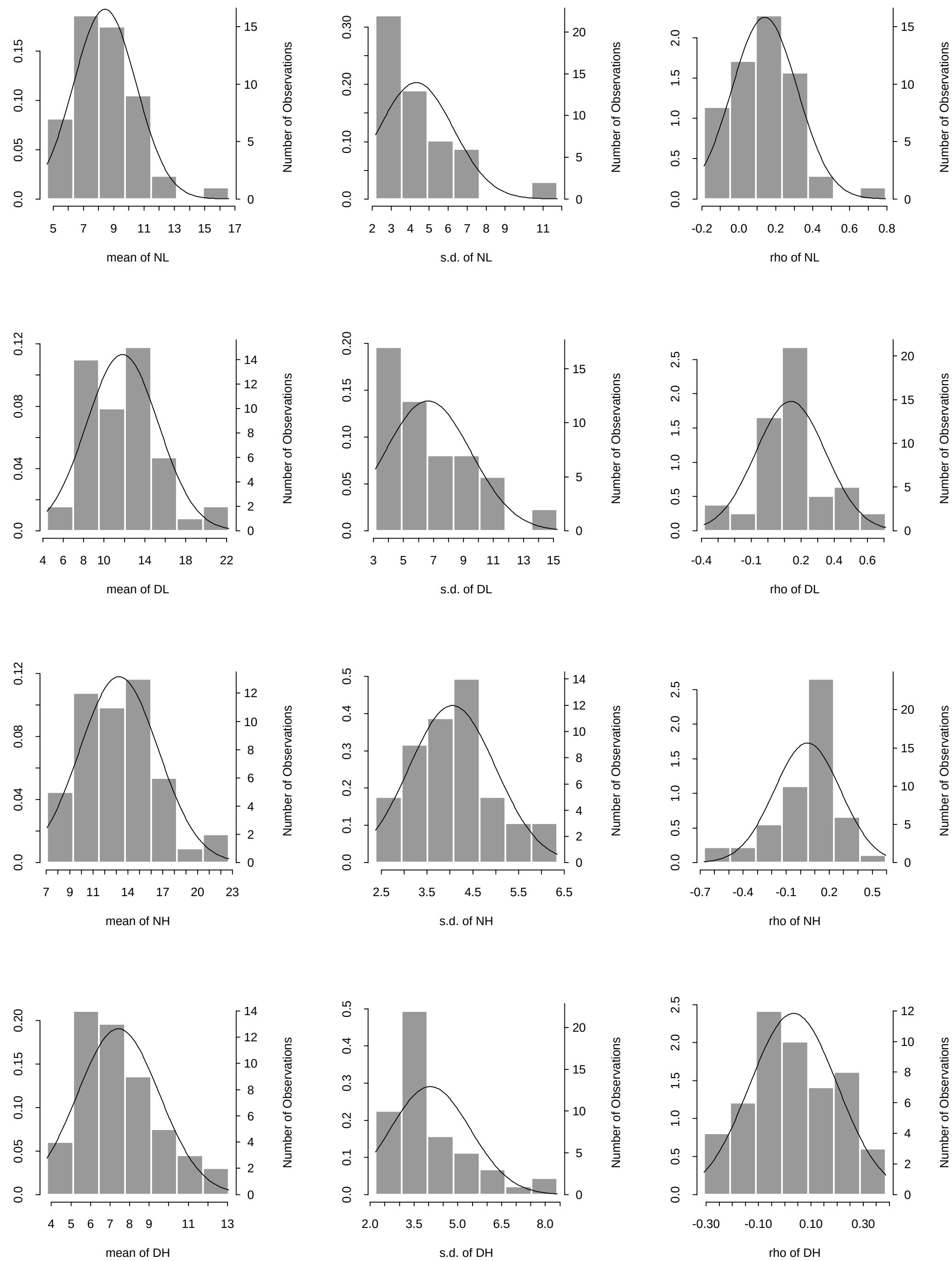

Figure 4.2: Distribution of the descriptive statistics across the stations

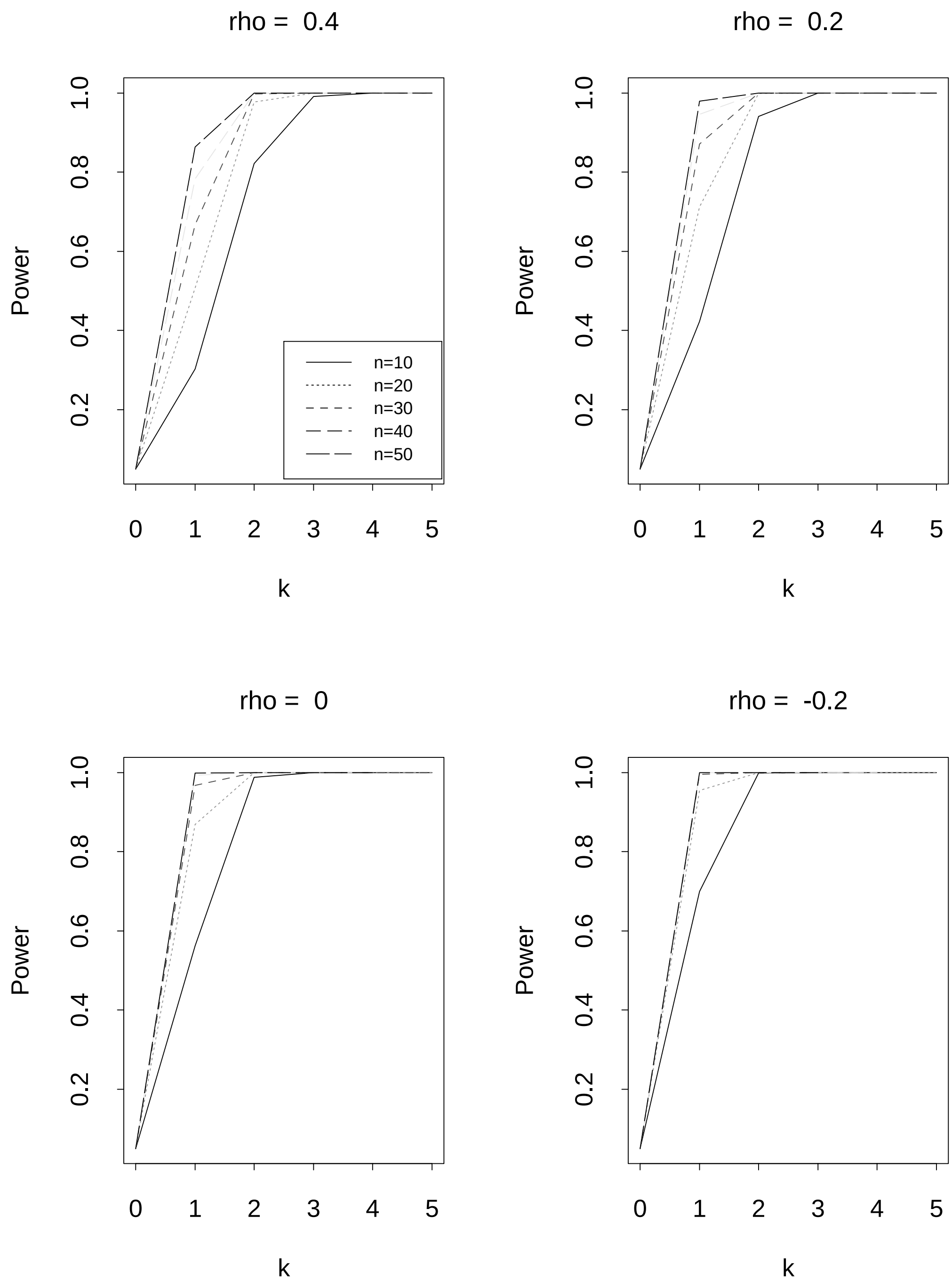

Figure 4.3: Theoretical powers for some values of ρ (Method II)

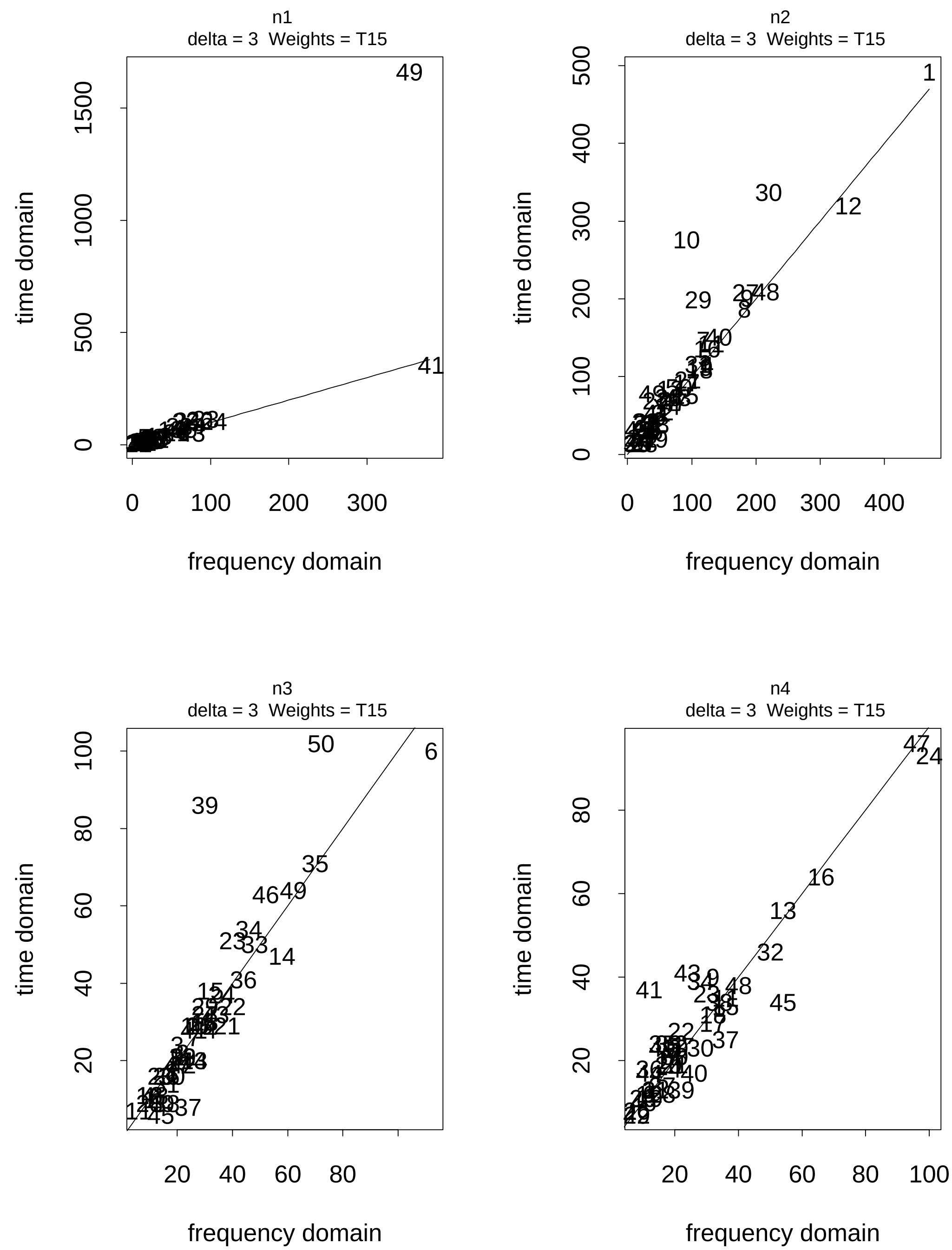

Figure 4.4: Comparison of sample sizes between TDM and FDM (triangular weights of length 15; serial numbers in Table 4.1 used for sites)

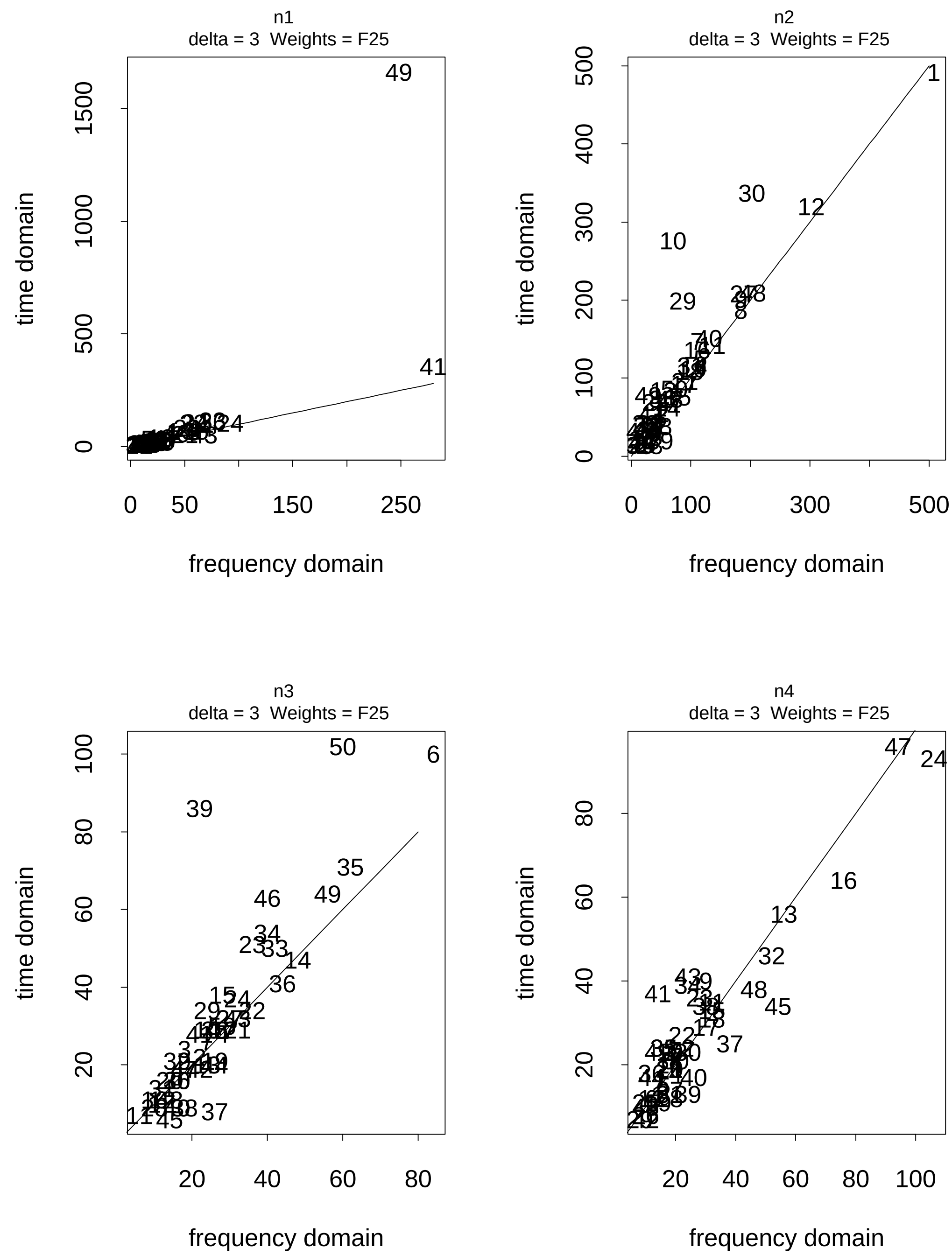

Figure 4.5: Comparison of sample sizes between TDM and FDM (flat weights of length 25; serial numbers in Table 4.1 used for sites)

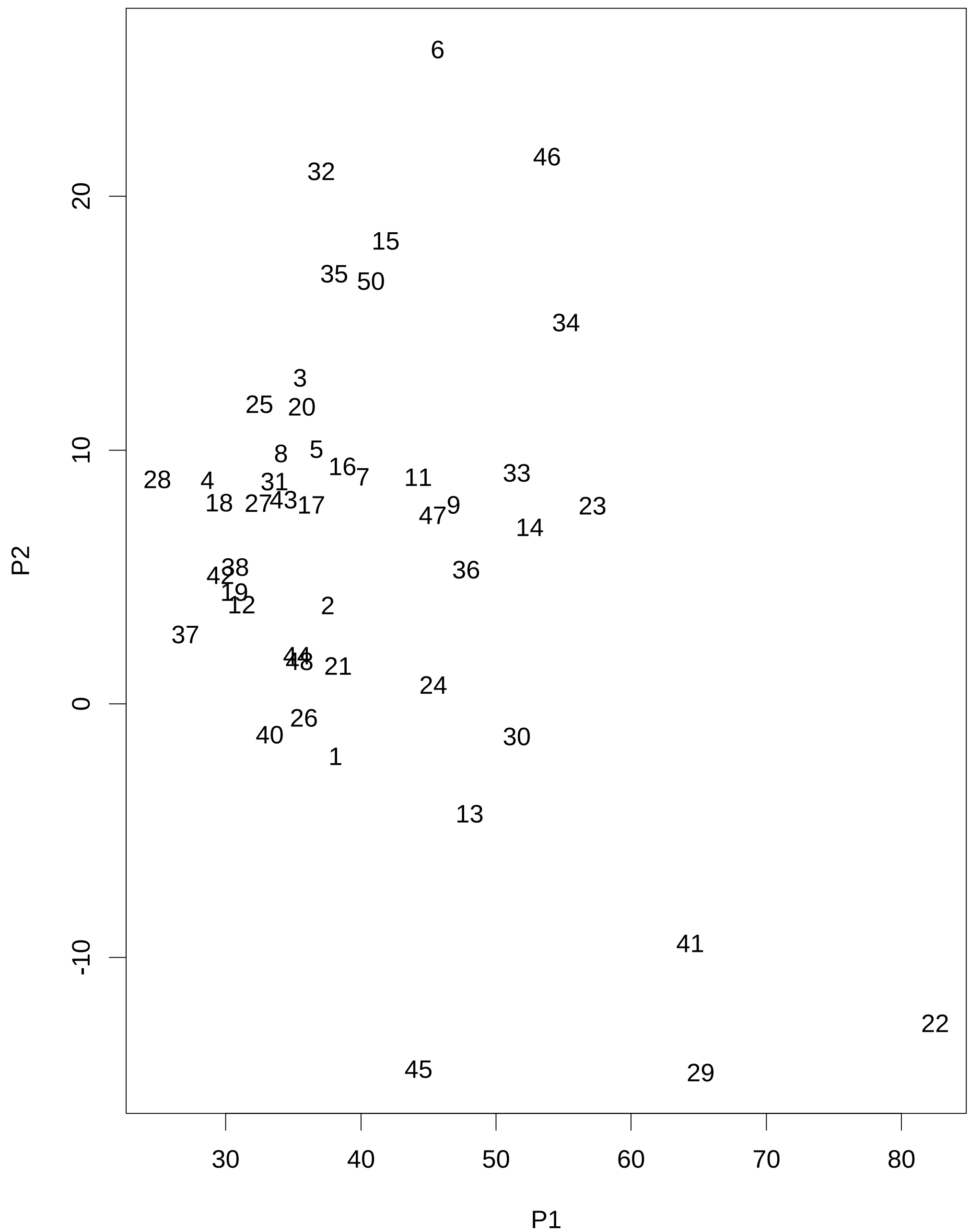

Figure 4.6: The first two principal components for the 4 variables n_1, n_2, n_3 and n_4 in Table 4.6 (outlying stations excluded)

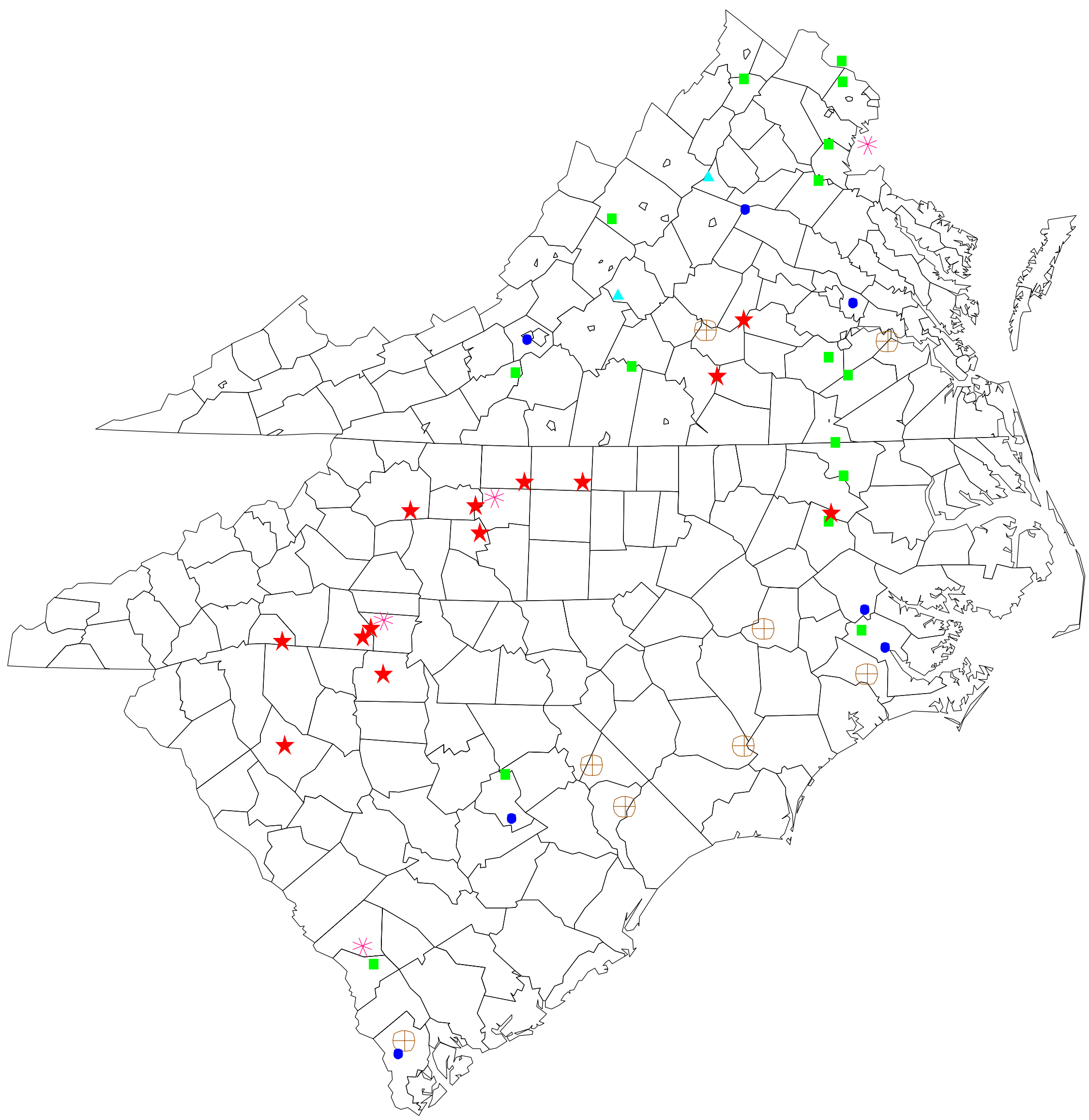

Figure 4.7: Clustering of the stations by the descriptive statistics (complete linkage)

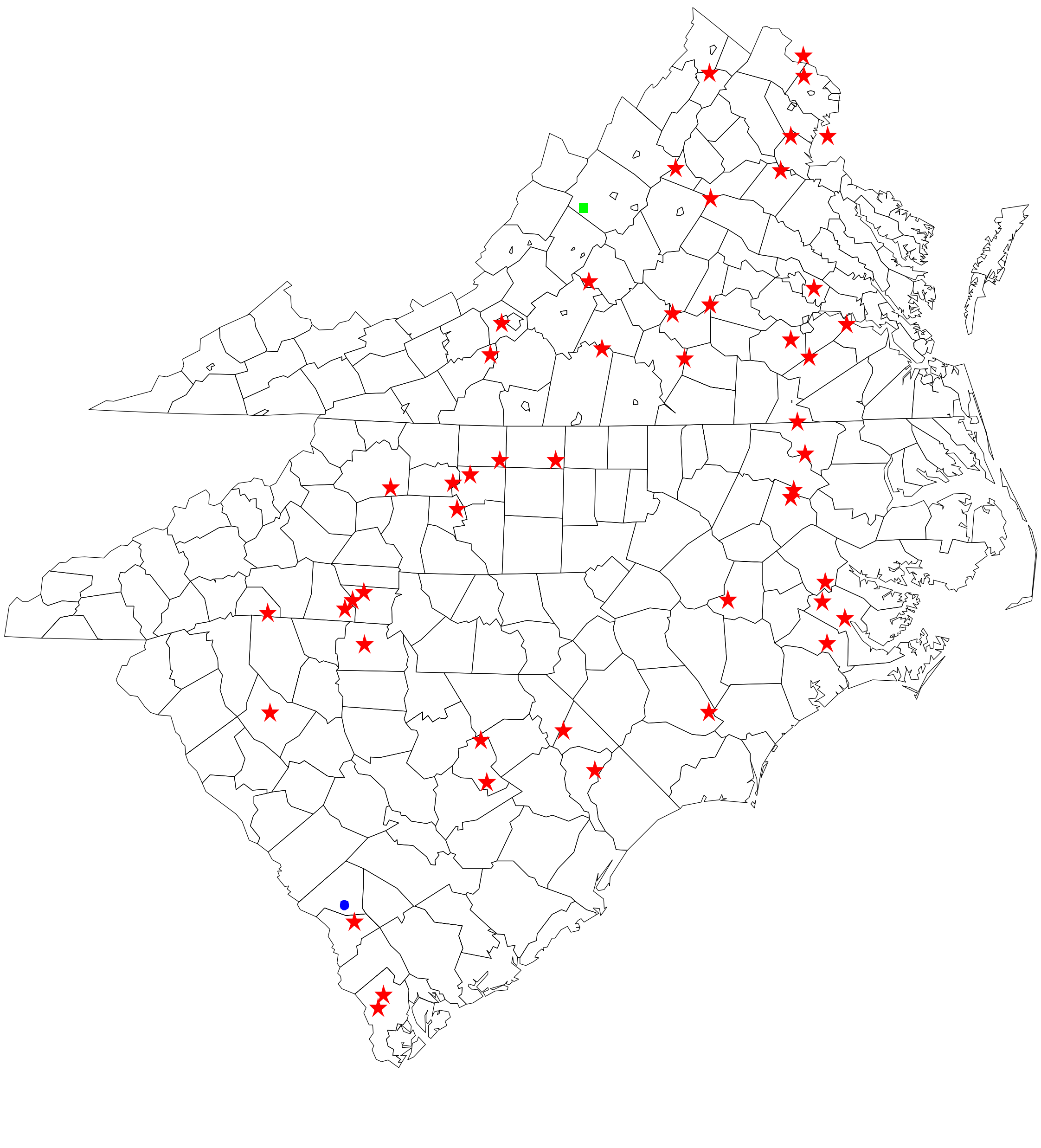

Figure 4.8: Clustering of the stations by the sample sizes (single linkage; maximizing CCC)

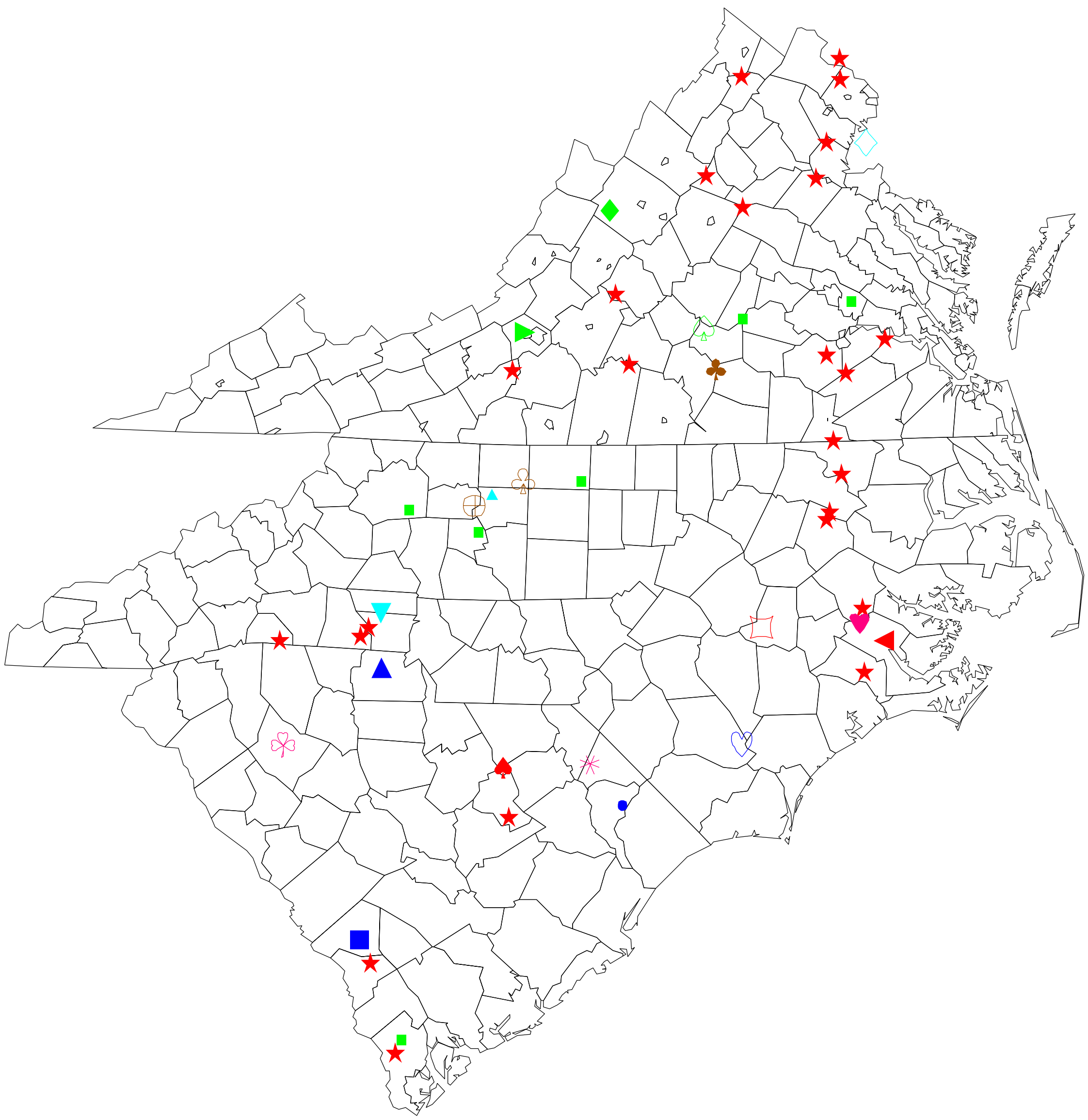

Figure 4.9: Clustering of the stations by the sample sizes (single linkage; The largest cluster contains no more than 25 stations.)

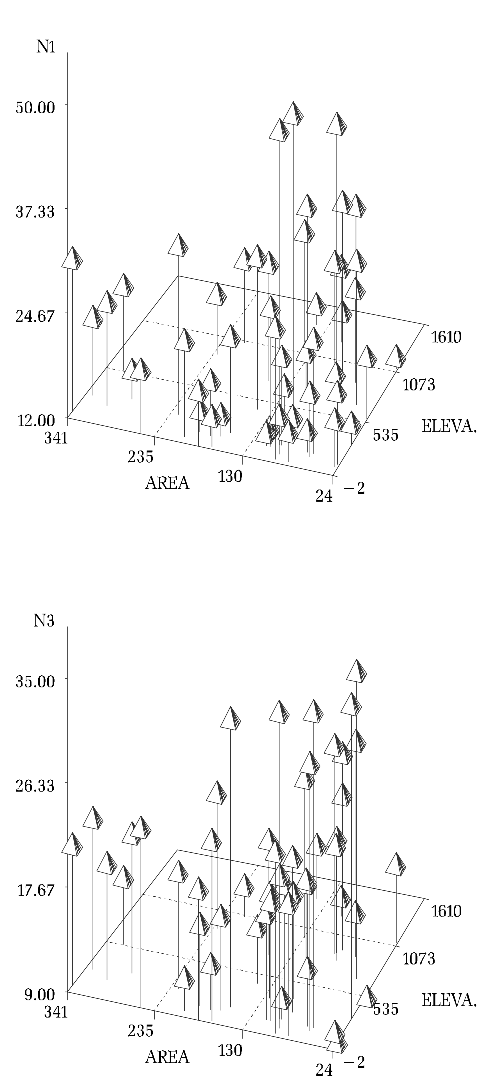

Figure 4.10: Scatter plots of elevation, drainage area and sample size from Method I ($k = 1$)

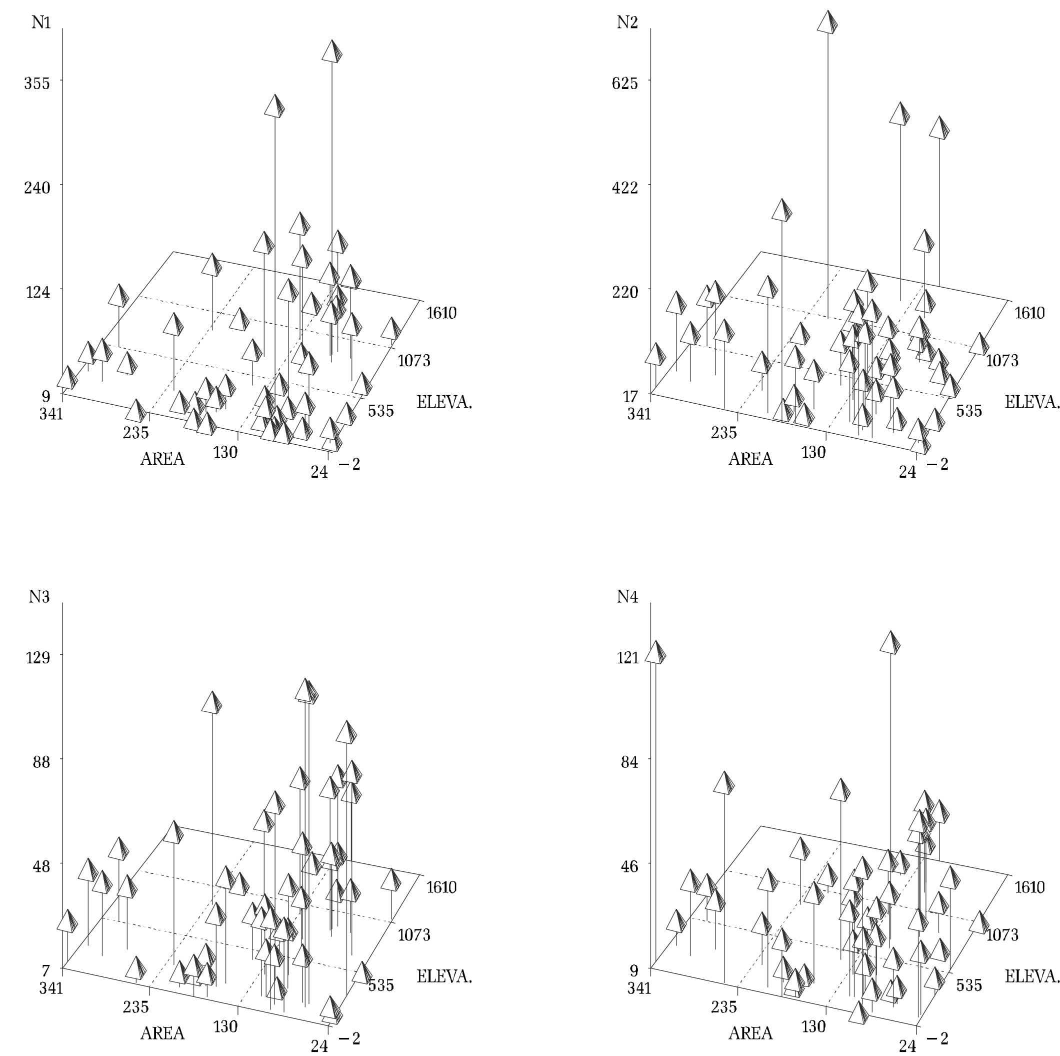

Figure 4.11: Scatter plots of elevation, drainage area and sample size from Method I ($\delta = 3$)

Chapter 5

Conclusion

This dissertation consists of 3 essays involving time series with trend breaks.

In Chapter 2, we examine the 'converse Perron phenomenon' which refers to the spurious rejection of the unit root null hypothesis when the data are generated from a $I(1)$ process with an early break. Empirical sizes of unit root tests are compared between the ordinary least squares estimator and the symmetric estimators. The symmetric estimators are found to have much less severe size distortion problems.

The spurious rejection problem is verified to disapper asymptotically at the same rate for all the estimators. To gain some clues with respect to the empirical results, the expected values of the quadratic forms constituting the test statistics are analyzed.

Augmented tests with lagged first differences are also considered to extend the procedures to more general processes. The tests in the higher order models behave quite similarly to those in the lag one models. Next our test statistics are modified to accomodate a level shift. It is seen that the weighted symmetric estimator never gives less power than the ordinary least squares estimator and sometimes shows power gains that are substantial.

In Chapter 3, a new procedure for testing the unit root null hypothesis considering a possible trend-break is introduced. The idea here is to divide the data in half and take the minimum of the resulting two unit root test statistics. Avoiding the necessity of searching for the break, our method is simpler than other methods in the literature. Although the method might seem inefficient in its use of the data, it shows reasonable power compared to other methods for testing for unit roots in the presence of trend breaks.

Simulation results for the empirical size and power are given. Empirical results obtained for the Nelson-Plosser (1982) data set are displayed along with those of other researchers.

In Chapter 4, we analyze temporal characteristics of hydrologic variables in stream flow data from the US Geological Survey. Based on USGS tradition, the target variables of our research are the counts and durations of high and low flows. We model the autocorrelation structure and check for stationarity and trends for each station.

Computing the required sample size to detect a certain level shift is discussed using 3 different methods. The consistency of the sample size results is checked between a model based approach and a nonparametric approach. A measure of accuracy for the sample size estimates is also examined. The cluster pattern based on the sample sizes does not correspond to any particular geographical features. A regression analysis is performed, failing to find any influence of physiographic variables on the sample sizes.

A possible direction for future research is to perform similar analyses on different hydrologic variables. There might be other hydrologic variables which are more appropriately represented as functions of geographic features. Another issue worth pursuing is the use of spatial analysis techniques on this data set.

References

Amsler, C. and Lee, J. (1995), An LM test for a unit root in the presence of a structural change. Econometric Theory 11, 359-368.

Banerjee, A., Lumsdaine, R. L. and Stock, J. H. (1992), Recursive and sequential tests of the unit root and trend-break hypotheses: Theory and international evidence. Journal of Business & Economic Statistics 10, 271-287.

Box, G. E. P., Jenkins, G. M. and Reinsel, G. C. (1994), *Time Series Analysis : Forecasting and Control*. Prentice Hall, Englewood Cilffs, NJ.

Brocklebank, J. C. and Dickey, D. A. (1986), *SAS System for Forecasting Time Series*. SAS Institute Inc., Cary, NC.

Christiano, L. J. (1992), Searching for a break in GNP. Journal of Business & Economic Statistics 10, 237-250.

Clausen, B. and Biggs, B. J. F. (1997), Relationships between benthic biota and hydrological indices in New Zealand streams. Freshwater Biology 38, 327-342.

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.

Dinardo, J., Johnston, J. and Johnston, J. (1996), *Econometric methods.* McGraw-Hill, New York, NY.

Fuller, W. A. (1996), *Introduction to Statistical Time Series.* Wiley, New York, NY.

Gonzalez-Farias, G. M. (1992), A new unit root test for autoregressive time series. Ph.D. dissertation, North Carolina State University, Raleigh, NC.

Hamilton, J. D. (1994), *Time Series Analysis.* Princeton University Press, Princeton, NJ.

Huh, S. and Dickey, D. A. (1999), Comparison of break induced size distortions among unit root test statistics. Unpublished manuscript, North Carolina State University, Raleigh, NC.

Leybourne, S. J., Mills, T. C. and Newbold, P. (1998), Spurious rejections by Dickey-Fuller tests in the presence of a break under the null. Journal of Econometrics 87, 191-203.

Nelson, C. R. and Plosser, C. I. (1982), Trends and random walks in macroeconomic time series: Some evidence and implications. Journal of Monetary Economics 10, 139-162.

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

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.

Perron, P. (1989), The great crash, the oil price shock, and the unit root hypothesis.

Econometrica 57, 1361-1401.

Perron, P. (1994), Trend, unit root and structural change in macroeconomic time series. In:Rao, B.B.(ed.), Cointegration for the Applied economist. Macmillan, New York, NY, pp113-146.

Perron, P. (1997), Further evidence on breaking trend functions in macroeconomic variables. Journal of Econometrics 80, 355-385.

Poff, N. L. and Allan, J. D. (1995), Functional organization of stream fish assemblages in relation to hydrological variability. Ecology 76, 606-627.

Poff, N. L., Allan, J. D., Bain, M. B, Karr, J. R., Prestegaard, K. L., Richter, B. D., Sparks, R. E. and Stromberg, J. C. (1997), The natural flow regime: a paradigm for river conservation and restoration. BioScience 47, 769-784.

Poff, N. L. and Ward, J. V. (1989), Implications of streamflow variability and predictability for lotic community structure: a regional analysis of stream flow patterns. Canadian Journal of Fisheries and Aquatic Sciences 46, 1805-1818.

Priestley, M. B. (1981), *Spectral analysis and time series.* Academic Press, New York, NY.

Rawlings, J. O., Pantula, S. G. and Dickey, D. A. (1998), *Applied regression analysis : A research tool.* Springer-Verlag, New York, NY.

Richter, B. D., Baumgartner, J. V., Powell, J. and Braun, D. P. (1996), A method for assessing hydrologic alteration within ecosystems. Conservation Biology 10, 1163-1174.

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. (1989), *SAS/STAT User's Guide*, Version 6 (4th ed.), Vol. 1. SAS

Institute Inc., Cary, NC.

SAS Institute Inc. (1994), *The Encore Series : SAS/GRAPH Software*. SAS Institute Inc., Cary, NC.

Vogelsang, T. J. and Perron, P. (1998), Additional tests for a unit root allowing for a break in the trend function at an unknown time. International Economic Review 39, 1073-1100.

Walker, K. F., Sheldon, F. and Puckridge, J. T. (1995), A perspective on dryland river ecosystems. Regulated Rivers: Research & Management 11, 85-104.

Wei, W. W. S. (1990), *Time Series Analysis : Univariate and Multivariate Methods*. Addison-Wesley, New York, NY.

Zivot, E. and Andrews, D. W. K. (1992), Further evidence on the great crash, the oil price shock, and the unit root hypothesis. Journal of Business & Economic Statistics 10, 251-270.

Appendices

Appendix A: Limiting distribution of test statistics allowing for a break based on the WS estimator

We want to derive the limiting distribution of test statistics allowing for a break based on the weighted symmetric estimator. Recall that we are considering a data generating process (DGP)

$$Y_t = m\sigma I(t > c) + X_t, \quad X_t = X_{t-1} + e_t, \quad t = 1, 2, ..., n$$

and define

$$\begin{aligned} \bar{X}_1 &= \frac{1}{c}\sum_{t=1}^{c} X_t \\ \text{and} \qquad \bar{X}_2 &= \frac{1}{n-c}\sum_{t=c+1}^{n} X_t. \end{aligned}$$

We know that the weighted symmetric estimator allowing for a break is

$$\tilde{\rho}_w = \frac{\sum_{t=2}^{n} r_t r_{t-1}}{\sum_{t=2}^{n-1} r_t^2 + \frac{1}{n}\sum_{t=1}^{n} r_t^2}$$

where r_t is the residual from the regression estimation of the model

$$Y_t = \beta_0 + \beta_1 I(t > c) + e_t.$$

Then,

$$r_t = \begin{cases} X_t - \bar{X}_1 & \text{if } t \le c = n\lambda \\ X_t - \bar{X}_2 & \text{if } t > c = n\lambda. \end{cases}$$

The associated pivotal statistic is

$$\tau_w = \frac{\tilde{\rho}_w - 1}{\text{s.e.}} = \frac{\tilde{\rho}_w - 1}{\sqrt{\tilde{\sigma}_w^2 (\sum_{t=2}^{n-1} r_t^2 + \frac{1}{n}\sum_{t=1}^{n} r_t^2)^{-1}}}$$

where

$$\begin{aligned} \tilde{\sigma}_w^2 &= \frac{1}{n-2}[\sum_{t=2}^{n} w_t (r_t - \tilde{\rho}_w r_{t-1})^2 + \sum_{t=1}^{n-1}(1 - w_{t+1})(r_t - \tilde{\rho}_w r_{t+1})^2] \\ &= \frac{1}{n-2}[\sum_{t=2}^{n}(r_t - \tilde{\rho}_w r_{t-1})^2 + (1 - \tilde{\rho}_w^2)(r_1^2 - \frac{1}{n}\sum_{t=1}^{n} r_t^2)]. \end{aligned}$$

and $w_t = \frac{t-1}{n}$ for $t = 1, \cdots, n$.

After some straightforward but tedious algebra, we can show that, for $0 < \lambda < 1$,

$$\begin{aligned} n(\tilde{\rho}_w - 1) &\xrightarrow{d} \frac{N(\lambda)}{D(\lambda)}, \\ \tilde{\sigma}_w^2 &\xrightarrow{d} \sigma^2 S(\lambda) \\ \text{and} \qquad \tau_w &\xrightarrow{d} \frac{N(\lambda)/D(\lambda)}{\sqrt{S(\lambda)/D(\lambda)}} = \frac{N(\lambda)}{\sqrt{S(\lambda) \cdot D(\lambda)}} \end{aligned}$$

where

$$\begin{aligned} N(\lambda) &= 0.5(T^2 - 1) - \frac{1}{1-\lambda} T H_2 + \frac{1}{\lambda(1-\lambda)} H_1 H_2 \\ &\quad - G + \frac{1}{\lambda} H_1^2 + \frac{1}{(1-\lambda)} H_2^2, \\ D(\lambda) &= G - \frac{1}{\lambda} H_1^2 - \frac{1}{1-\lambda} H_2^2 \\ \text{and} \qquad S(\lambda) &= 1 + \frac{1}{\lambda^2} H_1^2 - \frac{2}{\lambda(1-\lambda)} H_1 H_2 + \frac{1}{(1-\lambda)^2} H_2^2. \end{aligned}$$

Here

$$\begin{aligned} G &= \int_0^1 W^2(t)dt, \\ H &= \int_0^1 W(t)dt, \\ H_1 &= \int_0^\lambda W(t)dt, \\ H_2 &= \int_\lambda^1 W(t)dt, \\ \text{and} \qquad T &= W(1) \end{aligned}$$

where $W(t)$ is the standard Wiener process.

If we assume no break ($\lambda = 0$ or 1), then $\bar{X}_1 = \bar{X}_2 = \bar{X} = \frac{1}{n}\sum_{t=1}^n X_t$. We can easily show that

$$\begin{aligned} n(\tilde{\rho}_w - 1) &\xrightarrow{d} \frac{0.5(T^2 - 1) - TH - G + 2H^2}{G - H^2}, \\ \tilde{\sigma}_w^2 &\xrightarrow{d} \sigma^2 \\ \text{and} \qquad \tau_w &\xrightarrow{d} \frac{0.5(T^2 - 1) - TH - G + 2H^2}{\sqrt{G - H^2}}. \end{aligned}$$

These results match with the ones in Theorem 10.1.8, Fuller p 571.

Appendix B: Limiting distribution of the bisection test statistic

Consider a data generating process

$$Y_t = \theta\sigma I(t > c) + W_t, \quad W_t = \rho W_{t-1} + e_t, \quad t = 1, 2, \cdots, n = 100.$$

We are deriving the limiting distribution of $\tau_w^* \equiv \min(\tau_{w,1}, \tau_{w,2})$ under the null hypothesis of $H_0 : \rho = 1$.

We know that

$$\begin{aligned} \bar{Y} &= \bar{W} + (1-\lambda)m\sigma \\ Y_t - \bar{Y} &= \begin{cases} W_t - \bar{W} - (1-\lambda)m\sigma & \text{if } t \leq c = n\lambda \\ W_t - \bar{W} + \lambda m\sigma & \text{if } t > c = n\lambda. \end{cases} \end{aligned}$$

Under $\rho = 1$,

$$W_t = \sum_{j=1}^{t} e_j.$$

We divide the data into 2 groups with $\frac{n}{2} = 50$ observations each. Without loss of generality we can assume further that

$$51 < c < 100$$

,i.e., the break is in the second group. Then

$$\begin{aligned} W_1 &= e_1 \\ W_2 &= e_1 + e_2 \\ &\vdots \\ W_{50} &= e_1 + e_2 + \cdots + e_{50} \end{aligned}$$

and

$$W_{51} = W_{50} + e_{51}$$

$$
\begin{aligned}
W_{52} &= W_{50} + e_{51} + e_{52} \\
&\vdots \\
W_{100} &= W_{50} + e_{51} + e_{52} + \cdots + e_{100}.
\end{aligned}
$$

Let

$$
\begin{aligned}
\bar{W}_1 &= \frac{1}{50}\sum_{t=1}^{50} W_t \\
&= \frac{1}{50}\sum_{t=1}^{50}\sum_{j=1}^{t} e_j
\end{aligned}
$$

and

$$
\begin{aligned}
\bar{W}_2 &= \frac{1}{50}\sum_{t=51}^{100} W_t \\
&= \frac{1}{50}\sum_{t=51}^{100}(W_{50} + \sum_{j=51}^{t} e_j).
\end{aligned}
$$

Then we can easily show that

$$W_t - \bar{W}_1 = \text{function only of } (e_1, e_2, \cdots, e_{50}) \qquad t = 1, 2, ..., 50$$

and

$$W_t - \bar{W}_2 = \text{function only of } (e_{51}, e_{52}, \cdots, e_{100}) \qquad t = 51, 52, ..., 100.$$

Therefore we can conclude that $(W_1 - \bar{W}_1, W_2 - \bar{W}_1, \cdots, W_{50} - \bar{W}_1)$ and $(W_{51} - \bar{W}_2, W_{52} - \bar{W}_2, \cdots, W_{100} - \bar{W}_2)$ are independent.

Since $51 < c < 100$, for the first group,

$$
\begin{aligned}
Y_t &= W_t \qquad t = 1, 2, ..., 50 \\
\bar{Y}_1 &= \frac{1}{50}\sum_{t=1}^{50} Y_t = \bar{W}_1.
\end{aligned}
$$

Therefore

$$
\begin{aligned}
Y_t - \bar{Y}_1 &= W_t - \bar{W}_1 \qquad t = 1, 2, ..., 50 \\
&= \text{function only of } (W_1 - \bar{W}_1, W_2 - \bar{W}_1, \cdots, W_{50} - \bar{W}_1).
\end{aligned}
$$

For the second group,

$$\begin{aligned} Y_t &= \begin{cases} W_t & 51 \le t \le c \\ W_t + m\sigma & c < t \le 100 \end{cases} \\ \bar{Y}_2 &= \frac{1}{50}\sum_{t=51}^{100} Y_t \\ &= \bar{W}_2 + \frac{100-c}{50} m\sigma. \end{aligned}$$

and so

$$\begin{aligned} Y_t - \bar{Y}_2 &= \begin{cases} W_t - \bar{W}_2 - \frac{100-c}{50} m\sigma & 51 \le t \le c \\ W_t - \bar{W}_2 + \frac{c-50}{50} m\sigma & c < t \le 100 \end{cases} \\ &= \text{function only of } (W_{51} - \bar{W}_2, W_{52} - \bar{W}_2, \cdots, W_{100} - \bar{W}_2). \end{aligned}$$

Therefore we can conclude that $(Y_1 - \bar{Y}_1, Y_2 - \bar{Y}_1, \cdots, Y_{50} - \bar{Y}_1)$ and $(Y_{51} - \bar{Y}_2, Y_{52} - \bar{Y}_2, \cdots, Y_{100} - \bar{Y}_2)$ are independent as are

$$\tilde{\rho}_{w,1} = \frac{\sum_{t=2}^{50}(Y_t - \bar{Y}_1)(Y_{t-1} - \bar{Y}_1)}{\sum_{t=2}^{49}(Y_t - \bar{Y}_1)^2 + \frac{1}{50}\sum_{t=1}^{50}(Y_t - \bar{Y}_1)^2}$$

and

$$\tilde{\rho}_{w,2} = \frac{\sum_{t=52}^{100}(Y_t - \bar{Y}_2)(Y_{t-1} - \bar{Y}_2)}{\sum_{t=52}^{99}(Y_t - \bar{Y}_2)^2 + \frac{1}{50}\sum_{t=51}^{100}(Y_t - \bar{Y}_2)^2}.$$

Now let's think about $\tau_{w,1}$ and $\tau_{w,2}$. We also know that $\tau_{w,1}$ and $\tau_{w,2}$ are independent since, by definition,

$$\tau_{w,1} = \frac{\tilde{\rho}_{w,1} - 1}{\sqrt{\tilde{\sigma}^2_{w,1}\{\sum_{t=2}^{49}(Y_t - \bar{Y}_1)^2 + \frac{1}{50}\sum_{t=1}^{50}(Y_t - \bar{Y}_1)^2\}^{-1}}}$$

and

$$\tau_{w,2} = \frac{\tilde{\rho}_{w,2} - 1}{\sqrt{\tilde{\sigma}^2_{w,2}\{\sum_{t=52}^{99}(Y_t - \bar{Y}_2)^2 + \frac{1}{50}\sum_{t=51}^{100}(Y_t - \bar{Y}_2)^2\}^{-1}}}$$

where

$$\tilde{\sigma}^2_{w,1} = \frac{1}{48}[\sum_{t=2}^{50} w_t\{(Y_t - \bar{Y}_1) - \tilde{\rho}_{w,1}(Y_{t-1} - \bar{Y}_1)\}^2$$

$$
\begin{aligned}
& +\sum_{t=1}^{49}(1-w_{t+1})\{(Y_t-\bar{Y}_1)-\tilde{\rho}_{w,1}(Y_{t+1}-\bar{Y}_1)\}^2] \\
= \; & \frac{1}{48}[\sum_{t=2}^{50}\{(Y_t-\bar{Y}_1)-\tilde{\rho}_{w,1}(Y_{t-1}-\bar{Y}_1)\}^2 \\
& +(1-\tilde{\rho}^2_{w,1})\{(Y_1-\bar{Y}_1)^2-\frac{1}{50}\sum_{t=1}^{50}(Y_t-\bar{Y}_1)^2\}]
\end{aligned}
$$

and

$$
\begin{aligned}
\tilde{\sigma}^2_{w,2} \; = \; & \frac{1}{48}[\sum_{t=52}^{100} w_t\{(Y_t-\bar{Y}_2)-\tilde{\rho}_{w,2}(Y_{t-1}-\bar{Y}_2)\}^2 \\
& +\sum_{t=51}^{99}(1-w_{t+1})\{(Y_t-\bar{Y}_2)-\tilde{\rho}_{w,2}(Y_{t+1}-\bar{Y}_2)\}^2] \\
= \; & \frac{1}{48}[\sum_{t=52}^{100}\{(Y_t-\bar{Y}_2)-\tilde{\rho}_{w,2}(Y_{t-1}-\bar{Y}_2)\}^2 \\
& +(1-\tilde{\rho}^2_{w,2})\{(Y_{51}-\bar{Y}_2)^2-\frac{1}{50}\sum_{t=51}^{100}(Y_t-\bar{Y}_2)^2\}].
\end{aligned}
$$

Our example so far has $n = 100$. We can now generalize to n observations and split into 2 groups of $\frac{n}{2}$ observations each.

By Theorem 10.1.8, Fuller, $\tau_{w,1}$ and $\tau_{w,2}$ are independent with common limiting distribution

$$
\frac{0.5(T^2-1)-TH-G+2H^2}{\sqrt{G-H^2}}
$$

since we showed in Chapter 2 that a fixed level shift does not affect the limiting distribution of τ_w. Recall that we showed that the asymptotic distributions of the tests based on the symmetric estimators under $\rho = 1$ are unaffected by a structural break of fixed size, as suggested for the tests based on the OLS estimator by Amsler and Lee (1995).

Therefore

$$
\begin{aligned}
& P(\tau^*_w \le x) \\
= \; & 1-P(\min(\tau_{w,1},\tau_{w,2}) > x)
\end{aligned}
$$

$$\begin{aligned}
&= 1 - P(\tau_{w,1} > x, \tau_{w,2} > x) \\
&= 1 - P(\tau_{w,1} > x)P(\tau_{w,2} > x) \\
&= 1 - \{1 - P(\tau_{w,1} \le x)\}\{1 - P(\tau_{w,2} \le x)\} \\
&\overset{n\to\infty}{\to} 1 - \{1 - P(X \le x)\}\{1 - P(X \le x)\} \\
&= 2P(X \le x) - \{P(X \le x)\}^2
\end{aligned}$$

where

$$X \sim \frac{0.5(T^2-1) - TH - G + 2H^2}{\sqrt{G - H^2}}.$$

Here

$$\begin{aligned}
G &= \int_0^1 W^2(t)dt, \\
H &= \int_0^1 W(t)dt \\
\text{and} \qquad T &= W(1)
\end{aligned}$$

where $W(t)$ is the standard Wiener process.