VAASAN YLIOPISTO

TERHI LUOMA

# Nonparametric Event Study Tests for Testing Cumulative Abnormal Returns

ACTA WASAENSIA NO 254

---

STATISTICS 6

UNIVERSITAS WASAENSIS 2011

Reviewers

Professor Erkki Liski
Mathematics and Statistics
School of Information Sciences
FI–33014 University of Tampere
Finland

Professor Markku Lanne
University of Helsinki
Department of Political and Economic Studies
P.O. Box 17
FI–00014 University of Helsinki
Finland

**Julkaisija**
Vaasan yliopisto

**Julkaisupäivämäärä**
Marraskuu 2011

**Tekijä(t)**
Terhi Luoma

**Julkaisun tyyppi**
Monografia

**Julkaisusarjan nimi, osan numero**
Acta Wasaensia, 254

**Yhteystiedot**
Vaasan yliopisto
Teknillinen tiedekunta
Matemaattisten tieteiden yksikkö
PL700
65101 Vaasa

**ISBN**
978–952–476–372–1

**ISSN**
0355–2667, 1235–7936

**Sivumäärä**
88

**Kieli**
Englanti

**Julkaisun nimike**
Epäparametriset tapahtumatutkimuksen testit testattaessa kumulatiivisia epänormaaleja tuottoja

**Tiivistelmä**

Tämän tutkimuksen tarkoitus on kehittää uusia epäparametrisiä tapahtumatutkimuksen testejä kumulatiivisien epänormaalien tuottojen testaamiseen. Kumulatiivisia epänormaaleja tuottoja käytetään tapahtumatutkimuksissa, jotta voidaan ottaa huomioon mahdollinen epävarmuus tapahtuman esiintymisajankohdasta ja siitä kuinka nopeasti tapahtuma vaikuttaa osakkeiden hintoihin. Monet tapahtumatutkimukset pohjautuvat parametrisiin testeihin, mutta parametristen testien ongelmana on se, että ne sisältävät yksityiskohtaisia oletuksia tuottojen todennäköisyysjakaumasta. Epäparametriset testit eivät yleensä vaadi niin tarkkoja oletuksia tuottojen jakaumasta kuin parametriset testit. Jotkin epäparametriset testit on kuitenkin johdettu vain yhden päivän tapahtumaikkunalle.

Tutkimuksessa johdetaan uusia epäparametrisiä järjestysluku- ja merkkitestejä testaamaan kumulatiivisia epänormaaleja tuottoja. Tutkimuksessa johdetaan myös näiden testisuureiden asymptoottiset ominaisuudet. Simulaatiot, joissa käytetään todellisia tuottoja, osoittavat näiden uusien testisuureiden ominaisuudet verrattuna muihin hyvin tunnettuihin parametrisiin ja epäparametrisiin testisuureisiin.

Simulointitulokset osoittavat, että erityisesti uudet epäparametriset testisuureet SIGN-GSAR-T ja CUMRANK-T omaavat kilpailukykyisiä empiirisiä ominaisuuksia. Nämä testisuureet hylkäävät lähelle nimellistasoa, ovat robusteja tapahtumasta johtuvalle volatiliteetille ja omaavat hyvät empiiriset voimaominaisuudet. Lisäksi, mikäli tapahtumapäivät ovat keskenään klusteroituneita, testisuure SIGN-GSAR-T päihittää tarkastellut parametriset ja epäparametriset testisuureet.

**Asiasanat**
Tapahtumatutkimus, epäparametrinen, kumulatiiviset epänormaalit tuotot, järjestysluku, merkki, simulaatio

| **Publisher**<br>Vaasan yliopisto | **Date of publication**<br>November 2011 | |
|---|---|---|
| **Author(s)**<br>Terhi Luoma | **Type of publication**<br>Monograph | |
| | **Name and number of series**<br>Acta Wasaensia, 254 | |
| **Contact information**<br>University of Vaasa<br>Faculty of Technology<br>Department of Mathematics and Statistics<br>P.O. Box 700<br>FI–65101 Vaasa, Finland | **ISBN**<br>978–952–476–372–1<br>**ISSN**<br>0355–2667, 1235–7936 | |
| | **Number of pages**<br>88 | **Language**<br>English |

**Title of publication**
Nonparametric Event Study Tests for Testing Cumulative Abnormal Returns



**Abstract**

The contribution of this thesis is to develop new nonparametric event study tests for testing cumulative abnormal returns (CARs). CARs are used in event studies to account for potential imprecision in dating the event or uncertainty of the speed of the event's effect on security prices. Many event studies rely on parametric test statistics, but the disadvantage of parametric test statistics is that they embody detailed assumptions about the probability distribution of returns. Nonparametric statistics do not usually require as stringent assumptions about return distributions as parametric tests. Nonetheless, some of the nonparametric test statistics are derived only for one-day abnormal returns.

In the following, new nonparametric rank and sign test statistics for testing CARs are derived together with their asymptotic properties. Simulations with actual returns show the empirical properties of these new test statistics compared with other well-known parametric and nonparametric test statistics.

The simulation results reveal that the new nonparametric tests statistics SIGN-GSAR-T and CUMRANK-T in particular have competitive empirical properties. Those test statistics reject close to the nominal level, are robust against event-induced volatility and have good empirical power properties. Moreover, if the event-dates are clustered, the statistic SIGN-GSAR-T outperforms the examined parametric and nonparametric test statistics.



**Keywords**

Event study, nonparametric, cumulative abnormal return, rank, sign, simulation

# ACKNOWLEDGEMENTS

Summer 2007 was special for me, even though I spent it in Italy as I have done every summer in the past years. During that summer I finally decided that I would take a new step in my life. I knew that it would change my life, and it really did, even more than I have thought. After four years of hard work and some sleepless nights, I can say that I am happy that I made that decision in summer 2007 somewhere in Northern Italy. There have been many people who have helped me in taking that step and now I would like to thank them.

I would like to express my deepest gratitude to Professor Seppo Pynnönen, who introduced me to event studies and who has guided me in this thesis. I highly respect his professional instruction and the way how he has pushed me forward. I would like to thank the two preliminary examiners Professor Erkki Liski and Professor Markku Lanne for their careful and detailed examination. They have given me valuable ideas how to improve this thesis.

I would like to thank all my colleagues at the Department of Mathematics and Statistics and at the Department of Accounting and Finance for enabling me to work in an encouraging and inspiring working environment. My work has also benefitted from the suggestions of my colleagues at the PhD seminar organized by these departments. I would also like to offer some special thanks. Professor Seppo Hassi as the leader of the department has provided me an excellent framework to do my PhD studies and research. Dr. Bernd Pape has given me many valuable comments and helped me in editing this thesis. My colleague and friend Emilia Peni has given me so much advice and I highly value the discussions we have had.

Suomen Arvopaperimarkkinoiden Edistämissäätiö, OP-Pohjolaryhmän tutkimussäätiö, the Foundation of Evald and Hilda Nissi and Kauhajoen kulttuurisäätiö have all supported me financially, and I would therefore like to express my gratitude to those foundations. I would also like to thank the Finnish Doctoral Programme in Stochastics and Statistics for the significant financial support.

Conducting research is sometimes lonely, but I am happy to have people who have made me to shut down the computer for a while and who have also given me something else to think about. I would like to thank my family members, who have understood and supported me, each in their own way. Therefore, warm thanks to all my family members in the Luoma and Hagnäs families as well as Ginevra.

There have been days when I have doubted and not had any idea where I will end up with my studies and research. Nonetheless during these ten years of studies and work I have always had something stable in my life: a person who has always believed in me and my abilities, who has made me to continue to follow my dreams and who has given me faith, strength and love. My most sincere thanks go to that person, Ari.

Vaasa, November 2011

Terhi Luoma

# Contents

# List of Figures

# List of Tables

# 1 INTRODUCTION

Economists are frequently requested to measure the effect of some economic event on the value of a stock. The question could be, for example, what happens to the stock price at the reinvestment date. If the market is efficient, then on average the stock price falls by the amount of the dividend. Otherwise one has an opportunity for economic profit. The event study method is developed as a statistical tool for solving questions like this, which are focused on abnormal returns (ARs). The general applicability of the event study methodology has led to its wide use and nowadays it is one of the most frequently used analytical tools in financial research. Hence, in accounting and finance research, event studies have been applied to a variety of firm specific and economy wide events. Some examples include mergers and acquisitions, earnings announcements, issues of new debt or equity and announcements of macroeconomic variables such as trade deficit. However, applications in other fields are also abundant. Event studies are also used in the fields of law, economics, marketing, management, history and political science, among others.

Even though event study methodology has a number of different potential applications, for the most part this study is made from the viewpoint of financial events. The aim of this study is to present new nonparametric test statistics for testing cumulative abnormal returns (CARs), derive their asymptotical properties and consider the empirical properties of the new test statistics compared to other widely known parametric and nonparametric test statistics.

Section 2 focuses on the general background of nonparametric testing methods and Section 3 discusses the background of event study testing methods. The new nonparametric rank and sign test statistics are presented in Section 4. Also the asymptotical properties of these test statistics are presented in the cases, where the observations are independent and as well in the cases, where the event dates are clustered. In Section 5, the simulation construction and abnormal return model are presented together with the test statistics to which the new rank and sign test statistics are compared. Section 5 also presents the data to be used in the empirical simulations. Section 6 presents the empirical simulation results. The sample statistics, empirical distributions, rejection rates and powers of the tests are investigated. The conclusions of the study are discussed in Section 7.

# 2 GENERAL BACKGROUND OF NONPARAMETRIC TESTING METHODS

## 2.1 Definitions of Nonparametric Testing Methods

In much elementary statistic material, the subject matter of statistics is usually somewhat arbitrarily divided into two categories called descriptive and inductive statistics. Descriptive statistics usually relates only to the presentation of figures or calculations to summarize or characterize a dataset. For such procedures, no assumptions are made or implied, and there is no question of legitimacy of techniques. The descriptive statistics may be, for example, sample statistics like a mean, median and variance or a histogram. When sample descriptions are used to infer some information about the population, the subject is called inductive statistics or statistical inference. The two types of problems most frequently encountered in this kind of subject are estimation and testing of a hypothesis. The entire body of classical statistical inference techniques is based on fairly specific assumptions regarding the nature of the underlying population distribution: usually its form and some parameter values must be stated. However, in the reality everything does not come packaged with labels of population of origin and a decision must be made as to what population properties may judiciously be assumed for the model. An alternative set of techniques is also available and those may be classified as distribution-free and nonparametric procedure. [Gibbons and Chakraborti (1992)].

The definition of nonparametric varies slightly between authors. For example Gibbons (1976) has stated that statistical inferences that are not concerned with the value of one or more parameters would logically be termed nonparametric. Those inferences whose validity does not rest on a specific probability model in the population would logically be termed distribution-free. Also Bradley (1968) has concluded that the terms nonparametric and distribution-free are not synonymous. Broadly speaking, a nonparametric test is one which makes no hypothesis about the value of a parameter in a statistical density function, whereas a distribution-free test is one which makes no assumptions about the precise form of the sampled population. The definitions are not mutually exclusive and a test can be both distribution-free and parametric. [Bradley (1968)].

Many nonparametric procedures are described as rank or sign tests. Rank tests are based on ranked data and in those tests the data is ranked by ordering the observations from lowest to highest and assigning them, in order, the integer values from one to the sample size. Sign tests, on the other hand, use plus and minus signs of the observations rather than quantitative measures as its data.

## 2.2 General Advantages and Disadvantages of Nonparametric Testing Methods

According to Hettmansperger and McKean (2011) for example, nonparametric testing methods have a long and successful history extending back to early work by Wilcoxon (1945), who introduced rank-sum and signed rank tests. For example Daniel (1990) has concluded that nonparametric tests usually make less stringent demands on the data and since most nonparametric procedures depend on a minimum of assumptions, they are not usually improperly used. Gibbons (1976) has concluded that the attribute of nonparametric methods that may be most persuasive to the investigator who is not a professional statistician is that he is somewhat less likely to misuse statistics when applying nonparametric techniques than when using those methods that are parametric according to our definitions. The easiest way to abuse any statistical technique is to disregard or violate the assumptions necessary for the validity of the procedure.

Gibbons and Chakraborti (1992) have stated that when using nonparametric methods the basic data available need not be actual measurements. For example in many cases, if the test is to be based on ranks, only ranks are needed. Therefore, the process of collecting and compiling sample data then may be less expensive and time-consuming. Daniel (1990) has stated that for some nonparametric procedures, the computations can be quickly and easily performed. Therefore, researchers with minimum preparation in mathematics and statistics usually find the concepts and methods of nonparametric procedures easy to understand.

Daniel (1990) has also stated that although nonparametric procedures have a reputation for requiring only simple calculations, the arithmetic in many instances is tedious and laborious especially when samples are large and a high-powered computer is not available. For example Siegel (1956) has stated that if all the assumptions of the parametric statistical model are met in the data, and if the measurement is of the required strength, then nonparametric statistical tests are wasteful of data. Hence, some researchers think that the nonparametric procedures throw away information.

# 3 BACKGROUND OF EVENT STUDY TESTING METHODS

## 3.1 Overview of the History of the Event Study

As Campbell, Lo and MacKinlay (1997) and others have concluded, event studies have a long history. Perhaps the first published event study was conducted as early as the beginning of the 1930s by Dolley (1933). Dolley examined the price effects of stock splits, studying nominal price changes at the time of the split. Dolley planted a seed of event study that continues to flourish decades later. In the late 1960s seminal studies by Ball and Brown (1968), and Fama, Fisher, Jensen and Roll (1969) introduced the event study methodology to a broad audience of accounting and financial economists. That methodology is essentially the same as that which is in use today. Ball and Brown studied the information content of earnings while Fama, Jensen and Roll studied the effects of stock splits after removing the effects of simultaneous dividend increases.

Campbell, Lo and MacKinlay (1997) have also stated that in the years since those pioneering studies, several modifications of the basic methodology have been suggested, and two main changes in the methodology have taken place. First, the use of daily rather than monthly security return data has become relevant. Second, the methods used to estimate abnormal returns and calibrate their statistical significance have become more sophisticated. Useful papers which deal with the modifications of the event study methodology are the works by Brown and Warner published in 1980 and 1985. The former paper considers implementation issues for data sampled at a monthly interval and the later paper deals with issues for daily data.

It is not known precisely how many event studies have been published. Kothari and Warner (2007) report that over the period 1974–2000, five major finance journals published 565 articles containing event study results. As they concluded this is clearly a very conservative number as it does not include the many event studies published in accounting journals and other finance journals. Moreover, event studies are also published outside the realm of mainstream accounting and finance journals.

## 3.2 Outline of the Event Study

There have been many advances in event study methodology over the years, but the core elements of a typical event study are usually the same. As Campbell, Lo and MacKinlay (1997, Ch. 4.1) have presented, the event study analysis can be viewed as having seven steps.

The first step is to define the event of interest and identify the period over which the security prices of the firms involved in this event will be examined. That period is called the event window or event period. The second step is to determine the selection criteria for the inclusion of a given firm in the study. At this stage it is useful to summarize some characteristics of the data sample and note potential biases which may have been introduced through the sample selection.

To appraise the event's impact a measure of the abnormal return (AR) is required. The third step in the event study analysis is to define the normal returns and the ARs. The AR is the actual ex post return of the security over the event window minus the normal or expected return of the firm over the event window. The normal return is defined as the return that would be expected if the event did not take place. Once a normal performance model has been selected, the parameters of the model must be estimated using a subset of the data known as the estimation window or estimation period. Usually the estimation window is the period prior to the event window and usually the event window itself is not included in the estimation window to prevent the event from influencing the normal performance model parameter estimates. This step is the fourth step.

The fifth step is the defining of the testing framework for the ARs. Important considerations are defining the null hypothesis and determining the techniques for aggregating the ARs of individual firms. The sixth step is the presentation of the empirical results. In addition to presenting the basic empirical results, the presentation of diagnostics can be fruitful. The last step is interpretation and conclusions. As Campbell, Lo and MacKinlay (1997, Ch. 4.1) conclude, the empirical results will ideally lead to insights about the mechanism by which the event affects security prices.

## 3.3 Widely Known Event Study Test Statistics

### *3.3.1 Parametric test statistics*

There are numerous tests for evaluating the statistical significance of abnormal returns (ARs). Perhaps the most widely used parametric test statistics are an ordinary $t$-statistic and test statistics derived by Patell (1976), and Boehmer, Musumeci and Poulsen (BMP) (1991).

Patell (1976) proposed a test statistic, in which the event window ARs are standardized by the standard deviation of the estimation window ARs. This standardization reduces the effect of stocks with large return standard deviations on the test. Patell's test statistic assumes cross-sectional independence in the ARs, and it also assumes that the ARs are normally distributed. For example Campbell and Wasley (1993) have reported that the Patell's test rejects the true null hypothesis too often with Nasdaq samples due to the non-normality of Nasdaq returns, particularly lower priced and less liquid securities. Cowan and Sergeant (1996) also report such excessive rejections in Nasdaq samples in upper-tailed but not lower-tailed tests. Maynes and Rumsey (1993) report a similar misspecification of the test using the most thinly traded one-third of Toronto Exchange stocks. Also Kolari and Pynnönen (2010) have concluded that Patell's test is sensitive to event-induced volatility and rejects the null hypothesis too often.

BMP (1991) have introduced a variance-change corrected version of the Patell's test. Their test statistic has gained popularity over the Patell's statistic, because it has been found to be more robust with respect to possible volatility changes associated with the event. For example, BMP (1991) have reported that their test is correctly specified in NYSE-AMEX samples under null even when there is an increase in variance of stock returns on the event date.

We can conclude that due to their better power properties the standardized tests of Patell (1976) and BMP (1991) have gained in popularity over the conventional nonstandardized tests in testing event effects on mean security price performance. Harrington and Shrider (2007) have found that a short-horizon test focusing on mean ARs should always use tests that are robust against cross-sectional variation in the *true* AR [for discussion of *true* AR, see Harrington and Shrider (2007)]. They have found that the test statistic BMP is a good candidate for a robust, parametric test in conventional event studies.[1]

[1]The current research defines conventional event studies as those focusing only on mean stock price

### *3.3.2 Nonparametric test statistics*

The use of daily data in event studies is important for isolating stock price reactions against announcements. However, for example Fama (1976) has found that a potential problem with the use of daily returns is that daily stock returns depart from normality more than do monthly returns. The evidence generally suggests that distributions of daily returns are fat-tailed relative to a normal distribution [e.g. Fama (1976)]. Brown and Warner (1985) have shown that the same holds true for daily excess returns. However, generally the normality of abnormal returns is a key assumption underlying the use of parametric test statistics in event studies and therefore a disadvantage of parametric test statistics is that they embody detailed assumptions about the probability distribution of returns. Nonparametric statistics do not usually require such stringent assumptions about return distributions as parametric tests. [e.g. Cowan (1992)].

Corrado (1989) [and Corrado and Zivney (1992)] have introduced a nonparametric rank test based on standardized returns, which has proven to have very competitive and often superior power properties over the above mentioned standardized tests [e.g. Corrado (1989), Corrado and Zivney (1992), Campbell and Wasley (1993) and Kolari and Pynnönen (2010)]. Furthermore, the rank test of Corrado and Zivney (1992) based on the event period re-standardized returns has proven to be both robust against event-induced volatility [Campbell and Wasley (1993)] and to cross-correlation due to event-day clusterings [Kolari and Pynnönen (2010)].

Also sign tests are nonparametric tests, which are often used in event studies. Additionally, nonparametric procedures like the sign tests can be misspecified, if an incorrect assumption about the data is imposed. For example Brown and Warner (1980) and (1985), and Berry, Gallinger and Henderson (1990) have demonstrated that a sign test assuming an excess return median of zero is misspecified. Corrado and Zivney (1992) have introduced a sign test based on standardized excess returns that does not assume a median of zero, but instead uses a sample excess return median to calculate the sign of an event date excess return. The results of simulation experiments presented in Corrado and Zivney (1992) indicate that their sign test provides reliable and well-specified inferences in event studies. They also have reported that their version of the sign test is better specified than the ordinary $t$-test and has a power advantage over the ordinary

effects. As e.g. Kolari and Pynnönen (2011) have concluded, other types of event studies include the examination of return variance effects [Beaver (1968) and Patell (1976)], trading volume [Beaver (1968) and Campbell and Wasley (1996)], accounting performance [Barber and Lyon (1997)] and earnings management procedures [Dechow, Sloan and Sweeney (1995) and Kothari, Leone and Wasley (2005)].

$t$-test in detecting small levels of abnormal performance. In addition, for example, Corrado (2010) has summarized that nonparametric sign and rank tests are recommended for applications, where robustness against non-normally distributed data is desirable.

## 3.4 Testing for Cumulative Abnormal Returns

Identification of the correct event date is essential in event studies. A one-day event period that includes the announcement day only is the best choice, if the announcement date is known exactly. In practice, however, it is not always possible to pinpoint the time when the new information reaches investors. Consequently, there is a trade-off, because if the event window is too short, it may not include the time when investors truly learn about the event. On the other hand, if it is too long, other information will make the statistical detection onerous and less reliable. In practice, the period of interest is often expanded to several days, including at least the day of the announcement and some days before and after the announcement. Therefore, the accumulating of the ARs has an advantage when there is uncertainty about the event date. Many parametric tests, like the tests derived by Patell (1976) and BMP (1991) and the ordinary $t$-statistic can be rapidly applied to testing CARs over multiple day windows. However, many nonparametric tests are derived only for one-day ARs. Thus, there is demand for new improved nonparametric tests for event studies.

Campbell and Wasley (1993) have extended the event study rank test derived by Corrado (1989) for testing cumulative abnormal returns. The test statistic is hereafter called CAMPBELL-WASLEY. The ranks are dependent on construction, which introduces incremental bias into the standard error of the statistic in longer CARs. In Section 4 the bias will be corrected and a new $t$-ratio, which is called CUMRANK-T, will be derived. In addition a rank test statistic called CUMRANK-Z, which is essentially the same test statistic as proposed in Corrado and Truong (2008, p. 504), will be presented. Also asymptotic distributions for rank test statistics CAMPBELL-WASLEY, CUMRANK-T and CUMRANK-Z with fixed time series length will be derived. The statistic CUMRANK-T is well specified under the null hypothesis of no event mean effect and is robust to event-induced volatility. The simulation study with actual return data in Section 6 also will reveal that this test statistic has superior empirical power against the parametric tests considered. Again, consistent with the theoretical derivations, the simulation results with actual returns will confirm that in longer accumulation windows the test statistic tends to reject the null hypothesis closer to the nominal rate than the rank test based on approach suggested in Campbell and Wasley (1993). The test statistic CAMPBELL-WASLEY suffers from a small technical bias in the standard error of the statistic that does not harm the statistic in short period CARs but cause under-rejection of the null hypothesis in longer CARs.

Kolari and Pynnönen (2011) have also derived a nonparametric rank test of CARs, which is based on generalized standardized abnormal returns (GSARs). They have found that their rank test has superior (empirical) power relative to popular parametric tests both at short and long CAR-window lengths. Their test statistic has also been shown to be robust to abnormal return serial correlation and event-induced volatility. Kolari and Pynnönen (2011) have also suggested that GSARs derived by them can be used to extend the sign test in Corrado and Zivney (1992) for testing CARs. Hence, in Section 4 new sign test statistics (SIGN-GSAR-T and SIGN-GSAR-Z) based on GSARs, will be presented. These statistics can be used equally well for testing ARs and CARs. Cowan (1992) has also derived a sign test (called hereafter SIGN-COWAN) for testing CARs. The test statistic SIGN-COWAN compares the proportion of positive ARs around an event to the proportion from a period unaffected by the event. In this way the test statistic SIGN-COWAN takes account of a possible asymmetric return distribution under the null hypothesis. Cowan (1992) has reported that the test he derived is well specified for event windows of one to eleven days. He has also reported that the test is powerful and becomes relatively more powerful as the length of the CAR-window increases. The results of this study from the empirical simulations will show that the sign test statistic SIGN-GSAR-T especially has several advantages over many previous testing procedures, for example, being robust to the event-induced volatility and having good empirical power properties.

For example, according to Kolari and Pynnönen (2010) it is well known that event studies are prone to cross-sectional correlation among ARs when the event day is the same for sample firms. For this reason the test statistics cannot assume independence of ARs. They have also shown that even when cross-correlation is relatively low, event-date clustering is serious in terms of over-rejecting the null hypothesis of zero average ARs, when it is true. Also Section 6 will report that when the event-dates are clustered, many of the test statistics over-reject the null hypothesis for both short and long CAR-windows. Both new test statistics CUMRANK-T and SIGN-GSAR-T are quite robust against a certain degree of cross-sectional correlation caused by event day clustering. Thus, the new rank and sign procedures make available nonparametric tests for general application to the mainstream of event studies.

# 4 DERIVING THE RANK AND SIGN TEST STATISTICS FOR TESTING CUMULATIVE ABNORMAL RETURNS

This section will show that the variance estimator in the rank test statistic derived by Campbell and Wasley (1993) (CAMPBELL-WASLEY) is biased, and a new rank test (CUMRANK-T) based on a corrected variance estimator is suggested. Moreover, a modification of the test of Corrado and Truong (2008) for scaled ranks (CUMRANK-Z) is introduced. In addition to these rank tests, two new sign tests (SIGN-GSAR-T and SIGN-GSAR-Z) based on generalized standardized abnormal returns (GSARs) are proposed. The theoretical analysis of this section reveals that the CUMRANK-Z and SIGN-GSAR-Z are not robust with respect to cross-sectional correlation of the abnormal return series. Whereas, the CUMRANK-T and SIGN-GSAR-T tests are preferable when clustering is present.

Hence, this section first introduces necessary notations and concepts. Second, the distribution properties of the rank tests are derived and the rank test statistics CAMPBELL-WASLEY, CUMRANK-T, CUMRANK-Z, and the modified version of CAMPBELL-WASLEY are presented. Third, the GSAR is presented and the sign of the GSAR is derived. Fourth, the sign test statistics SIGN-GSAR-T and SIGN-GSAR-Z are presented. Fifth, the asymptotic distributions of the rank and sign test statistics for independent observations as well as for clustered event dates are derived.

## 4.1 Basic Concepts

The autocorrelations of the stock returns are assumed to be negligible and the following assumption is made:

**Assumption 1.** *Stock returns $r_{it}$ are weak white noise continuous random variables with*

$$\begin{aligned} E[r_{it}] &= \mu_i \textit{ for all } t, \\ \operatorname{var}[r_{it}] &= \sigma_i^2 \textit{ for all } t, \\ \operatorname{cov}[r_{it}, r_{is}] &= 0 \textit{ for all } t \neq s, \end{aligned} \tag{1}$$

*and where $i$ refers to the $i^{th}$ stock, and $t$ and $s$ are time indexes. Furthermore $i = 1, \ldots, n$, $t = 1, \ldots, T$ and $s = 1, \ldots, T$.*

Let $\mathrm{AR}_{it}$ represent the abnormal return of security $i$ on day $t$, and let day $t = 0$ indicate the event day.[2] The days $t = T_0 + 1, T_0 + 2, \ldots, T_1$ represent the estimation window days relative to the event day, and the days $t = T_1 + 1, T_1 + 2, \ldots, T_2$ represent event window days, again relative to the event day. Furthermore $L_1$ represents the estimation window length and $L_2$ represents the event window length. Standardized abnormal return (SAR) is defined as

$$\mathrm{SAR}_{it} = \mathrm{AR}_{it}/S(\mathrm{AR}_i), \tag{2}$$

where $S(\mathrm{AR}_i)$ is the standard deviation of the regression prediction errors in the abnormal returns computed as in Campbell, Lo and MacKinlay (1997, Sections 4.4.2–4.4.3).

The cumulative abnormal return (CAR) from day $\tau_1$ to $\tau_2$ with $T_1 < \tau_1 \leq \tau_2 \leq T_2$ is defined as

$$\mathrm{CAR}_{i,\tau_1,\tau_2} = \sum_{t=\tau_1}^{\tau_2} \mathrm{AR}_{it}, \tag{3}$$

and the time period from $\tau_1$ to $\tau_2$ is often called a CAR-window or a CAR-period.

Then the corresponding standardized cumulative abnormal return (SCAR) is defined as

$$\mathrm{SCAR}_{i,\tau_1,\tau_2} = \frac{\mathrm{CAR}_{i,\tau_1,\tau_2}}{S(\mathrm{CAR}_{i,\tau_1,\tau_2})}, \tag{4}$$

where $S(\mathrm{CAR}_{i,\tau_1,\tau_2})$ is the standard deviation of the CARs adjusted for forecast error [see Campbell, Lo and MacKinlay (1997, Section 4.4.3)]. Under the null hypothesis of no event effect both $\mathrm{SAR}_{it}$ and $\mathrm{SCAR}_{i,\tau_1,\tau_2}$ are distributed with mean zero and (approximately) unit variance.

[2]There are different ways to define the abnormal returns ($AR_{it}$). One quite often used method is to use market model to estimate the abnormal returns. Section 5 presents how the abnormal returns can be calculated with the help of the market model.

## 4.2 The Rank Tests for Testing Cumulative Abnormal Returns

### 4.2.1 *Distribution properties of the rank tests to be developed*

For the purpose of accounting for the possible event induced volatility, the re-standardized abnormal return is defined in the manner of BMP (1991) [see also Corrado and Zivney (1992)] as

$$\mathrm{SAR}'_{it} = \begin{cases} \mathrm{SAR}_{it}/S_{\mathrm{SAR}_t} & \text{in CAR-window} \\ \mathrm{SAR}_{it} & \text{otherwise,} \end{cases} \tag{5}$$

where

$$S_{\mathrm{SAR}_t} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(\mathrm{SAR}_{it} - \overline{\mathrm{SAR}}_t)^2} \tag{6}$$

is the cross-sectional standard deviation of $\mathrm{SAR}_{it}$s, $\overline{\mathrm{SAR}}_t = \frac{1}{n}\sum_{i=1}^{n}\mathrm{SAR}_{it}$, and $n$ is the number of stocks in the portfolio. Furthermore, let $R_{it} = \mathrm{rank}(\mathrm{SAR}'_{it})$ denote the rank number of re-standardized abnormal series $\mathrm{SAR}'_{it}$, where $R_{it} \in \{1,\ldots,T\}$, for all $t = 1,\ldots,T$ and $i = 1,\ldots,n$. With Assumption 1 and under the null hypothesis of no event effect, each value of $R_{it}$ is equally likely, implying $\Pr[R_{it} = k] = 1/T$, for all $k = 1,\ldots,T$. That is, the ranks have a discrete uniform distribution between values 1 and $T$, for which the expectation and variance are

$$E\left[R_{it}\right] = \frac{T+1}{2} \tag{7}$$

and

$$\mathrm{var}\left[R_{it}\right] = \frac{T^2-1}{12}. \tag{8}$$

Because each observation is associated to a unique rank, the ranks are not independent.[3] It is straightforward to show that the covariance of the ranks is [see e.g. Gibbons and Chakraborti (1992)]

$$\mathrm{cov}[R_{it}, R_{is}] = -\frac{(T+1)}{12}. \tag{9}$$

[3]Thus, if abnormal return $\mathrm{AR}_{it}$ has a rank value $R_{it} = m$, then a return at any other point in time can have any other rank value of the remaining $T-1$ ones, again equally likely.

With these results the major statistical properties of the cumulative ranks can be derived. These and more general moment properties can also be found in the classics research of Wilcoxon (1945) and Mann and Whitney (1947).

Cumulative rank for individual return series is defined as

$$S_{i,\tau_1,\tau_2} = \sum_{t=\tau_1}^{\tau_2} R_{it}, \tag{10}$$

where $i = 1,\ldots,n$ and $T_1 < \tau_1 \leq \tau_2 \leq T_2$. Using (7), the expectation of the cumulative rank is

$$E[S_{i,\tau_1,\tau_2}] = \tau\frac{T+1}{2}, \tag{11}$$

where $\tau = \tau_2 - \tau_1 + 1$ is the number of event days over which $S_{i,\tau_1,\tau_2}$ is accumulated. Because

$$\mathrm{var}[S_{i,\tau_1,\tau_2}] = \sum_{t=\tau_1}^{\tau_2} \mathrm{var}[R_{it}] + \sum_{t=\tau_1}^{\tau_2} \sum_{\substack{s=\tau_1 \\ s\neq t}}^{\tau_2} \mathrm{cov}[R_{it}, R_{is}], \tag{12}$$

using equations (8) and (9) it is straightforward to show that the variance of cumulative ranks is

$$\mathrm{var}[S_{i,\tau_1,\tau_2}] = \frac{\tau(T-\tau)(T+1)}{12}, \tag{13}$$

where $\tau \in \{1,\ldots,T\}$.

In particular, if the available observation on the estimation period varies from one series to another, it is more convenient to deal with scaled ranks. Following Corrado and Zivney (1992) the definition is:

**Definition 1.** *Scaled ranks are defined as*

$$K_{it} = R_{it}/(T+1). \tag{14}$$

Utilizing the above results for unscaled ranks, from (7), (8), and (9) the following proposition is immediately obtained:

**Proposition 1.** *Under the null hypothesis of no event effect the expectation, variance, and covariance of the scaled ranks defined in* (14) *are*

$$E[K_{it}] = \frac{1}{2}, \tag{15}$$

$$\sigma_K^2 = \mathrm{var}[K_{it}] = \frac{T-1}{12(T+1)} \tag{16}$$

*and*

$$\operatorname{cov}[K_{it}, K_{is}] = -\frac{1}{12(T+1)}, \tag{17}$$

*where $i = 1, \ldots, n$, $t \neq s$ and $t, s = 1, \ldots, T$.*

**Remark 1.** *An important result of Proposition 1 is that due to the (discrete) uniform null distribution of the rank numbers with $Pr[K_{it} = t/(T+1)] = 1/T$, $t = 1, \ldots T$, the expected value and the variance of the (scaled) ranks exactly match the sample mean and the sample variance. That is,*

$$\bar{K}_i = \frac{1}{T} \sum_{t=T_0+1}^{T_2} K_{it} = \frac{1}{2} = E\left[K_{it}\right] \tag{18}$$

*and*

$$s^2_{K_i} = \frac{1}{T} \sum_{t=T_0+1}^{T_2} \left(K_{it} - \frac{1}{2}\right)^2 = \frac{T-1}{12(T+1)} = \operatorname{var}\left[K_{it}\right]. \tag{19}$$

Next the cumulative scaled ranks of individual stocks are derived.

**Definition 2.** *The cumulative scaled ranks of a stock i over the event days window from $\tau_1$ to $\tau_2$ are defined as*

$$U_{i,\tau_1,\tau_2} = \sum_{t=\tau_1}^{\tau_2} K_{it}, \tag{20}$$

*where $T_1 < \tau_1 \leq \tau_2 \leq T_2$.*

The expectation and variance of $U_{i,\tau_1,\tau_2}$ $[= S_{i,\tau_1,\tau_2}/(T+1)]$ are again obtained directly by using (11) and (12). The results are summarized in the following proposition:

**Proposition 2.** *The expectation and variance of the cumulative scaled ranks under the null hypothesis of no event effect are*

$$\mu_{i,\tau_1,\tau_2} = E\left[U_{i,\tau_1,\tau_2}\right] = \frac{\tau}{2} \tag{21}$$

*and*

$$\sigma^2_{i,\tau_1,\tau_2} = \operatorname{var}\left[U_{i,\tau_1,\tau_2}\right] = \frac{\tau(T-\tau)}{12(T+1)}, \tag{22}$$

*where $i = 1, \ldots, n$, $T_1 < \tau_1 \leq \tau_2 \leq T_2$ and $\tau = \tau_2 - \tau_1 + 1$.*

Rather than investigating individual (cumulative) returns, the practice in event studies is to aggregate the individual returns into equally weighted portfolios such that:

**Definition 3.** *The cumulative scaled rank is defined as the equally weighted portfolio of the individual cumulative standardized rank defined in* (20)*,*

$$\bar{U}_{\tau_1,\tau_2} = \frac{1}{n}\sum_{i=1}^{n} U_{i,\tau_1,\tau_2}, \tag{23}$$

*or equivalently*

$$\bar{U}_{\tau_1,\tau_2} = \sum_{t=\tau_1}^{\tau_2} \bar{K}_t, \tag{24}$$

*where* $T_1 < \tau_1 \leq \tau_2 \leq T_2$ *and*

$$\bar{K}_t = \frac{1}{n}\sum_{i=1}^{n} K_{it} \tag{25}$$

*is the time t average of scaled ranks.*

The expectation is the same as the expectation of the cumulative rank of individual securities, because

$$E[\bar{U}_{\tau_1,\tau_2}] = \frac{1}{n}\sum_{i=1}^{n} E\left[U_{i,\tau_1,\tau_2}\right] = \frac{\tau}{2}.$$

If the event days are not clustered, the cross-correlations of the return series are zero (or at least negligible). In such a case the variance of (23) is straightforward to calculate. The situation is not much more complicated, if the event days are clustered, which implies cross-correlation. In such a case, recalling that the variances of $U_{i,\tau_1,\tau_2}$ given in equation (22) are constants (independent of $i$), the cross-covariance of $U_{i,\tau_1,\tau_2}$ and $U_{j,\tau_1,\tau_2}$ can be written as

$$\operatorname{cov}\left[U_{i,\tau_1,\tau_2}, U_{j,\tau_1,\tau_2}\right] = \frac{\tau(T-\tau)}{12(T-1)}\rho_{ij,\tau_1,\tau_2}, \tag{26}$$

where $\rho_{ij,\tau_1,\tau_2}$ is the cross-correlation of $U_{i,\tau_1,\tau_2}$ and $U_{j,\tau_1,\tau_2}$, $i,j = 1,\ldots,n$. Utilizing this and the variance-of-a-sum formula, straightforwardly

$$\begin{aligned}
\text{var}\left[\bar{U}_{\tau_1,\tau_2}\right] &= \text{var}\left[\frac{1}{n}\sum_{i=1}^{n} U_{i,\tau_1,\tau_2}\right] \\
&= \frac{1}{n^2}\sum_{i=1}^{n}\text{var}\left[U_{i,\tau_1,\tau_2}\right] + \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j\neq i}^{n}\text{cov}\left[U_{i,\tau_1,\tau_2}, U_{j,\tau_1,\tau_2}\right] \\
&= \frac{1}{n^2}\sum_{i=1}^{n}\frac{\tau(T-\tau)}{12(T-1)} + \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j\neq i}^{n}\frac{\tau(T-\tau)}{12(T-1)}\rho_{ij,\tau_1,\tau_2} \\
&= \frac{\tau(T-\tau)}{12(T+1)n}\left(1+(n-1)\bar{\rho}_{n,\tau_1,\tau_2}\right), \qquad (27)
\end{aligned}$$

where

$$\bar{\rho}_{n,\tau_1,\tau_2} = \frac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{\substack{j=1 \\ j\neq i}}^{n}\rho_{ij,\tau_1,\tau_2} \qquad (28)$$

is the average cross-correlation of the cumulated ranks. This is the main result of this derivation to be utilized later. Therefore, it is summarized in the following theorem:

**Theorem 1.** *Under the null hypothesis of no event effect the expectation and variance of the average cumulated scaled ranks $\bar{U}_{\tau_1,\tau_2}$, defined in* (23)*, are*

$$E\left[\bar{U}_{\tau_1,\tau_2}\right] = \frac{\tau}{2} \qquad (29)$$

*and*

$$\text{var}\left[\bar{U}_{\tau_1,\tau_2}\right] = \frac{\tau(T-\tau)}{12(T+1)n}\left(1+(n-1)\bar{\rho}_{n,\tau_1,\tau_2}\right), \qquad (30)$$

*where $\tau = \tau_2 - \tau_1 + 1$, $T_1 < \tau_1 \leq \tau_2 \leq T_2$, and $\bar{\rho}_{n,\tau_1,\tau_2}$ is defined in* (28).

From a practical point of view a crucial result in Theorem 1 is that the only unknown parameter to be estimated is the average cross-correlation $\bar{\rho}_{n,\tau_1,\tau_2}$. There are potentially several different ways to estimate the cross-correlation. An obvious and straightforward strategy is to construct first $\tau$ period multi-day series from individual scaled rank series and compute the average cross-correlation of them. This is, however, computationally expensive. The situation simplifies materially if the cross-correlation of cumulated ranks are assumed to be the same as the cross-correlation of single day correlations. As will be seen in such a case the average cross-correlation becomes estimated implicitly by using a suitable variance estimator.

### 4.2.2 *The test statistics CUMRANK-Z, CAMPBELL-WASLEY and CUMRANK-T*

If the event periods are non-clustered, the returns can be assumed cross-sectionally independent in particular in the event period, and thus the variance of the average cumulative ranks $\bar{U}_{\tau_1,\tau_2}$ defined in equation (23) reduces to var $[\bar{U}_{\tau_1,\tau_2}] = \tau(T-\tau)/(12(T+1)n)$ in equation (30). Thus, in order to test the null hypothesis of no event mean effect, which in terms of the ranks reduces to testing the hypothesis,

$$H_0 : \mu_{\tau_1,\tau_2} = \frac{1}{2}\tau, \tag{31}$$

and an appropriate $z$-ratio (called hereafter CUMRANK-Z) is

$$Z_1 = \frac{\bar{U}_{\tau_1,\tau_2} - \frac{1}{2}\tau}{\sqrt{\frac{\tau(T-\tau)}{12(T+1)n}}}. \tag{32}$$

This is the same statistic as $T_R^*$ proposed in Corrado and Truong (2008, p. 504) with non-scaled ranks.

**Remark 2.** *If the series are of different lengths such that there are $T^i$ observations available for series i, the average*

$$\overline{\text{var}}\,[\bar{U}_{\tau_1,\tau_2}; \bar{\rho}_{n,\tau_1,\tau_2} = 0] = \frac{1}{n}\sum_{i=1}^{n} \frac{\tau(T^i-\tau)}{12(T^i+1)n} \tag{33}$$

*is recommended for use in place of $\tau(T-\tau)/(12(T+1)n)$ in the denominator of* (32).

Even though the theoretical variance is known when the ranks are cross-sectionally independent, Corrado and Zivney (1992) propose estimating the variance for the event day average standardized rank $\bar{K}_t$ defined in equation (25) through the sample variance of the equally weighted portfolio

$$\tilde{s}_{\bar{K}}^2 = \widehat{\text{var}}[\bar{K}_t] = \frac{1}{T}\sum_{t=T_0+1}^{T_2} \frac{n_t}{n}\left(\bar{K}_t - \frac{1}{2}\right)^2, \tag{34}$$

where $T = T_2 - T_0$ is the combined length of the estimation period and the event period and $n_t$ is the number of observations in the mean $\bar{K}_t$ at time point $t$. As will be discussed later, an advantage of the sample estimator over the theoretical variance is that it is more robust than the theoretical variance to possible cross-sectional correlation of the returns. Cross-sectional correlation is in particular an issue when the event days are clustered.

The results of Kolari and Pynnönen (2010) show that just a small cross-correlation seriously biases the test results if not properly accounted for.

In terms of the estimator in (34), the variance of the cumulative ranks $\bar{U}_{\tau_1,\tau_2}$ is estimated in practice by simply ignoring the serial dependency between rank numbers and multiplying the single day rank variance by the number of the accumulated ranks such that

$$\tilde{s}^2_{\tau_1,\tau_2} = \frac{\tau}{T} \sum_{t=T_0+1}^{T_2} \frac{n_t}{n} \left( \bar{K}_t - \frac{1}{2} \right)^2 = \tau \tilde{s}^2_{\bar{K}}. \tag{35}$$

The implied $z$-ratio for testing the null hypothesis in (31) is

$$Z_2 = \frac{\bar{U}_{\tau_1,\tau_2} - \frac{1}{2}\tau}{\sqrt{\tau}\,\tilde{s}_{\bar{K}}}. \tag{36}$$

This statistic for testing CARs by the rank statistic is suggested in Campbell and Wasley (1993, p. 85), and it is called CAMPBELL-WASLEY. For a single day return the statistic reduces to the single period rank test suggested in Corrado (1989) and Corrado and Zivney (1992).

However, as will be demonstrated below, the autocorrelation between the ranks implies slight downward bias into the variance estimator $\tilde{s}^2_{\tau_1,\tau_2}$. The bias increases as the length, $\tau = \tau_2 - \tau_1 + 1$, of the period over which the ranks are accumulated, grows. Also, for fixed $T$ the asymptotic distributions of CUMRANK-Z and CAMPBELL-WASLEY (as well as Corrado's single period rank test) prove to be theoretically quite different. It is straightforward to show that the variance estimator $\tilde{s}^2_{\tau_1,\tau_2}$ in (35), utilized in the CAMPBELL-WASLEY statistic $Z_2$ in (36), is a biased estimator of the population variance $\mathrm{var}[\bar{U}_{\tau_1,\tau_2}]$ in equation (30). Assuming $n_t = n$ for all $t$, the bias can be computed, because $\mathrm{var}[\bar{K}_t] = E[(\bar{K}_t - 1/2)^2]$ such that

$$E\left[\tilde{s}^2_{\tau_1,\tau_2}\right] = \frac{\tau}{T} \sum_{t=T_0+1}^{T_1} E\left[(\bar{K}_t - 1/2)^2\right] = \frac{\tau}{T} \sum_{t=T_0+1}^{T_1} \mathrm{var}\left[\bar{K}_t\right].$$

Utilizing then equation (30) with $\tau_1 = \tau_2$ (in the equation), following proposition will be obtained:

**Proposition 3.** *Assuming $n_t = n$ for all $t = T_0 + 1, \ldots, T_1$, then under the null hypothesis of no event effects the expected value of $\tilde{s}^2_{\tau_1,\tau_2}$ defined in* (35) *is*

$$E[\tilde{s}^2_{\tau_1,\tau_2}] = \frac{\tau(T-1)}{12(T+1)n} (1 + (n-1)\bar{\rho}_n) \tag{37}$$

*and the bias is*

$$\begin{aligned}\text{Bias}\left[\tilde{s}^2_{\tau_1,\tau_2}\right] &= E\left[\tilde{s}^2_{\tau_1,\tau_2}\right] - \sigma^2_{\tau_1,\tau_2} \\ &= \frac{\tau(\tau-1)}{12(T+1)n}\left[1+(n-1)\bar{\rho}_n\right] \\ &\quad + \frac{\tau(T-\tau)}{12(T+1)n}\left\{1+(n-1)\left[\bar{\rho}_n - \bar{\rho}_{n,\tau_1,\tau_2}\right]\right\}, \end{aligned} \tag{38}$$

*where* $\bar{\rho}_n$ *is the average cross-correlation of the single day ranks* $K_{it}$ *and* $\bar{\rho}_{n,\tau_1,\tau_2}$ *is the average cross-correlation of* $\tau = \tau_2 - \tau_1 + 1$ *period cumulated ranks.*

In practice the average cross-correlation, $\bar{\rho}_n$, of the single day ranks and the average cross-correlation, $\bar{\rho}_{n,\tau_1,\tau_2}$ of $\tau$-period cumulated ranks is likely to be approximately the same, i.e., $\bar{\rho}_{n,\tau_1,\tau_2} \approx \bar{\rho}_n$, such that the bias reduces to

$$\text{Bias}\left[\tilde{s}^2_{\tau_1,\tau_2}\right] = \frac{\tau(\tau-1)}{12(T+1)n}\left[1+(n-1)\bar{\rho}_n\right]. \tag{39}$$

In this case the bias is easily fixed by multiplying $\tilde{s}^2_{\tau_1,\tau_2}$ defined in equation (35) by the factor $(T-\tau)/(T-1)$ yielding an estimator

$$\hat{s}^2_{\tau_1,\tau_2} = \frac{\tau(T-\tau)}{T(T-1)}\sum_{t=T_0+1}^{T_2}\frac{n_t}{n}\left(\bar{K}_t - \frac{1}{2}\right)^2 = \frac{T-\tau}{T-1}\tilde{s}^2_{\tau_1,\tau_2}. \tag{40}$$

Utilizing this correction gives a modification of the CAMPBELL-WASLEY statistic, such that

$$Z_3 = \frac{\bar{U}_{\tau_1,\tau_2} - \frac{1}{2}\tau}{\hat{s}_{\tau_1,\tau_2}} = \sqrt{\frac{T-1}{T-\tau}}Z_2. \tag{41}$$

Rather than using this, the small sample distributional properties (in terms of the number of time series observations, $T$) turn out to better by using the following modified statistic, which is called CUMRANK-T

$$Z_4 = Z_3\sqrt{\frac{T-2}{T-1-(Z_3)^2}}. \tag{42}$$

An advantage of the above CUMRANK-T statistic over the CUMRANK-Z statistic, $Z_1$, defined in equation (32), is its better robustness against cross-sectional correlation, because the variance estimator in equation (40), which is used in the denominator, implicitly accounts the possible cross-correlation. The downside, however, is loss of some power, which, however, proves to be small as the simulations with actual returns will demonstrate. For fixed $T$ the asymptotic distributions of these statistics are, however, different.

## 4.3 The Sign Tests for Testing Cumulative Abnormal Returns

### *4.3.1 The sign of the GSAR*

In order to account for the possible event-induced volatility Kolari and Pynnönen (2011) re-standardize the SCARs like BMP (1991) with the cross-sectional standard deviation to get re-standardized SCAR

$$\mathrm{SCAR}'_{i,\tau_1,\tau_2} = \frac{\mathrm{SCAR}_{i,\tau_1,\tau_2}}{S(\mathrm{SCAR}_{\tau_1,\tau_2})}, \tag{43}$$

where

$$S(\mathrm{SCAR}_{\tau_1,\tau_2}) = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(\mathrm{SCAR}_{i,\tau_1,\tau_2} - \overline{\mathrm{SCAR}}_{\tau_1,\tau_2})^2} \tag{44}$$

is the cross-sectional standard deviation of $\mathrm{SCAR}_{i,\tau_1,\tau_2}$s and

$$\overline{\mathrm{SCAR}}_{\tau_1,\tau_2} = \frac{1}{n}\sum_{i=1}^{n}\mathrm{SCAR}_{i,\tau_1,\tau_2}. \tag{45}$$

Again $\mathrm{SCAR}'_{i,\tau_1,\tau_2}$ is a zero mean and unit variance random variable. The generalized standardized abnormal returns (GSARs) are defined similar to Kolari and Pynnönen (2011):

**Definition 4.** *The generalized standardized abnormal return (GSAR) is defined as*

$$\mathrm{GSAR}_{it} = \begin{cases} \mathrm{SCAR}'_{i,\tau_1,\tau_2}, & \textit{in } \mathrm{CAR}\textit{-window} \\ \mathrm{SAR}_{it}, & \textit{otherwise}, \end{cases} \tag{46}$$

*where* $\mathrm{SCAR}'_{i,\tau_1,\tau_2}$ *is defined in equation (43) and* $\mathrm{SAR}_{it}$ *is defined in equation (2).*

Thus the CAR-window is considered as one time point in which the GSAR equals the re-standardized cumulative abnormal return defined in equation (43), and for other time points GSAR equals the usual standardized abnormal returns defined in equation

(2). The time indexing is redefined, similar as Kolari and Pynnönen (2011), such that the CAR-window of length $\tau_2 - \tau_1 + 1$ is squeezed into one observation with time index $t = 0$. Thus, considering the standardized cumulative abnormal return as one observation, in the testing procedure there are again $L_1 + 1$ observations of which the first $L_1$ are the estimation window abnormal returns and the last one is the cumulative abnormal return. Kolari and Pynnönen (2011) have suggested that the GSARs can be used to extend the sign test in Corrado and Zivney (1992) for testing CARs. This can be achieved by defining the sign of the GSAR like:

**Definition 5.** *The sign of the generalized standardized abnormal return* $\mathrm{GSAR}_{it}$ *is*

$$G_{it} = \mathrm{sign}[\mathrm{GSAR}_{it} - \mathrm{median}(\mathrm{GSAR}_i)], \tag{47}$$

*where* $\mathrm{sign}(x)$ *is equal to +1, 0, -1 as x is* $> 0$, $= 0$ *or* $< 0$.

If $T = L_1 + 1$ is even, the corresponding probabilities for the sign of the GSAR for values +1, 0 and -1 are

$$\Pr[G_{it} = 1] = \Pr[G_{it} = -1] = \frac{1}{2} \tag{48}$$

and

$$\Pr[G_{it} = 0] = 0. \tag{49}$$

If $T = L_1 + 1$ is odd, the corresponding probabilities for the sign of the GSAR for values +1, 0 and -1 are

$$\Pr[G_{it} = 1] = \Pr[G_{it} = -1] = \frac{T-1}{2T} \tag{50}$$

and

$$\Pr[G_{it} = 0] = \frac{1}{T}. \tag{51}$$

The expectations, variances and covariances of the sign of GSAR are presented in A Appendix for even and odd $T$, and summarized in Proposition 4.

**Proposition 4.** *The expectation for the sign of the* GSAR *defined in (47) is*

$$E[G_{it}] = 0 \tag{52}$$

*for T being even or odd. Furthermore the variance and covariance of the sign of the* GSAR *are*

$$\mathrm{var}[G_{it}] = \begin{cases} 1, & \textit{for even } T \\ \frac{T-1}{T}, & \textit{for odd } T \end{cases} \tag{53}$$

*and*

$$\mathrm{cov}[G_{it}, G_{is}] = \begin{cases} -\frac{1}{T-1}, & \textit{for even } T \\ -\frac{1}{T}, & \textit{for odd } T. \end{cases} \tag{54}$$

*Furthermore* $i = 1, \ldots, n$ *and* $t \neq s$.

### *4.3.2 The test statistics SIGN-GSAR-T and SIGN-GSAR-Z*

The null hypothesis of no mean event effect, reduces in terms of sign test statistics to

$$H_0 : \mu = 0, \tag{55}$$

where $\mu$ is the expectation of the (cumulative) abnormal return. Like Kolari and Pynnönen (2011) suggested, a new sign test statistic (called hereafter SIGN-GSAR-T) based on GSARs will be introduced and this test statistic can be used for testing the presented null hypothesis. The test statistic SIGN-GSAR-T is defined as

$$Z_5 = \frac{Z_6\sqrt{T-2}}{\sqrt{T-1-Z_6^2}}, \tag{56}$$

where

$$Z_6 = \frac{\overline{G_0}}{S(G)}, \tag{57}$$

with

$$S(G) = \sqrt{\frac{1}{T}\sum_{t\in\mathscr{T}} \frac{n_t}{n}\overline{G_t}^2} \tag{58}$$

and

$$\overline{G}_t = \frac{1}{n_t}\sum_{i=1}^{n_t} G_{it} \tag{59}$$

in which $n_t$ is the number of nonmissing returns in the cross-section of $n$-firms on day $t$ and $\mathscr{T} = \{T_0+1,\ldots,T_1,0\}$. The test statistic $Z_6$ in equation (57) is the sign test derived by Corrado and Zivney (1992) for testing single event-day abnormal returns.

Using facts about statistics based on signs (see A Appendix) and assuming $n_t = n$ for all $t \in \mathscr{T}$, it is easy to show that

$$\operatorname{var}[\overline{G_0}] = \begin{cases} \frac{1}{n}, & \text{for even } T \\ \frac{T-1}{nT} \approx \frac{1}{n}, & \text{for odd } T. \end{cases} \tag{60}$$

Thus, under the assumption that $\operatorname{var}[\overline{G_0}] = \frac{1}{n}$, another useful test statistic for the null hypothesis (55) is

$$Z_7 = \frac{\overline{G_0}}{\sqrt{\operatorname{var}[\overline{G_0}]}} = \overline{G_0}\sqrt{n}. \tag{61}$$

Henceforth this statistic is referred as SIGN-GSAR-Z.

The simplicity of the test statistic SIGN-GSAR-Z makes it an attractive alternative to the test statistic SIGN-GSAR-T. This is particularly the case when the event days across the sample firms are not clustered. However, as will be seen in the presence of event day clustering, which causes cross-sectional correlations between the returns, the SIGN-GSAR-T can be expected to be much more robust than the SIGN-GSAR-Z test statistic.

## 4.4 Asymptotic Distributions of the Rank and Sign Test Statistics

In event studies asymptotics can be dealt with both in time series and in cross-section dimensions. In the former the length of the estimation period is increasing while in the latter the number of firms is allowed to increase towards infinity. In most cases the interest is in the latter when the number of firms, $n$, is increasing. In the following this convention is adopted and it is assumed that all series in the sample have a fixed number $T$ time series observations such there are no missing returns. The asymptotic properties depend on, whether the event dates are clustered or non-clustered.

#### *4.4.1 Independent observations*

When the event days are non-clustered, such that the event period observations are cross-sectionally independent, the Central Limit Theorem (CLT) can be applied to get the following result.

**Theorem 2.** (Asymptotic normality of CUMRANK-Z): *If the even days are non-clustered such that the cumulative standardized ranks $U_{i,\tau_1,\tau_2}$ defined in* (20) *are independent and identically distributed random variables with mean $E[U_{i,\tau_1,\tau_2}] = \tau/2$ and variance $\mathrm{var}[U_{i,\tau_1,\tau_2}] = \tau(T-\tau)/(12(T+1))$, $i = 1,\ldots,n$, then under the null hypothesis of no event effect, as $n \to \infty$*

$$Z_1 = \frac{\bar{U}_{\tau_1,\tau_2} - \tau/2}{\sigma_{\tau_1,\tau_2}} \xrightarrow{d} N(0,1), \tag{62}$$

*where*

$$\sigma_{\tau_1,\tau_2} = \sqrt{\mathrm{var}[\bar{U}_{\tau_1,\tau_2}; \bar{\rho}_{n,\tau_1,\tau_2} = 0]} = \sqrt{\frac{\tau(T-\tau)}{12(T+1)n}}, \tag{63}$$

*$T_0 < \tau_1 \le \tau_2 \le T_2$, $T = T_2 - T_0$, $\tau = \tau_2 - \tau_1 + 1$, and "$\xrightarrow{d}$" denotes convergence in distribution.*

Similarly CLT can be applied to get following result.

**Theorem 3.** (Asymptotic normality of SIGN-GSAR-Z): *If the event days are again non-clustered such that the signs $G_{i0}$ defined in* (47) *(with t=0, the signs of the event days) are independent and identically distributed random variables with mean $E[G_{i0}] = 0$ and variance $\mathrm{var}[G_{i0}] \approx 1$, $i = 1,\ldots,n$, then under the null hypothesis of no event effect, as $n \to \infty$*

$$Z_7 = \frac{\overline{G_0}}{\sqrt{\mathrm{var}[\overline{G_0}]}} = \overline{G_0}\sqrt{n} \xrightarrow{d} N(0,1), \tag{64}$$

*where $\mathrm{var}[\overline{G_0}] \approx \frac{1}{n}$ and "$\xrightarrow{d}$" denotes convergence in distribution.*

Proofs of the following theorems (Theorem 4 and Theorem 5) regarding the asymptotic distributions of $Z_3$ (modified CAMPBELL-WASLEY), $Z_6$, $Z_4$ (CUMRANK-T) and $Z_5$ (SIGN-GSAR-T) defined in equations (41), (57), (42) and (56), respectively, are presented in B Appendix.

**Theorem 4.** (Asymptotic distribution of the modified CAMPBELL-WASLEY and $Z_6$): *For a fixed $T$, under the assumption of cross-sectional independence, the density function of the asymptotic distribution of the modified CAMPBELL-WASLEY statistic $Z_3$ defined in equation (41) and $Z_6$ statistic defined in equation (57) when $n \to \infty$ are*

$$f_{Z_3}(z) = f_{Z_6}(z) = \frac{\Gamma[(T-1)/2]}{\Gamma[(T-2)/2]\sqrt{(T-1)\pi}} \left(1 - \frac{z^2}{T-1}\right)^{\frac{1}{2}(T-2)-1}, \tag{65}$$

*for $|z| \leq \sqrt{T-1}$ and zero elsewhere, where $\Gamma(\cdot)$ is the Gamma function.*

Thus, Theorem 4 implies that $(Z_3)^2/(T-1)$ and $(Z_6)^2/(T-1)$ are asymptotically Beta distributed with parameters $1/2$ and $(T-2)/2$. Corrado and Zivney (1992) conjecture that for sufficiently large sample size the Central Limit Theorem implies that the distribution of $Z_6$ should converge to normality. By Theorem 4 we can conclude that the asymptotic normality of $Z_6$ and the modified CAMPBELL-WASLEY $(Z_3)$ hold only if also $T$ is large enough. This follows from the fact that in equation (65)

$$\left(1 - \frac{z^2}{T-1}\right)^{\frac{1}{2}(T-2)-1} \to e^{-\frac{1}{2}z^2} \tag{66}$$

and the normalizing constant

$$\frac{\Gamma[(T-1)/2]}{\Gamma[(T-2)/2]\sqrt{(T-1)\pi}} \to 1/\sqrt{2\pi} \tag{67}$$

as $T \to \infty$, implying the limiting $N(0,1)$-distribution.

The distribution of test statistic CAMPBELL-WASLEY ($Z_2$) defined in (36) is obtained via the transformation in equation (41).

**Theorem 5.** (Asymptotic distribution of CUMRANK-T and SIGN-GSAR-T): *Under the assumptions of Theorem* 4,

$$Z_4 = Z_3\sqrt{\frac{T-2}{T-1-(Z_3)^2}} \xrightarrow{d} t_{T-2} \tag{68}$$

*and*

$$Z_5 = Z_6\sqrt{\frac{T-2}{T-1-(Z_6)^2}} \xrightarrow{d} t_{T-2}, \tag{69}$$

*as $n \to \infty$, where $Z_3$ is defined in equation (41), $Z_6$ is defined in equation (57), and $t_{T-2}$*

*denotes the Student t-distribution with $T-2$ degrees of freedom.*

Given that the $t$-distribution approaches the $N(0,1)$-distribution as the degrees of freedom $T-2$ increases, also the null distributions of the test statistics CUMRANK-T and SIGN-GSAR-T approach the standard normal distribution as $T\to\infty$.

#### 4.4.2 Cross-sectional dependence (clustered event dates)

Cross-sectional dependence due to clustered event days (the same event days across the firms) changes materially the asymptotic properties of the test statistics and in particular those statistics that do not account for the cross-sectional dependence.

As stated in Lehmann (1999, Sec. 2.8), it is still, however, frequently true that the asymptotic normality holds provided that the average cross-correlation, $\overline{\rho}_n$, tends to zero rapidly enough such that

$$\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1,i\neq j}^{n}\rho_{ij}\to\gamma \tag{70}$$

as $n\to\infty$. In financial applications this would be the case if there are a finite number of firms in each industry and the return correlations between industries are zeros. In fact this is a special case of so called $m$-independence. Generally, a sequence of random variables $X_1$, $X_2$,..., is said to be $m$-independent, if $X_i$ and $X_j$ are independent if $|i-j|>m$. In cross-sectional analysis this would mean that the variables can be ordered such that when the index difference is larger than $m$, then the variables are independent. [See Kolari and Pynnönen (2011)].

It can be shown in the same manner as in Kolari and Pynnönen (2011) that the result (70) holds. In the case of rank test statistics it is assumed that for any fixed $t$, $U_{i,\tau_1,\tau_2}$ defined in equation (24) are $m$-independent, $i=1,2,\dots,n$, $(n>m)$ then the correlation matrix of $U_{1,\tau_1,\tau_2},\dots,U_{n,\tau_1,\tau_2}$ is band-diagonal such that all $\rho_{ij}$ with $|i-j|>m$ are zeros. Similarly, in the case of sign test statistics it is assumed that for any fixed $t$, $G_{it}$ defined in equation (47) are $m$-independent, $i=1,2,\dots,n$ $(n>m)$, then the correlation matrix of $G_{it},\dots,G_{nt}$ is band-diagonal such that all $\rho_{ij}$ with $|i-j|>m$ are zeros. Then it is straightforward to see that in these kinds of correlation matrixes there are $m(2n-m-1)$ nonzero correlations in addition to the $n$ ones on the diagonal. Thus, in the double summation (70) there are $m(2n-m-1)$ non-zero elements, and it can be written such that

$$\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n}\rho_{ij} = \frac{m(2n-m-1)}{n}\widetilde{\rho}_n \to \gamma, \tag{71}$$

where $\widetilde{\rho}_n$ is the average of the $m(2n-m-1)$ cross-correlations in the band-diagonal correlations matrix and $\gamma = 2m\widetilde{\rho}$ is a finite constant with $\widetilde{\rho} = \lim_{n\to\infty}\widetilde{\rho}_n$ and $2m = \lim_{n\to\infty} m(2n-m-1)/n$.

Thus, under the $m$-independence the asymptotic distribution of the test statistic CUMRANK-Z ($Z_4$) and SIGN-GSAR-Z ($Z_7$) are

$$Z_4 \to N(0, 1+\gamma) \tag{72}$$

and

$$Z_7 \to N(0, 1+\gamma). \tag{73}$$

This implies that the test statistics CUMRANK-Z and SIGN-GSAR-Z are not robust to cross-sectional correlation of the return series. Typically, $\gamma > 0$, which means that these test statistics will tend to over-reject the null hypothesis.

The limiting distributions of the test statistics CAMPBELL-WASLEY ($Z_2$), modified CAMPBELL-WASLEY ($Z_3$), CUMRANK-T ($Z_4$) and SIGN-GSAR-T ($Z_5$) turn out to apply also under $m$-independence. This follows from the fact that if the asymptotic normality holds under the $m$-independence such that the limiting correlation effect is $1+\gamma$, then using the scaled variables, $(K_{it} - 1/2)/\sqrt{1+\gamma}$ and $G_{it}/\sqrt{1+\gamma}$, in place of the original variables, all the results in Theorem 4 and Theorem 5 follow, because $Z_3$, $Z_4$, $Z_5$ and $Z_6$ are invariant to scaling of the observations. Hence, the theoretical derivations indicate that when the event-dates are clustered, the test statistics CAMPBELL-WASLEY, CUMRANK-T and SIGN-GSAR-T behave better than the test statistics CUMRANK-Z and SIGN-GSAR-Z.

# 5 THE SIMULATION DESIGN

In the empirical simulations the well-known simulation approach presented by Brown and Warner (1980) is adopted. This approach is also widely used in several other methodological studies [e.g. Brown and Warner (1985), Corrado (1989), Cowan (1992), Campbell and Wasley (1993), and Cowan and Sergeant (1996)]. Like Corrado (2010) has concluded Brown and Warner -type studies are categorically distinguished by their use of computer simulation experiments based on random drawings from a large population of actual security returns data. An alternative approach could be the analytical approach where the data is generated by the process modeled. Nevertheless, Brown and Warner (1980) have concluded that even if it were possible to analytically derive and compare the properties of alternative methods for measuring abnormal performance in event studies, conclusions from the comparison would not necessarily be valid if the actual data used in event studies were generated by a process which differed from that which the comparison assumed. Also Dyckman, Philbrick and Stephan (1984) have concluded that simulations provide a tractable means of dealing with simulations, where an analytical approach may yield results suggesting direction but not magnitude or where such techniques are unusually cumbersome. Furthermore, because it is desirable that the results of this study can be compared with most other studies, the simulation study is conducted to investigate the empirical behaviour of the new test statistics presented in Section 4 and to compare these test statistics with other widely known test statistics.

As Kolari and Pynnönen (2011) for example, have concluded, the optimality of a test can be judged on the basis of its size and power. Within a class of tests of given size (Type I error probability), the one that has the maximum power (minimum Type II error probability) is the best. A testing procedure is robust if the Type I error rate is not affected by real data issues such as non-normality, event-induced volatility, autocorrelation and cross-correlation of returns. Consequently, the aim of the simulations is to focus on the robustness and power properties of the tests. Non-normality, autocorrelation, and other data issues are captured in the simulation by using actual return data instead of artificially simulated data. Event-induced volatility effects are investigated by introducing volatility change within the event period, and the effect of cross-sectional correlation is examined by setting the same event day in the return series for each firm in the sample.

In Section 5 first the sample constructions are described. Second, the abnormal return model is presented. Third, some widely known parametric and nonparametric test statistics are presented and fourth the data of the simulation study is described.

## 5.1 Sample Constructions

The simulation study is conducted by selecting 1,000 samples of $n = 50$ return series with replacement from the data base. Each time a security is selected, a hypothetical event day is generated. The events are assumed to occur with equal probability on each trading day. The event day is denoted as day 0. The results are reported for event day $t = 0$ abnormal return AR(0), together with cumulative abnormal returns $\text{CAR}(-1,+1)$, $\text{CAR}(-5,+5)$ and $\text{CAR}(-10,+10)$. The estimation window is normally comprised of 239 days (the days from -249 to -11) and the event window is comprised of 21 days (the days -10 to +10). Therefore, the estimation window and the event window altogether consist of 260 days, which is close to one years trading days. For a security to be included in the sample there should be no missing return data in the last 30 days, that is, in days -19 to +10.

In earlier studies [e.g. Charest (1978), Mikkelson (1981), Penman (1982), and Rosenstein and Wyatt (1990)] it has been found that the event window standard deviation is about 1.2 to 1.5 times the estimation window standard deviation. Therefore, the increased volatility is introduced by multiplying the cumulated event window returns by a factor $\sqrt{c}$ with values $c = 1.5$ for an approximate 20 percent increased volatility, $c = 2.0$ for an approximate 40 percent increased volatility and $c = 3.0$ for an approximate 70 percent increased volatility due to the event effect.[4] To add realism the volatility factors $c$ are generated for each stock based on the following uniform distributions $U[1,2]$, $U[1.5,2.5]$ and $U[2.5,3.5]$. This generates on average the variance effects of 1.5, 2.0 and 3.0. Furthermore for the no volatility effect experiment $c = 1.0$ is fixed.

A similar method for investigating the power properties is used as is used for example in Campbell and Wasley (1993). Therefore, for single-day event period [AR(0)] the abnormal performance is artificially introduced by adding the indicated percent (a constant) to the day-0 return of each security. While, in the multiday setting [$\text{CAR}(-1,+1)$, $\text{CAR}(-5,+5)$ and $\text{CAR}(-10,+10)$], abnormal performance is introduced by selecting one day of the CAR-window at random and adding the particular level of abnormal

[4] Because $\sqrt{1.5} \approx 1.2$, $\sqrt{2.0} \approx 1.4$ and $\sqrt{3.0} \approx 1.7$.

performance to that day's return. By this the aim is to mimic the real situations, where there can be the information leakage and delayed adjustment. That is, if the markets are inefficient, information may leak before the event which shows up as abnormal behaviour before the event day. Delays in the event information processing show up as abnormal return behavior after the event day.

Also the effect of event-date clustering on the test statistics is studied. The effect of event-date clustering is examined by constructing 1,000 portfolios each of 50 stocks again from the data base, but all stocks in the portfolio have exactly the same event date.

## 5.2 Abnormal Return Model

The abnormal behavior of security returns can be estimated for example through the market model

$$r_{it} = \alpha_i + \beta_i r_{mt} + \varepsilon_{it}, \tag{74}$$

where again $r_{it}$ is the return of stock $i$ at time $t$, $r_{mt}$ is the market index return at time $t$ and $\varepsilon_{it}$ is a white noise random component, which is not correlated with $r_{mt}$. The resulting abnormal returns are obtained as differences of realized and predicted returns on day $t$ in the event window

$$\mathrm{AR}_{it} = r_{it} - (\hat{\alpha}_i + \hat{\beta}_i r_{mt}), \tag{75}$$

where the parameters $\hat{\alpha}_i$ and $\hat{\beta}_i$ are estimated from the estimation window with ordinary least squares. Hence, the abnormal returns are calculated by first estimating the parameters $\hat{\alpha}_i$ and $\hat{\beta}_i$ from the estimation window with ordinary least squares. Then the abnormal returns for the event window are achieved by setting the event window returns of stock $i$ and the event window market index returns to the equation (75) together with the estimated parameters $\hat{\alpha}_i$ and $\hat{\beta}_i$.

According to Campbell, Lo and MacKinlay (1997) the market model represents a potential improvement over the traditional constant-mean-return model, because by removing the portion of the return that is related to variation in the market's return, the variance of the abnormal return is reduced. This can lead to increased ability to detect event effects. A number of other statistical models have been proposed for modeling the normal return. A common type of statistical model is the factor model. The market model is an example of an one-factor model, but in a multifactor model one might for example include industry indexes in addition to the market. Nevertheless, Campbell, Lo and MacKinlay (1997) concludes that in practice the gains from employing multifactor models for event studies are limited.

## 5.3 Test Statistics

In Section 6 the nonparametric test statistics presented in Section 4 are compared with some widely-known parametric and nonparametric tests statistics. Therefore, those widely-known test statistics are briefly presented in this section.

### *5.3.1 Parametric test statistics*

The ordinary $t$-test (ORDIN) is defined as

$$Z_{\text{ORDIN}} = \frac{\overline{\text{CAR}}_{\tau_1,\tau_2}}{S(\overline{\text{CAR}}_{\tau_1,\tau_2})}, \tag{76}$$

where

$$\overline{\text{CAR}}_{\tau_1,\tau_2} = \frac{1}{n}\sum_{i=1}^{n} \text{CAR}_{i,\tau_1,\tau_2}, \tag{77}$$

in which $\text{CAR}_{i,\tau_1,\tau_2}$ is defined in equation $(3)$ and $S(\overline{\text{CAR}}_{\tau_1,\tau_2})$ is the standard error of the average cumulative abnormal return $\overline{\text{CAR}}_{\tau_1,\tau_2}$ adjusted for the prediction error [see again Campbell, Lo and MacKinlay (1997, Section 4.4.3)]. The ordinary $t$-test statistic is asymptotically $N(0,1)$-distributed under the null hypothesis of no event effect.

Patell (1976) test statistic (PATELL) is

$$Z_{\text{PATELL}} = \sqrt{\frac{n(L_1-4)}{L_1-2}}\,\overline{\text{SCAR}}_{\tau_1,\tau_2}, \tag{78}$$

where $\overline{\text{SCAR}}_{\tau_1,\tau_2}$ is the average of the SCAR defined in equation $(4)$, and $L_1$ is again the length of the estimation window. Also the test statistic derived by Patell is asymptotically $N(0,1)$-distributed under the null hypothesis of no event effect.

The Boehmer, Musumeci and Poulsen (1991) test statistic (BMP) is

$$Z_{\text{BMP}} = \frac{\overline{\text{SCAR}}_{\tau_1,\tau_2}\sqrt{n}}{S(\text{SCAR}_{\tau_1,\tau_2})}, \tag{79}$$

where again $S(\text{SCAR}_{\tau_1,\tau_2})$ is the cross-sectional standard deviation of SCARs defined in (44), and $\overline{\text{SCAR}}_{\tau_1,\tau_2}$ is defined in equation (4). Also the test statistic $Z_{\text{BMP}}$ is asymptotically $N(0,1)$-distributed under the null hypothesis of no event effect.

### *5.3.2 Nonparametric test statistics*

Like Kolari and Pynnönen (2011) defined the demeaned standardized abnormal ranks of the GSARs, those are also here defined as

$$V_{it} = \text{Rank}(\text{GSAR}_{it})/(T+1) - 1/2, \tag{80}$$

where $i = 1,\ldots,n$. Furthermore, $t \in \mathscr{T} = \{T_0+1,\ldots,T_1,0\}$ is the set of time indexes including the estimation window for $t = T_0+1,\ldots,T_1$ and to the cumulative abnormal return for $t = 0$, with $T_0+1$ and $T_1$ being the first and last observation on the estimation window. Hence, the total number of observations is $T = L_1+1 = T_1-T_0+1$. Then the test statistic GRANK is defined as

$$Z_{\text{GRANK}} = Z_8\sqrt{\frac{T-2}{T-1-Z_8^2}}, \tag{81}$$

where

$$Z_8 = \frac{\overline{V}_0}{S_{\overline{V}}} \tag{82}$$

with

$$S_{\overline{V}} = \sqrt{\frac{1}{T}\sum_{t\in\mathscr{T}}\frac{n_t}{n}\overline{V}_t^2} \tag{83}$$

and

$$\overline{V}_t = \frac{1}{n_t}\sum_{i=1}^{n_t} V_{it}. \tag{84}$$

Furthermore $n_t$ is the number of valid GSARs available at time point $t$ and $\overline{V}_0$ is the mean $\overline{V}_t$ for $t = 0$ (cumulative abnormal return). According to Kolari and Pynnönen (2011) the asymptotic distribution of the test statistic GRANK is Student $t$-distribution with $T-2$ degrees of freedom. Again given that the $t$-distribution approaches the $N(0,1)$-distribution as the degrees of freedom $T-2$ increases, also the null hypothesis of the test statistic $Z_{\text{GRANK}}$ approach the standard normal distribution as $T \to \infty$.

The generalized sign test statistic presented by Cowan (1992) (SIGN-COWAN) is

$$Z_{\text{COWAN}} = \frac{w - n\widehat{p}}{\sqrt{n\widehat{p}(1-\widehat{p})}}, \tag{85}$$

where $w$ is the number of stocks in the event window for which the cumulative abnormal return is positive and $n$ is again the number of stocks. Furthermore

$$\widehat{p} = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{m_i}\sum_{t=T_0+1}^{T_1} S_{it}, \tag{86}$$

where $m_i$ is the number of non-missing returns in the estimation window for security-event $i$ and

$$S_{it} = \begin{cases} 1 & \text{if } \mathrm{AR}_{it} > 0 \\ 0 & \text{otherwise.} \end{cases} \tag{87}$$

According to Cowan (1992) the test statistic SIGN-COWAN is asymptotically $N(0,1)$-distributed under the null hypothesis.

## 5.4 The Data

The data used consists of randomly drawn observations from U.S. stock market returns. Hence, a parametric model as a data generating process is not used, which guarantees that the simulated data have the typical properties of the actual financial data. By using actual (rather than artificial) stock returns in repeated simulations, a reliable and realistic view about the comparative real data performance of the test statistics in true applications is attained. Of course this procedure excludes the option of examining how the results would change with normally distributed data, for example. Therefore, while information is obtained on the behavior of the tests with financial data, the results may not always apply to some other fields where event studies can be applied.

In other words, the data in this simulation design consists of daily closing prices of 1,500 U.S. traded stocks that make up the S&P 400, S&P 500 and S&P 600 indexes.

S&P 400 covers the mid-cap range of stocks, S&P 500 the large-cap range of stocks and S&P 600 the small-cap range of stocks. The five percent of the stocks having the smallest trading volume are excluded. Hence, 72 stocks from S&P 600, two stocks from S&P 400 and one stock from S&P 500 are excluded. In the simulations the S&P 500 index stands for the market index. The sample period spans from the beginning of July, 1991 to the end of October, 2009. S&P 400 index was launched in June in 1991, which is the reason why the sample period starts in the beginning of July, 1991. Official holidays and observance days are excluded from the data.

The returns are defined as log-returns

$$r_{it} = \log(P_{it}) - \log(P_{it-1}), \tag{88}$$

where $P_{it}$ is the closing price for stock $i$ at time $t$.

# 6 THE SIMULATION RESULTS

This section discusses the results from the simulation study. First, the sample statistics of the ARs, the CARs and the test statistics are presented. Second, the properties of the empirical distributions of the test statistics are presented. The properties of the empirical distributions are presented as well in the cases where the estimation period is shortened. Third, the rejection rates are reported. The rejection rates are also investigated in the situations, where the event induced volatility is present and in the cases where the estimation period is shortened. Fourth, the power properties of the test statistics are presented. The power properties are also reported in the cases where the event days are clustered.

## 6.1 Sample Statistics

Table 1 reports sample statistics from 1,000 simulations for the event day abnormal returns, AR$(0)$, and for the cumulative abnormal returns in cases: CAR$(-1,+1)$, CAR$(-5,+5)$ and CAR$(-10,+10)$. It also reports the sample statistics for the test statistics for AR$(0)$, CAR$(-1,+1)$, CAR$(-5,+5)$ and CAR$(-10,+10)$. Under the null hypothesis of no event effect all the test statistics, except CUMRANK-T, GRANK and SIGN-GSAR-T, should be approximately $N(0,1)$-distributed. Strictly speaking, the asymptotic distribution of CUMRANK-T, GRANK and SIGN-GSAR-T should be $t$-distribution with $T-2$ degrees of freedom. However, with $T-2$ equal to 258, for CUMRANK-T, and 238, for GRANK and SIGN-GSAR-T, the normal approximation should be valid, and so the null distribution of these test statistics approach as well the standard normal distribution. Hence, it can be concluded that under the null hypothesis of no event effect, all the test statistics should have zero mean and (approximately) unit variance.

Considering only the single abnormal returns, AR$(0)$, in Panel A of Table 1, it can be noticed that means of all test statistics are statistically close to zero. For example (in absolute value) the largest mean of $-0.024$ for the PATELL statistic is only 1.113 standard errors away from zero. Furthermore, the standard deviations of the test statistics are close to one in Panel A of Table 1.

In longer event windows the means of the test statistics, albeit small, start to deviate significantly away from the theoretical value of zero. Considering the 3-day cumulative abnormal returns, CAR$(-1,+1)$, in Panel B of Table 1, it can be seen that only the means for the test statistics PATELL and BMP deviate significantly away from zero. While, considering the 11- and 21-day cumulative abnormal returns, CAR$(-5,+5)$ and CAR$(-10,+10)$, in Panels C and D of Table 1, it is noticable that means for almost all the test statistics deviate significantly away from the theoretical value. Nonetheless, it can be seen that the means of the test statistics PATELL and BMP deviate more rapidly

**Table 1.** Sample statistics in event tests for 1,000 random portfolios on $n = 50$ securities belonging to S&P400-, S&P500- and S&P600-indexes. Superscripts $a$, $b$ and $c$ in the second column correspond to the significance levels 0.10, 0.05 and 0.01.

| Test statistics | | | | | | | |
|---|---|---|---|---|---|---|---|
| Panel A: AR(0) | Mean | Med. | Std. | Skew. | Kurt. | Min. | Max. |
| Abnormal return, % | 0.004 | -0.008 | 0.413 | -0.082 | 1.018 | -1.688 | 1.641 |
| ORDIN | 0.008 | -0.019 | 1.053 | -0.079 | 0.701 | -3.878 | 3.694 |
| PATELL | -0.024 | -0.036 | 1.113 | -0.193 | 1.170 | -6.178 | 3.837 |
| BMP | -0.013 | -0.033 | 1.000 | -0.013 | 0.146 | -4.000 | 3.777 |
| CAMPBELL-WASLEY | 0.000 | -0.015 | 0.956 | 0.037 | 0.036 | -3.480 | 3.322 |
| CUMRANK-T | 0.000 | -0.015 | 0.959 | 0.038 | 0.084 | -3.558 | 3.388 |
| CUMRANK-Z | 0.001 | -0.015 | 0.979 | 0.044 | 0.145 | -3.811 | 3.576 |
| GRANK | 0.002 | -0.010 | 0.974 | 0.056 | 0.071 | -3.518 | 3.375 |
| SIGN-COWAN | -0.002 | -0.042 | 0.958 | 0.059 | -0.120 | -3.475 | 2.999 |
| SIGN-GSAR-T | -0.016 | 0.000 | 0.990 | 0.041 | -0.206 | -2.630 | 2.997 |
| SIGN-GSAR-Z | -0.016 | 0.000 | 1.082 | 0.037 | -0.226 | -2.828 | 3.111 |
| Panel B: CAR$(-1,+1)$ | Mean | Med. | Std. | Skew. | Kurt. | Min. | Max. |
| CAR$(-1,+1)$, % | -0.010 | -0.029 | 0.671 | -0.019 | 0.146 | -2.288 | 2.096 |
| ORDIN | -0.018 | -0.040 | 0.988 | -0.028 | 0.306 | -3.759 | 3.329 |
| PATELL | -0.067[b] | -0.085 | 1.077 | 0.133 | 0.113 | -3.380 | 4.059 |
| BMP | -0.054[a] | -0.088 | 1.023 | 0.159 | 0.083 | -3.208 | 3.856 |
| CAMPBELL-WASLEY | -0.010 | 0.003 | 0.961 | 0.072 | -0.053 | -2.760 | 3.138 |
| CUMRANK-T | -0.010 | 0.003 | 0.696 | 0.075 | -0.013 | -2.808 | 3.206 |
| CUMRANK-Z | -0.009 | 0.003 | 0.987 | 0.088 | -0.013 | -2.884 | 3.177 |
| GRANK | -0.001 | 0.011 | 1.021 | 0.088 | 0.189 | -3.332 | 3.963 |
| SIGN-COWAN | 0.042 | 0.108 | 1.020 | 0.063 | 0.127 | -3.475 | 3.832 |
| SIGN-GSAR-T | 0.028 | 0.000 | 1.036 | 0.050 | 0.082 | -3.225 | 3.610 |
| SIGN-GSAR-Z | 0.031 | 0.000 | 1.128 | 0.038 | 0.045 | -3.677 | 3.960 |

(*continues*)

Table 1 (continues).

| Test statistics | | | | | | | |
|---|---|---|---|---|---|---|---|
| Panel C: $CAR(-5,+5)$ | Mean | Med. | Std. | Skew. | Kurt. | Min. | Max. |
| CAR(-5,+5), % | -0.076 | -0.027 | 1.269 | -0.114 | 0.346 | -5.178 | 4.455 |
| ORDIN | -0.060[b] | -0.020 | 0.959 | -0.183 | 0.433 | -3.977 | 3.441 |
| PATELL | -0.132[c] | -0.108 | 1.107 | -0.034 | 0.363 | -4.005 | 3.992 |
| BMP | -0.113[c] | -0.117 | 1.036 | 0.088 | 0.157 | -3.417 | 3.603 |
| CAMPBELL-WASLEY | -0.067[b] | -0.094 | 0.922 | 0.144 | 0.116 | -2.893 | 3.158 |
| CUMRANK-T | -0.068[b] | -0.096 | 0.944 | 0.148 | 0.165 | -2.996 | 3.281 |
| CUMRANK-Z | -0.067[b] | -0.094 | 0.959 | 0.170 | 0.144 | -3.020 | 3.328 |
| GRANK | 0.016 | 0.062 | 1.038 | 0.011 | 0.113 | -3.073 | 3.334 |
| SIGN-COWAN | 0.085[b] | 0.102 | 0.973 | -0.028 | -0.081 | -2.764 | 3.334 |
| SIGN-GSAR-T | 0.065[b] | 0.000 | 0.993 | 0.018 | -0.005 | -2.804 | 3.284 |
| SIGN-GSAR-Z | 0.071[b] | 0.000 | 1.084 | 0.030 | 0.034 | -3.111 | 3.677 |
| Panel D: $CAR(-10,+10)$ | Mean | Med. | Std. | Skew. | Kurt. | Min. | Max. |
| CAR(-10,+10), % | -0.056 | -0.029 | 1.800 | -0.136 | 0.018 | -5.442 | 4.845 |
| ORDIN | -0.038 | -0.015 | 0.967 | -0.225 | 0.149 | -3.332 | 2.749 |
| PATELL | -0.130[c] | -0.108 | 1.105 | -0.287 | 0.852 | -5.208 | 4.129 |
| BMP | -0.100[c] | -0.117 | 1.042 | 0.011 | 0.033 | -3.092 | 3.759 |
| CAMPBELL-WASLEY | -0.048[a] | -0.044 | 0.866 | 0.008 | -0.071 | -2.450 | 3.035 |
| CUMRANK-T | -0.050[a] | -0.046 | 0.904 | 0.009 | -0.039 | -2.578 | 3.216 |
| CUMRANK-Z | -0.048[a] | -0.046 | 0.923 | 0.003 | -0.033 | -2.895 | 3.061 |
| GRANK | 0.071[b] | 0.063 | 1.053 | -0.049 | 0.282 | -3.645 | 3.942 |
| SIGN-COWAN | 0.180[c] | 0.162 | 1.013 | -0.052 | 0.295 | -3.296 | 3.536 |
| SIGN-GSAR-T | 0.148[c] | 0.241 | 0.997 | -0.059 | 0.370 | -3.275 | 3.423 |
| SIGN-GSAR-Z | 0.158[c] | 0.283 | 1.090 | -0.088 | 0.413 | -3.677 | 3.677 |

and clearly from the theoretical value of zero than means of the other test statistics. It can also be seen that the mean of the test statistic GRANK seems to deviate more slowly from the theoretical value of zero than the means of the other test statistics. Importantly, again all standard deviations of the test statistics are close to their theoretical values of unity.

## 6.2 Empirical Distributions

Table 2 reports Cramer-von Mises normality tests for ORDIN, PATELL, BMP, CAMPBELL-WASLEY, CUMRANK-Z, SIGN-COWAN and SIGN-GSAR-Z, and Cramer-von Mises tests of CUMRANK-T, GRANK and SIGN-GSAR-T against a $t$-distribution with appropriate degrees of freedom, depending on the length of the estimation window. Panel D of Table 2 deals the case where the estimation window consists of 239 days, and Panels A to C report the results in the cases where the estimation window is shortened consisting only of 25, 50 and 100 days.

**Table 2.** Cramer-von Mises tests of the distributions. Superscripts *a* and *b* correspond to the significance levels 0.05 and 0.01.

| Panel A: $L_1$=25 | AR(0) | CAR(-1,+1) | CAR(-5,+5) | CAR(-10,+10) |
|---|---|---|---|---|
| ORDIN | 0.441 | 0.393 | 1.090[b] | 2.194[b] |
| PATELL | 0.626[a] | 0.280 | 0.138 | 0.433 |
| BMP | 0.027 | 0.189 | 0.188 | 0.046 |
| CAMPBELL-WASLEY | 0.304 | 0.168 | 1.183[b] | 4.995[b] |
| CUMRANK-T | 0.374 | 0.125 | 0.289 | 0.710[b] |
| CUMRANK-Z | 0.371 | 0.124 | 0.282 | 0.619[a] |
| GRANK | 0.311 | 0.084 | 0.194 | 0.463[a] |
| SIGN-COWAN | 0.334 | 0.212 | 1.294[b] | 1.734[b] |
| SIGN-GSAR-T | 0.250 | 0.210 | 1.243[b] | 1.140[b] |
| SIGN-GSAR-Z | 0.986[b] | 0.848[b] | 2.044[b] | 2.010[b] |
| Panel B: $L_1$=50 | AR(0) | CAR(-1,+1) | CAR(-5,+5) | CAR(-10,+10) |
| ORDIN | 0.097 | 0.393 | 4.365[b] | 0.743[b] |
| PATELL | 0.307 | 0.280 | 0.332 | 0.208 |
| BMP | 0.060 | 0.189 | 0.288 | 0.081 |
| CAMPBELL-WASLEY | 0.140 | 0.086 | 0.851[b] | 2.213[b] |
| CUMRANK-T | 0.164 | 0.070 | 0.369 | 0.423 |
| CUMRANK-Z | 0.153 | 0.145 | 0.382 | 0.387 |
| GRANK | 0.125 | 0.038 | 0.177 | 0.624[a] |
| SIGN-COWAN | 0.357 | 0.164 | 0.826[b] | 1.828[b] |
| SIGN-GSAR-T | 0.097 | 0.116 | 0.612[a] | 1.205[b] |
| SIGN-GSAR-Z | 0.533[a] | 0.659[a] | 1.191[b] | 1.941[b] |

(*continues*)

Table 2 (continues).

| Panel C: $L_1$=100 | AR(0) | CAR(-1,+1) | CAR(-5,+5) | CAR(-10,+10) |
|---|---|---|---|---|
| ORDIN | 0.059 | 0.166 | 0.355 | 0.387 |
| PATELL | 0.109 | 0.527[a] | 0.686[a] | 0.450 |
| BMP | 0.037 | 0.434 | 0.575[a] | 0.319 |
| CAMPBELL-WASLEY | 0.096 | 0.090 | 0.781[b] | 1.368[b] |
| CUMRANK-T | 0.110 | 0.084 | 0.603[a] | 0.581[a] |
| CUMRANK-Z | 0.100 | 0.077 | 0.591[a] | 0.497[a] |
| GRANK | 0.105 | 0.036 | 0.167 | 0.657[a] |
| SIGN-COWAN | 0.207 | 0.090 | 0.961[b] | 3.011[b] |
| SIGN-GSAR-T | 0.179 | 0.124 | 0.747[b] | 2.450[b] |
| SIGN-GSAR-Z | 0.772[b] | 0.582[a] | 1.214[b] | 3.080[b] |
| Panel D: $L_1$=239 | AR(0) | CAR(-1,+1) | CAR(-5,+5) | CAR(-10,+10) |
| ORDIN | 0.054 | 0.218 | 0.350 | 0.196 |
| PATELL | 0.164 | 0.795[b] | 1.488[b] | 1.104[b] |
| BMP | 0.066 | 0.625 | 1.277[b] | 0.985[b] |
| CAMPBELL-WASLEY | 0.085 | 0.092 | 0.849[b] | 0.919[b] |
| CUMRANK-T | 0.089 | 0.088 | 0.783[b] | 0.629[a] |
| CUMRANK-Z | 0.054 | 0.051 | 0.725[a] | 0.487[a] |
| GRANK | 0.074 | 0.029 | 0.143 | 0.541[a] |
| SIGN-COWAN | 0.136 | 0.270 | 0.916[b] | 2.994[b] |
| SIGN-GSAR-T | 0.361 | 0.387 | 0.855[b] | 2.400[b] |
| SIGN-GSAR-Z | 0.918[b] | 1.006[b] | 1.288[b] | 2.871[b] |

When the estimation window consists of 239 days, Panel D of Table 2, the departures from normality ($t$-distribution for CUMRANK-T, GRANK and SIGN-GSAR-T) of the statistics are typically not statistically significant in for the AR$(0)$ and CAR$(-1,+1)$, i.e., in the short CAR-windows. The departures from normality are statistically significant only for PATELL, for CAR$(-1,+1)$, and for SIGN-GSAR-Z for both AR$(0)$ and CAR$(-1,+1)$. In the longer CAR-windows the normality (or $t$-distribution) is rejected in many of the cases, in particular, for PATELL, BMP, CAMPBELL-WASLEY, SIGN-COWAN, SIGN-GSAR-T and SIGN-GSAR-Z. Hence, the normality of the test statistic SIGN-GSAR-Z is rejected for both short and long CAR-windows. For longer CAR-windows of CAR$(-5,+5)$ and CAR$(-10,+10)$ the non-normality of the CAMPBELL-WASLEY, in particular, is most likely due to the increasing bias in the standard error shown in equation (40). The failing normality (or $t$-distribution) of PATELL, BMP, SIGN-COWAN, SING-GSAR-T and SIGN-GSAR-Z with the long CAR-windows is not that clear.

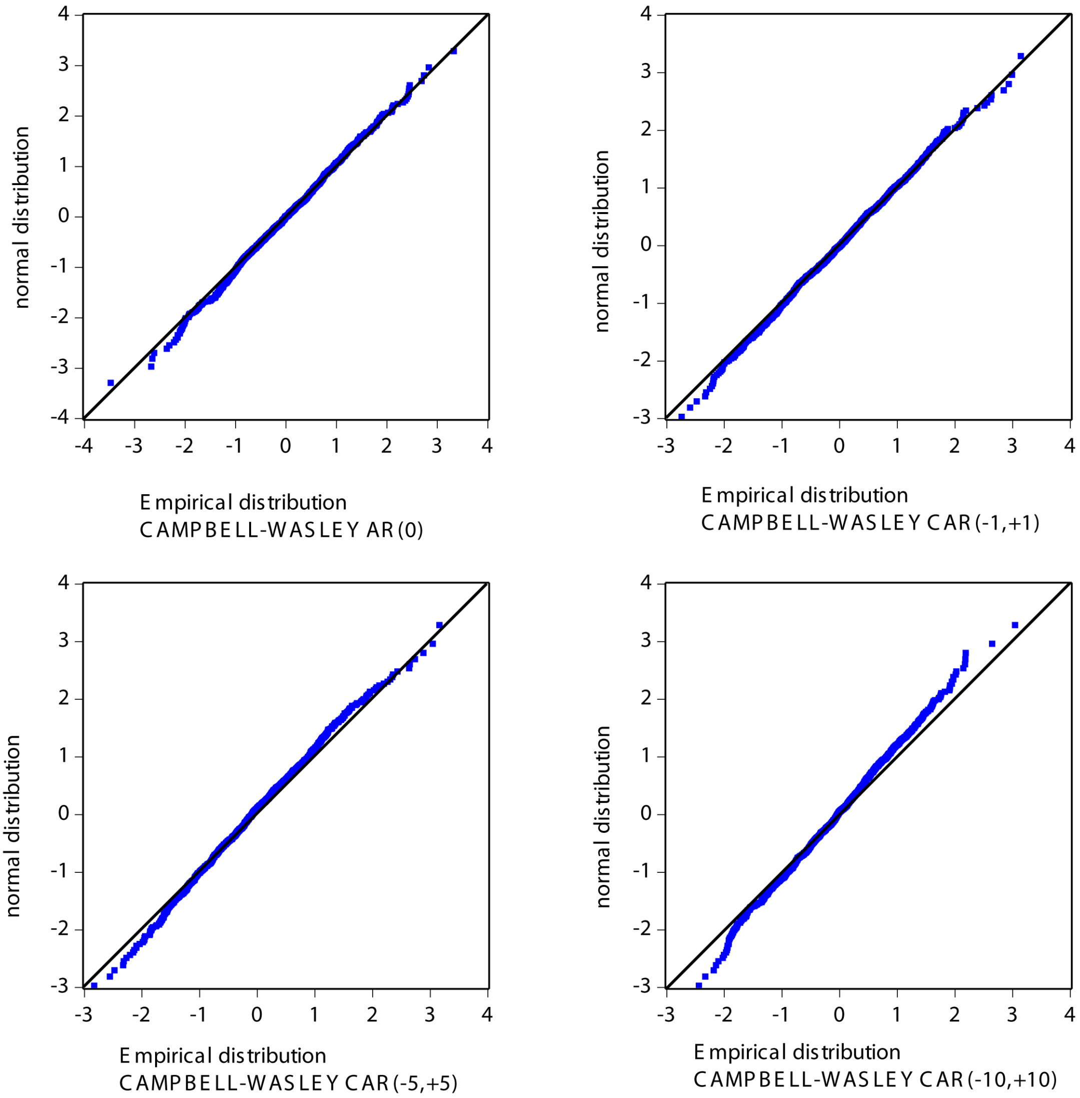


**Figure 1.** The Q-Q plots for the test statistic CAMPBELL-WASLEY in cases where the estimation window consists of 239 days and the event window consists of 21 days.

In order to get a closer view of the null-distributions of the nonparametric test statistics of CAMPBELL-WASLEY, CUMRANK-T, CUMRANK-Z, GRANK, SIGN-COWAN, SIGN-GSAR-T and SIGN-GSAR-Z Figures 1 to 7 plot empirical quantiles from 1,000 simulations against theoretical quantiles of those test statistics for abnormal return $\mathrm{AR}(0)$ and for cumulative abnormal returns in cases $\mathrm{CAR}(-1,+1)$, $\mathrm{CAR}(-5,+5)$ and $\mathrm{CAR}(-10,+10)$ under the null hypothesis of no event effect. The quantile-quantile (Q-Q) plot is a simple and powerful tool for comparing the empirical distributions against

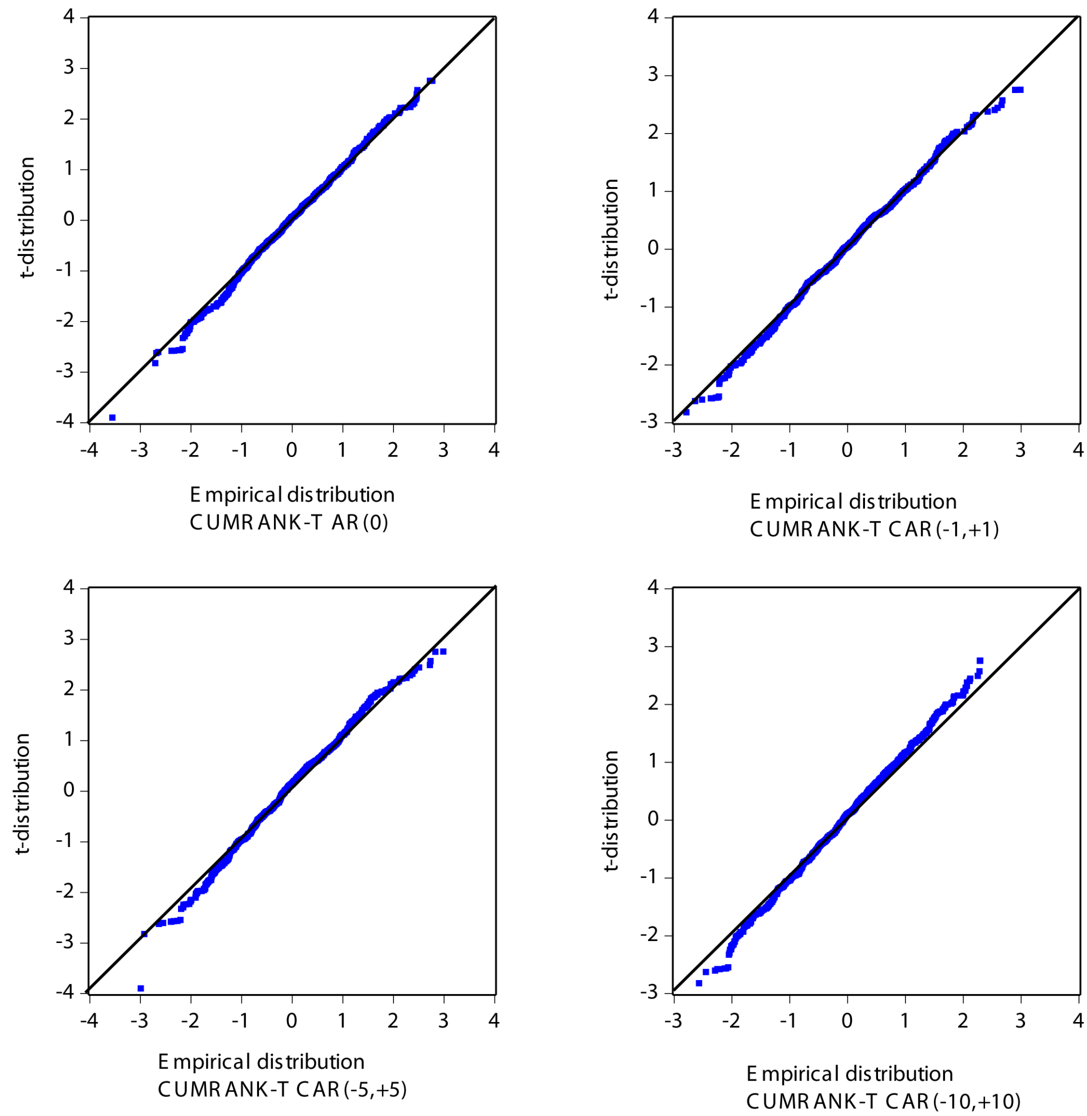


**Figure 2.** The Q-Q plots for the test statistic CUMRANK-T in cases where the estimation window consists of 239 days and the event window consists of 21 days.

their theoretical distributions.[5] If the statistics follow the theoretical distributions, the plots should be close to the 45 degree diagonal line.

According to Figures 1 to 7 all the nonparametric test statistics considered and the appropriate distributions seem to match quite well for short CAR-windows. Furthermore for longer CAR-windows the empirical and theoretical distributions of those test

[5]See e.g. Cleveland (1993) to get more information.

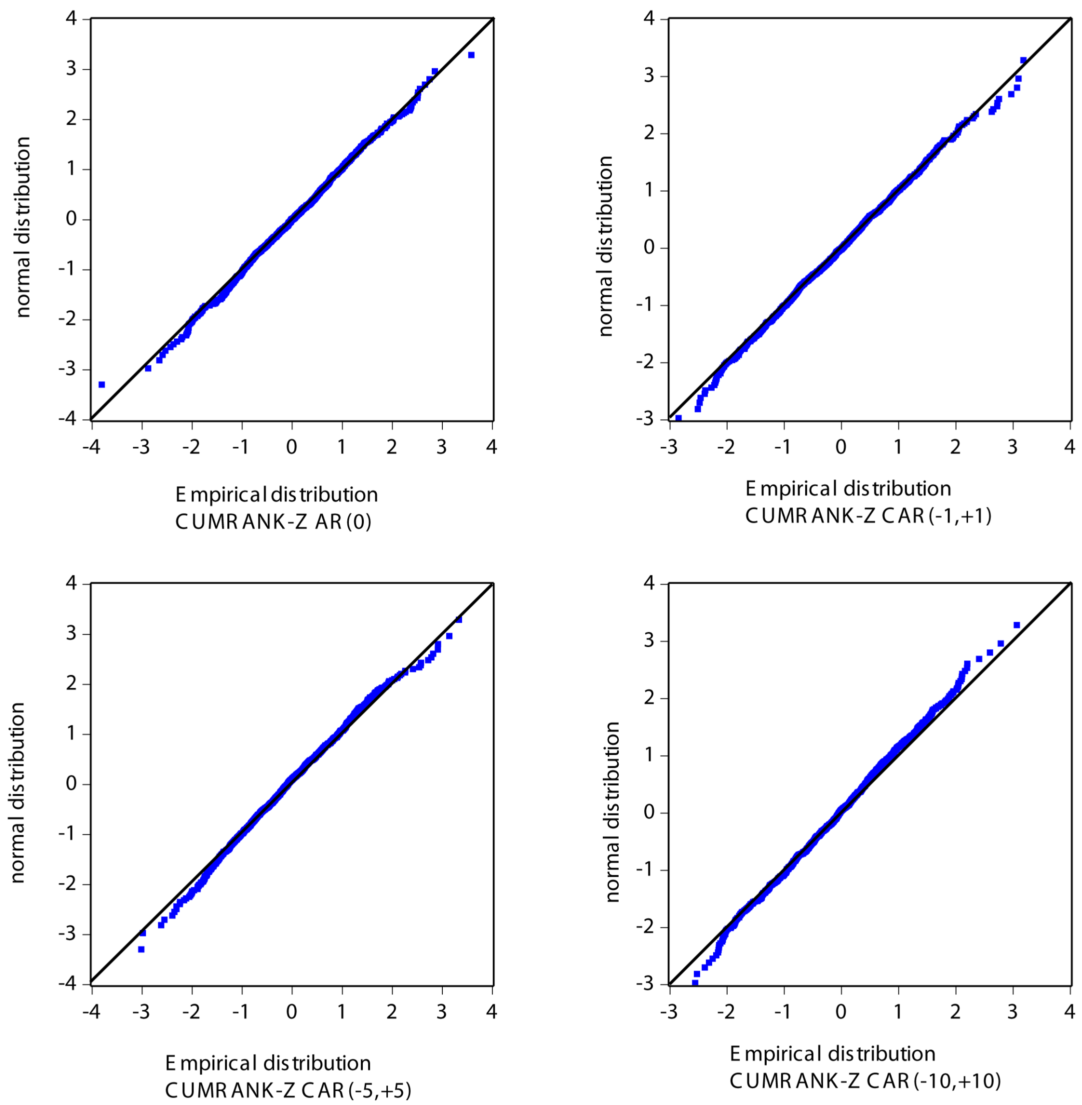


**Figure 3.** The Q-Q plots for the test statistic CUMRANK-Z in cases where the estimation window consists of 239 days and the event window consists of 21 days.

statistics seem to match quite well too, but in most of the cases not as well as for shorter CAR-windows. According to Figures 1 to 3, especially in the case of CAR$(-10,+10)$, the Q-Q plots fall on a straight line in the middle but curve downward on the left and upward on the right, which indicates that the tails of the empirical distributions of the test statistics CAMPBELL-WASLEY, CUMRANK-T and CUMRANK-Z may be somewhat thinner than the tails of the theoretical distributions for CAR$(-10,+10)$. While, according to Figures 4 and 5, again especially in the case of CAR$(-10,+10)$, the Q-

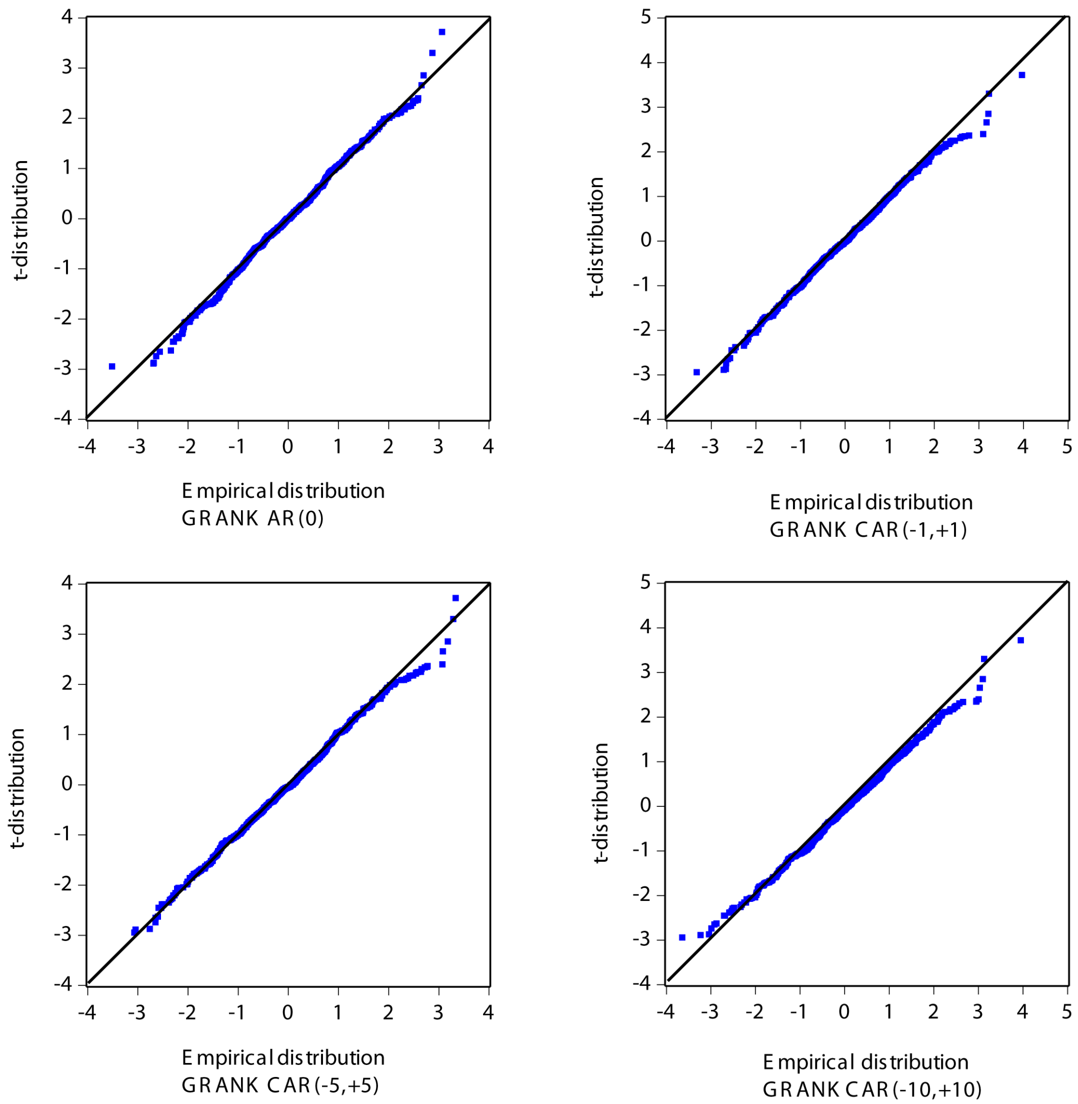


**Figure 4.** The Q-Q plots for the test statistic GRANK in cases where the estimation window consists of 239 days and the event window consists of 21 days.

Q plots fall on a straight line in the middle but curve slightly upward on the left and downward on the right, which indicates that the tails of the empirical distributions of the test statistics GRANK and SIGN-COWAN may be somewhat thicker than the tails of the theoretical distributions for $\mathrm{CAR}(-10,+10)$. According to Figure 6 the Q-Q plots for the test statistic SIGN-GSAR-T do not curve so clearly upwards or downwards. According to Figure 7 the Q-Q plots of the test statistic SIGN-GSAR-Z, for $\mathrm{AR}(0)$, $\mathrm{CAR}(-1,+1)$, $\mathrm{CAR}(-5,+5)$ and $\mathrm{CAR}(-10,+10)$, curve upward on the left and downward on the right, which indicates that the tails of the empirical distributions

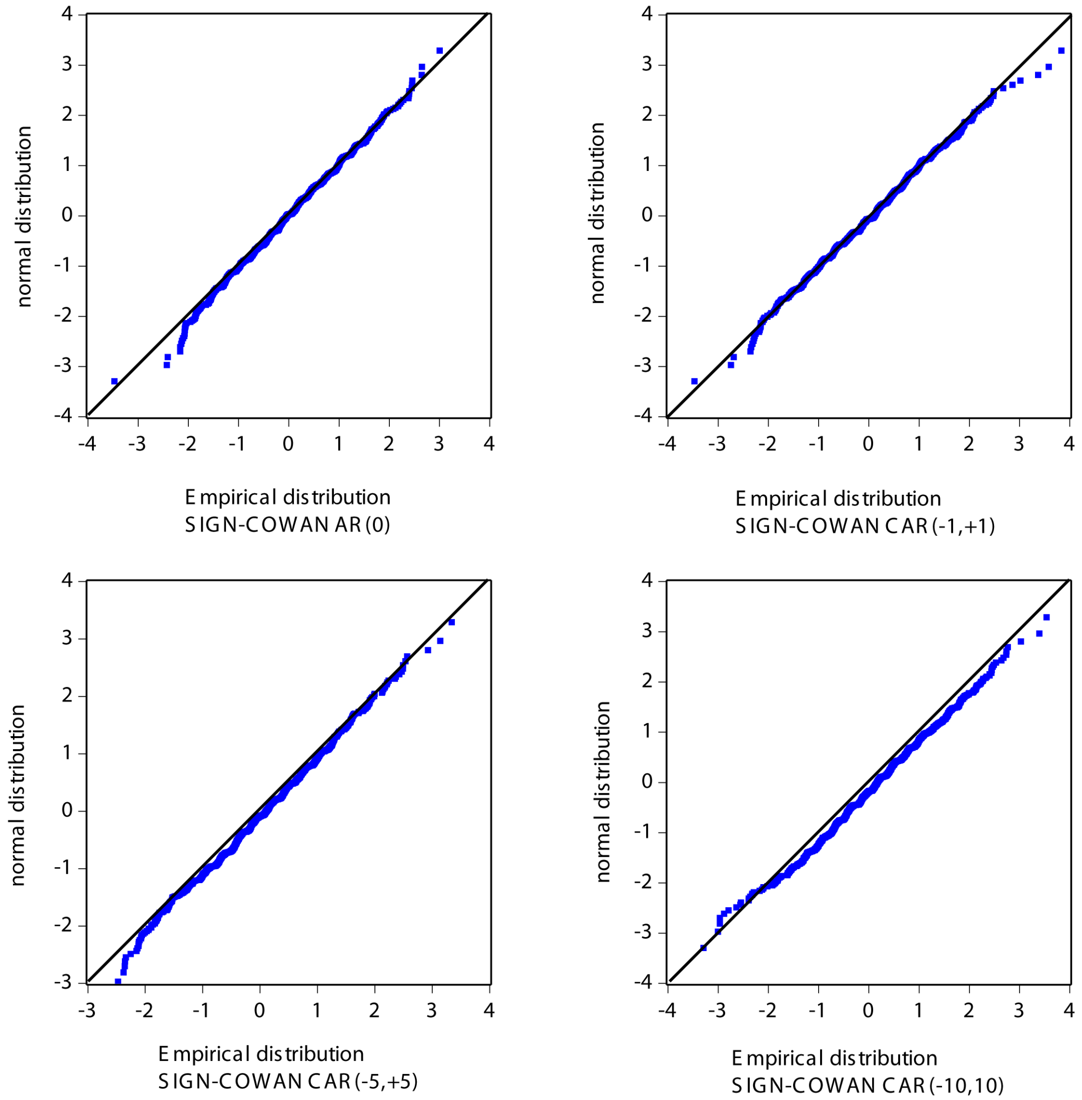


**Figure 5.** The Q-Q plots for the test statistic SIGN-COWAN in cases where the estimation window consists of 239 days and the event window consists of 21 days.

of SIGN-GSAR-Z are in every case thicker than the tails of the normal distribution.

Table 2, Panels A to C, also reports the Cramer-von Mises test statistics in the cases where the estimation window is shortened for consisting only of 25, 50 and 100 days. Again the departures from normality ($t$-distribution for CUMRANK-T, GRANK and SIGN-GSAR-T) of the statistics are typically not statistically significant in for AR$(0)$ and CAR$(-1,+1)$. Again the normality of the test statistic PATELL is rejected in some cases for short CAR-windows. Furthermore, the normality of the test statistic

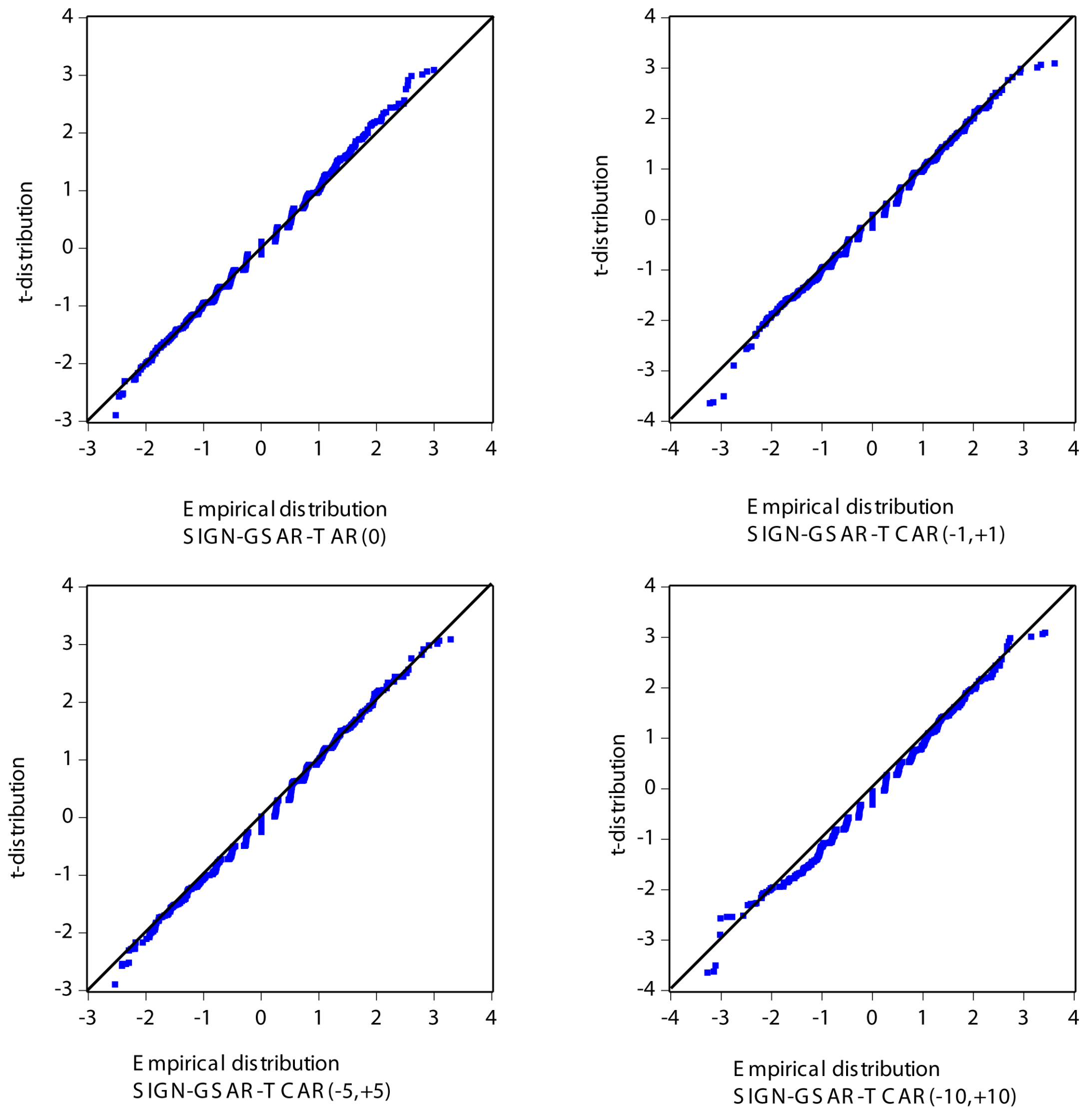


**Figure 6.** The Q-Q plots for the test statistic SIGN-GSAR-T in cases where the estimation window consists of 239 days and the event window consists of 21 days.

SIGN-GSAR-Z is rejected for both AR$(0)$ and CAR$(-1,+1)$. The normality of the test statistic ORDIN is not rejected when the estimation window consists of 239 days, Panel D of Table 2, but is rejected for longer CAR-windows when the estimation window consists only of 25 or 50 days, Panels A and B of Table 2. Hence, it seems that the shortened estimation window does not have such a strong effect on the rejection of the Cramer-von Mises test statistics for the other test statistics except the test statistic ORDIN.

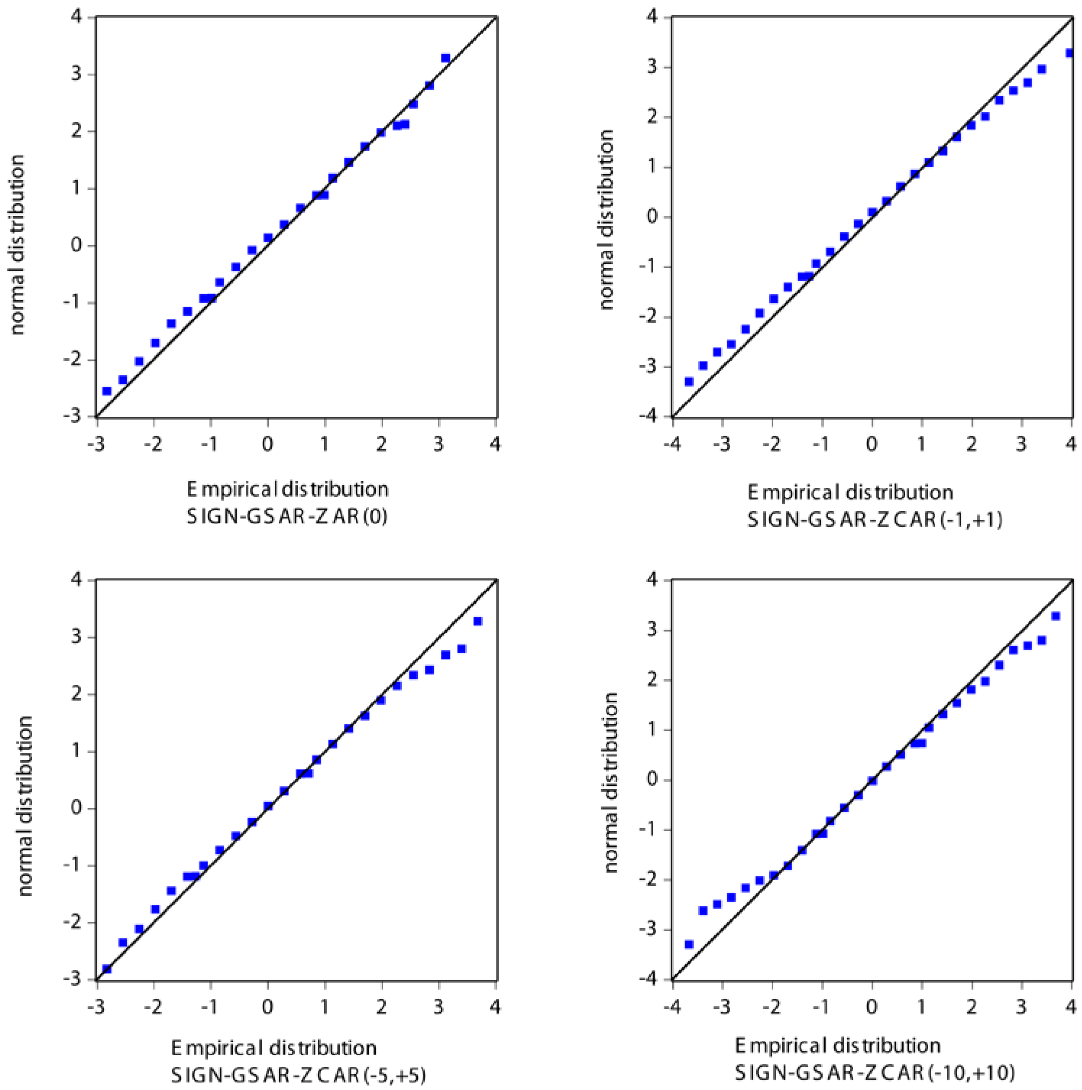


**Figure 7.** The Q-Q plots for the test statistic SIGN-GSAR-Z in cases where the estimation window consists of 239 days and the event window consists of 21 days.

## 6.3 Rejection Rates

Columns 2–4 in Table 3 report the left tail, right tail and two-tailed rejection rates (Type I errors) at the 5 percent level under the null hypothesis of no event mean effect with no event-induced volatility. Almost all rejection rates are close to the nominal rate of 0.05 for short CAR-windows of AR(0) and CAR$(-1,+1)$. Only PATELL statistics tend to over-reject the null hypothesis for two-tailed tests and SIGN-GSAR-Z statistics tend to over-reject for left and right tail tests as well as two-tailed tests. For the longer CAR-windows of CAR$(-5,+5)$ and CAR$(-10,10)$ again ORDIN, GRANK and SIGN-GSAR-T reject close to the nominal rate with rejection rates that are well within the approximate 99 percent confidence interval of [0.032, 0.068]. For the longer CAR-windows PATELL tends to over-reject in addition to the two-tailed tests on the left tail also. The BMP statistic tends to somewhat over-reject the null hypothesis for two-tailed tests for CAR$(-10,+10)$ and the SIGN-COWAN statistic tends to over-reject the null hypothesis for CAR$(-10,+10)$ for right tail tests. For longer CAR-windows the test statistic CAMPBELL-WASLEY under-rejects the null hypothesis for right tail and two-tailed tests. In addition, CUMRANK-T and CUMRANK-Z seem to somewhat under-reject the null hypothesis for CAR$(-10,+10)$ for right tail tests. For longer CAR-windows the SIGN-GSAR-Z statistic over-rejects the null hypothesis again for left and right tail tests as well as two-tailed tests. Hence, SIGN-GSAR-Z seems always to over-reject the null hypothesis.

As the Q-Q plots also have indicated, the distributions of the nonparametric test statistics CAMPBELL-WASLEY, CUMRANK-T, CUMRANK-Z and SIGN-COWAN for CAR$(-10,+10)$ seem to differ from their theoretical counterparts. It seems that the tails of the test statistic SIGN-GSAR-Z are fat in every case as well the Q-Q plots have indicated. The fat tails of the test statistics SIGN-GSAR-Z may be the reason why the Cramer-von Mises test rejects the normality of the test statistic in every case.

Columns 5–13 in Table 3 report the rejection rates under the null hypothesis in the cases where the event-induced variance is present. ORDIN and PATELL tests over-reject when the variance increases, which is a well-known outcome. At the highest factor of $c = 3.0$ the Type I errors for both ORDIN and PATELL are in the range from 0.20 to 0.30 in two-tailed testing, that is, five to six times the nominal rate. For all other test statistics, except ORDIN and PATELL, the rejection rates are quite similar in the cases where the event induced volatility is present as in the cases where there is no event induced volatility. Note that because test statistic SIGN-COWAN takes account only of the sign of the difference between abnormal return and zero, and not for example the sign of the difference between abnormal return and its median, the event-induced volatility does not have any impact on the rejection rates of the test statistic SIGN-COWAN. Hence, it can be concluded that all the other test statistics, except ORDIN and PATELL, seem to behave quite well in the cases where the event induced volatility is present.

**Table 3.** Left tail, right tail and two-tailed rejection rates of test statistics at the 5% level of the null hypothesis of no mean event effects with different event windows and different levels of event-induced volatility. The 99 percent confidence interval around the 0.05 rejection rate is [0.032,0.068].

| | c=1.0 | | | c=1.5 | | | c=2.0 | | | c=3.0 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Test statistic | Left | Right | Two-tail | Left | Right | Two-tail | Left | Right | Two-tail | Left | Right | Two-tail |
| Panel A: AR(0) | | | | | | | | | | | | |
| ORDIN | 0.050 | 0.055 | 0.058 | 0.085 | 0.091 | 0.113 | 0.131 | 0.127 | 0.162 | 0.174 | 0.188 | 0.261 |
| PATELL | 0.064 | 0.057 | 0.070 | 0.109 | 0.103 | 0.132 | 0.136 | 0.136 | 0.194 | 0.183 | 0.180 | 0.291 |
| BMP | 0.048 | 0.042 | 0.045 | 0.051 | 0.044 | 0.041 | 0.051 | 0.043 | 0.045 | 0.048 | 0.042 | 0.044 |
| CAMPBELL-WASLEY | 0.046 | 0.046 | 0.043 | 0.041 | 0.046 | 0.042 | 0.045 | 0.044 | 0.043 | 0.046 | 0.045 | 0.044 |
| CUMRANK-T | 0.046 | 0.046 | 0.043 | 0.041 | 0.046 | 0.042 | 0.045 | 0.044 | 0.043 | 0.046 | 0.044 | 0.044 |
| CUMRANK-Z | 0.046 | 0.046 | 0.051 | 0.043 | 0.047 | 0.047 | 0.044 | 0.046 | 0.048 | 0.046 | 0.046 | 0.049 |
| GRANK | 0.044 | 0.048 | 0.047 | 0.041 | 0.048 | 0.046 | 0.042 | 0.046 | 0.048 | 0.043 | 0.046 | 0.045 |
| SIGN-COWAN | 0.040 | 0.044 | 0.037 | 0.040 | 0.044 | 0.037 | 0.040 | 0.044 | 0.037 | 0.040 | 0.044 | 0.037 |
| SIGN-GSAR-T | 0.048 | 0.041 | 0.042 | 0.046 | 0.044 | 0.041 | 0.048 | 0.042 | 0.042 | 0.046 | 0.042 | 0.042 |
| SIGN-GSAR-Z | 0.087 | 0.071 | 0.085 | 0.084 | 0.073 | 0.086 | 0.087 | 0.074 | 0.086 | 0.087 | 0.070 | 0.085 |
| Panel B: CAR(-1,+1) | | | | | | | | | | | | |
| ORDIN | 0.052 | 0.044 | 0.047 | 0.085 | 0.090 | 0.108 | 0.113 | 0.127 | 0.165 | 0.163 | 0.165 | 0.250 |
| PATELL | 0.067 | 0.058 | 0.076 | 0.112 | 0.092 | 0.139 | 0.149 | 0.128 | 0.190 | 0.203 | 0.164 | 0.293 |
| BMP | 0.053 | 0.054 | 0.054 | 0.051 | 0.052 | 0.051 | 0.054 | 0.053 | 0.054 | 0.054 | 0.055 | 0.054 |
| CAMPBELL-WASLEY | 0.047 | 0.041 | 0.043 | 0.046 | 0.040 | 0.041 | 0.046 | 0.044 | 0.044 | 0.047 | 0.047 | 0.048 |
| CUMRANK-T | 0.047 | 0.041 | 0.044 | 0.046 | 0.040 | 0.041 | 0.046 | 0.044 | 0.045 | 0.047 | 0.042 | 0.044 |
| CUMRANK-Z | 0.051 | 0.046 | 0.049 | 0.047 | 0.048 | 0.046 | 0.049 | 0.049 | 0.051 | 0.048 | 0.042 | 0.045 |
| GRANK | 0.049 | 0.051 | 0.052 | 0.052 | 0.050 | 0.055 | 0.051 | 0.048 | 0.054 | 0.048 | 0.052 | 0.053 |
| SIGN-COWAN | 0.051 | 0.062 | 0.055 | 0.051 | 0.062 | 0.055 | 0.051 | 0.062 | 0.055 | 0.051 | 0.062 | 0.055 |
| SIGN-GSAR-T | 0.056 | 0.057 | 0.058 | 0.056 | 0.057 | 0.059 | 0.057 | 0.056 | 0.057 | 0.056 | 0.056 | 0.058 |
| SIGN-GSAR-Z | 0.082 | 0.091 | 0.105 | 0.083 | 0.092 | 0.105 | 0.081 | 0.090 | 0.105 | 0.083 | 0.091 | 0.104 |

(*continues*)

Table 3 (continues).

| | c=1.0 | | | c=1.5 | | | c=2.0 | | | c=3.0 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Test statistic | Left | Right | Two-tail | Left | Right | Two-tail | Left | Right | Two-tail | Left | Right | Two-tail |
| Panel C: CAR(-5,+5) | | | | | | | | | | | | |
| ORDIN | 0.052 | 0.033 | 0.042 | 0.092 | 0.065 | 0.098 | 0.127 | 0.083 | 0.148 | 0.173 | 0.132 | 0.227 |
| PATELL | 0.080 | 0.054 | 0.080 | 0.130 | 0.093 | 0.141 | 0.166 | 0.112 | 0.200 | 0.217 | 0.149 | 0.299 |
| BMP | 0.066 | 0.046 | 0.060 | 0.063 | 0.044 | 0.057 | 0.072 | 0.043 | 0.060 | 0.067 | 0.041 | 0.062 |
| CAMPBELL-WASLEY | 0.037 | 0.031 | 0.032 | 0.036 | 0.029 | 0.030 | 0.037 | 0.030 | 0.033 | 0.040 | 0.029 | 0.045 |
| CUMRANK-T | 0.042 | 0.032 | 0.036 | 0.041 | 0.030 | 0.034 | 0.041 | 0.032 | 0.036 | 0.041 | 0.030 | 0.037 |
| CUMRANK-Z | 0.042 | 0.032 | 0.036 | 0.047 | 0.032 | 0.035 | 0.044 | 0.037 | 0.046 | 0.034 | 0.038 | 0.037 |
| GRANK | 0.056 | 0.054 | 0.058 | 0.053 | 0.050 | 0.058 | 0.057 | 0.051 | 0.058 | 0.055 | 0.053 | 0.057 |
| SIGN-COWAN | 0.042 | 0.045 | 0.042 | 0.042 | 0.045 | 0.042 | 0.042 | 0.045 | 0.042 | 0.042 | 0.045 | 0.042 |
| SIGN-GSAR-T | 0.041 | 0.053 | 0.039 | 0.041 | 0.053 | 0.038 | 0.040 | 0.055 | 0.039 | 0.040 | 0.054 | 0.037 |
| SIGN-GSAR-Z | 0.076 | 0.079 | 0.091 | 0.076 | 0.080 | 0.090 | 0.076 | 0.077 | 0.092 | 0.076 | 0.078 | 0.090 |
| Panel D: CAR(-10,+10) | | | | | | | | | | | | |
| ORDIN | 0.053 | 0.033 | 0.048 | 0.090 | 0.078 | 0.099 | 0.120 | 0.101 | 0.154 | 0.174 | 0.146 | 0.233 |
| PATELL | 0.086 | 0.050 | 0.075 | 0.125 | 0.089 | 0.145 | 0.0159 | 0.115 | 0.198 | 0.213 | 0.161 | 0.289 |
| BMP | 0.067 | 0.047 | 0.069 | 0.065 | 0.050 | 0.066 | 0.067 | 0.047 | 0.065 | 0.066 | 0.044 | 0.069 |
| CAMPBELL-WASLEY | 0.038 | 0.024 | 0.020 | 0.039 | 0.030 | 0.022 | 0.042 | 0.030 | 0.021 | 0.039 | 0.027 | 0.020 |
| CUMRANK-T | 0.049 | 0.031 | 0.034 | 0.044 | 0.032 | 0.030 | 0.047 | 0.031 | 0.033 | 0.049 | 0.030 | 0.034 |
| CUMRANK-Z | 0.049 | 0.031 | 0.038 | 0.046 | 0.036 | 0.038 | 0.048 | 0.037 | 0.039 | 0.050 | 0.032 | 0.039 |
| GRANK | 0.053 | 0.063 | 0.064 | 0.053 | 0.063 | 0.061 | 0.051 | 0.062 | 0.063 | 0.053 | 0.065 | 0.063 |
| SIGN-COWAN | 0.032 | 0.073 | 0.060 | 0.032 | 0.073 | 0.060 | 0.032 | 0.073 | 0.060 | 0.032 | 0.073 | 0.060 |
| SIGN-GSAR-T | 0.032 | 0.063 | 0.054 | 0.033 | 0.066 | 0.065 | 0.034 | 0.064 | 0.054 | 0.034 | 0.063 | 0.054 |
| SIGN-GSAR-Z | 0.044 | 0.092 | 0.089 | 0.046 | 0.094 | 0.093 | 0.044 | 0.092 | 0.092 | 0.046 | 0.092 | 0.091 |

**Table 4.** Two-tailed rejection rates of the test statistics at the 5% level of the null hypothesis of no mean event effects with different event windows and different length of the estimation period. The 99 percent confidence interval around the 0.05 rejection rate is [0.032,0.068].

| | Length of the estimation period | | | | | |
|---|---|---|---|---|---|---|
| Test statistic | 25 days | 50 days | 100 days | 25 days | 50 days | 100 days |
| | Panel A: AR(0) | | | Panel B: CAR(-1,+1) | | |
| ORDIN | 0.066 | 0.066 | 0.063 | 0.040 | 0.048 | 0.044 |
| PATELL | 0.104 | 0.090 | 0.068 | 0.076 | 0.071 | 0.064 |
| BMP | 0.053 | 0.052 | 0.042 | 0.058 | 0.055 | 0.055 |
| CAMPBELL-WASLEY | 0.037 | 0.040 | 0.045 | 0.033 | 0.040 | 0.046 |
| CUMRANK-T | 0.037 | 0.040 | 0.045 | 0.040 | 0.042 | 0.047 |
| CUMRANK-Z | 0.037 | 0.040 | 0.045 | 0.038 | 0.046 | 0.049 |
| GRANK | 0.042 | 0.039 | 0.046 | 0.053 | 0.051 | 0.053 |
| SIGN-COWAN | 0.046 | 0.042 | 0.039 | 0.064 | 0.054 | 0.050 |
| SIGN-GSAR-T | 0.045 | 0.036 | 0.052 | 0.063 | 0.059 | 0.058 |
| SIGN-GSAR-Z | 0.091 | 0.077 | 0.083 | 0.099 | 0.090 | 0.089 |
| | Panel C: CAR(-5,+5) | | | Panel D: CAR(-10,+10) | | |
| ORDIN | 0.021 | 0.019 | 0.038 | 0.011 | 0.028 | 0.036 |
| PATELL | 0.043 | 0.066 | 0.066 | 0.066 | 0.044 | 0.068 |
| BMP | 0.053 | 0.054 | 0.060 | 0.061 | 0.061 | 0.062 |
| CAMPBELL-WASLEY | 0.010 | 0.023 | 0.028 | 0.005 | 0.010 | 0.012 |
| CUMRANK-T | 0.031 | 0.037 | 0.033 | 0.029 | 0.025 | 0.028 |
| CUMRANK-Z | 0.031 | 0.034 | 0.037 | 0.026 | 0.025 | 0.027 |
| GRANK | 0.048 | 0.052 | 0.059 | 0.058 | 0.054 | 0.058 |
| SIGN-COWAN | 0.051 | 0.059 | 0.054 | 0.050 | 0.055 | 0.052 |
| SIGN-GSAR-T | 0.054 | 0.058 | 0.050 | 0.046 | 0.054 | 0.051 |
| SIGN-GSAR-Z | 0.100 | 0.086 | 0.084 | 0.098 | 0.094 | 0.102 |

Table 4 reports the two-tailed rejection rates (Type I errors) at the 5 percent level under the null hypothesis of no event mean effect when the estimation window consists only of 25, 50 and 100 days. Generally the rejection rates are not very sensitive to the length of the estimation window, when the CAR-window is short. For AR(0) and CAR$(-1,+1)$ PATELL test statistic over-rejects the null hypothesis in cases where the estimation window is as short as 25 and 50 days. For CAR$(-5,+5)$ and CAR$(-10,+10)$ test statistics ORDIN and CAMPBELL-WASLEY clearly under-reject the null hypothesis when the estimation window is shortened. Hence, the sample size in particular seems to substantially affect the under-rejection of CAMPBELL-WASLEY in the longer CAR-windows. For example, with CAR$(-10,+10)$ the empirical rejection rate is 0.005 with estimation window length of 25, which is one tenth of the nominal rate. The test statistics CUMRANK-T and CUMRANK-Z also somewhat under-reject the null hypothesis when the estimation window is short, but clearly not as much as the test statistic CAMPBELL-WASLEY. Again, when the estimation window is 25, 50 or 100 days, the SIGN-GSAR-Z test statistic over-rejects the null hypothesis. Hence, we can conclude that according to the rejection rates the test statistics BMP, GRANK, SIGN-COWAN and SIGN-GSAR-T together with the test statistics CUMRANK-T and CUMRANK-Z are the best options in special cases where the estimation window is shorter than 239 days.

## 6.4 Power of the Tests

### *6.4.1 Non-clustered event days*

The power results of the test statistics for two-tailed tests are shown in Panels A to D of Table 5 and most of them are graphically depicted in Figures 8–11. The zero abnormal return line in each panel of Table 5 indicates the Type I error rate and replicates column 4 of Table 3 (i.e. there is no event-induced volatility). The rest of the lines of Table 5 indicate the rejection rates for the respective abnormal returns shown in the first column. For AR(0) the abnormal performance is artificially introduced by adding the indicated percentage (a constant) to the day-0 return of each security. While, for CAR$(-1,+1)$, CAR$(-5,+5)$ and CAR$(-10,+10)$, the abnormal performance is introduced by selecting one day of the CAR-window at random and adding the particular level of abnormal performance to that day's return.

**Table 5.** Non-clustered event days: Powers of the selected test statistics in two-tailed testing at the 5% significance from 1,000 random portfolios of $n = 50$ securities. The zero abnormal return line in each panel of this table indicates the Type I error rates. The rest of the lines of this table indicate the rejection rates for the respective abnormal returns shown in the first column.

| Panel A: AR(0) | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| AR | ORDIN | PATELL | BMP | CAMPBELL-WASLEY | CUM-RANK-T | CUM-RANK-Z | GRANK | SIGN-COWAN | SIGN-GSAR-T | SIGN-GSAR-Z |
| −3.0 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| −2.0 | 0.996 | 1.000 | 0.999 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.997 | 0.999 |
| −1.0 | 0.720 | 0.960 | 0.909 | 0.958 | 0.958 | 0.964 | 0.958 | 0.912 | 0.856 | 0.910 |
| ±0.0 | 0.058 | 0.070 | 0.045 | 0.043 | 0.043 | 0.051 | 0.047 | 0.037 | 0.042 | 0.085 |
| +1.0 | 0.729 | 0.942 | 0.899 | 0.971 | 0.971 | 0.977 | 0.969 | 0.928 | 0.886 | 0.928 |
| +2.0 | 0.994 | 0.999 | 0.996 | 0.999 | 0.999 | 0.999 | 0.999 | 1.000 | 1.000 | 1.000 |
| +3.0 | 1.000 | 1.000 | 0.999 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| Panel B: CAR(-1,+1) | | | | | | | | | | |
| AR | ORDIN | PATELL | BMP | CAMPBELL-WASLEY | CUM-RANK-T | CUM-RANK-Z | GRANK | SIGN-COWAN | SIGN-GSAR-T | SIGN-GSAR-Z |
| −3.0 | 0.994 | 1.000 | 0.999 | 0.999 | 0.999 | 1.000 | 1.000 | 0.999 | 0.998 | 1.000 |
| −2.0 | 0.983 | 0.965 | 0.971 | 0.971 | 0.974 | 0.974 | 0.984 | 0.957 | 0.909 | 0.954 |
| −1.0 | 0.313 | 0.595 | 0.572 | 0.591 | 0.593 | 0.610 | 0.656 | 0.528 | 0.446 | 0.575 |
| ±0.0 | 0.047 | 0.076 | 0.054 | 0.043 | 0.044 | 0.049 | 0.052 | 0.055 | 0.058 | 0.105 |
| +1.0 | 0.297 | 0.535 | 0.514 | 0.600 | 0.600 | 0.625 | 0.638 | 0.572 | 0.479 | 0.611 |
| +2.0 | 0.834 | 0.981 | 0.958 | 0.978 | 0.978 | 0.980 | 0.993 | 0.971 | 0.928 | 0.964 |
| +3.0 | 0.987 | 1.000 | 0.999 | 1.000 | 1.000 | 1.000 | 1.000 | 0.999 | 0.995 | 0.998 |

(*continues*)

Table 5 (continues).

| Panel C: CAR(-5,+5) | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| AR | ORDIN | PATELL | BMP | CAMPBELL-WASLEY | CUM-RANK-T | CUM-RANK-Z | GRANK | SIGN-COWAN | SIGN-GSAR-T | SIGN-GSAR-Z |
| −3.0 | 0.633 | 0.912 | 0.892 | 0.697 | 0.707 | 0.722 | 0.913 | 0.782 | 0.716 | 0.799 |
| −2.0 | 0.326 | 0.619 | 0.614 | 0.489 | 0.508 | 0.533 | 0.628 | 0.481 | 0.438 | 0.550 |
| −1.0 | 0.120 | 0.244 | 0.239 | 0.195 | 0.212 | 0.230 | 0.233 | 0.155 | 0.146 | 0.221 |
| ±0.0 | 0.042 | 0.080 | 0.060 | 0.032 | 0.036 | 0.037 | 0.058 | 0.042 | 0.039 | 0.091 |
| +1.0 | 0.085 | 0.182 | 0.171 | 0.165 | 0.174 | 0.188 | 0.234 | 0.195 | 0.175 | 0.254 |
| +2.0 | 0.312 | 0.535 | 0.538 | 0.434 | 0.451 | 0.480 | 0.659 | 0.572 | 0.490 | 0.614 |
| +3.0 | 0.616 | 0.867 | 0.843 | 0.687 | 0.696 | 0.708 | 0.920 | 0.865 | 0.791 | 0.856 |
| Panel D: CAR(-10,+10) | | | | | | | | | | |
| AR | ORDIN | PATELL | BMP | CAMPBELL-WASLEY | CUM-RANK-T | CUM-RANK-Z | GRANK | SIGN-COWAN | SIGN-GSAR-T | SIGN-GSAR-Z |
| −3.0 | 0.355 | 0.680 | 0.670 | 0.373 | 0.402 | 0.423 | 0.655 | 0.471 | 0.422 | 0.551 |
| −2.0 | 0.184 | 0.379 | 0.378 | 0.235 | 0.261 | 0.273 | 0.368 | 0.223 | 0.212 | 0.299 |
| −1.0 | 0.088 | 0.165 | 0.154 | 0.100 | 0.119 | 0.119 | 0.141 | 0.077 | 0.067 | 0.139 |
| ±0.0 | 0.048 | 0.075 | 0.069 | 0.020 | 0.034 | 0.038 | 0.064 | 0.060 | 0.054 | 0.089 |
| +1.0 | 0.069 | 0.126 | 0.118 | 0.081 | 0.090 | 0.101 | 0.161 | 0.150 | 0.124 | 0.203 |
| +2.0 | 0.162 | 0.316 | 0.320 | 0.208 | 0.236 | 0.251 | 0.398 | 0.349 | 0.300 | 0.403 |
| +3.0 | 0.351 | 0.609 | 0.598 | 0.361 | 0.402 | 0.439 | 0.699 | 0.613 | 0.539 | 0.637 |

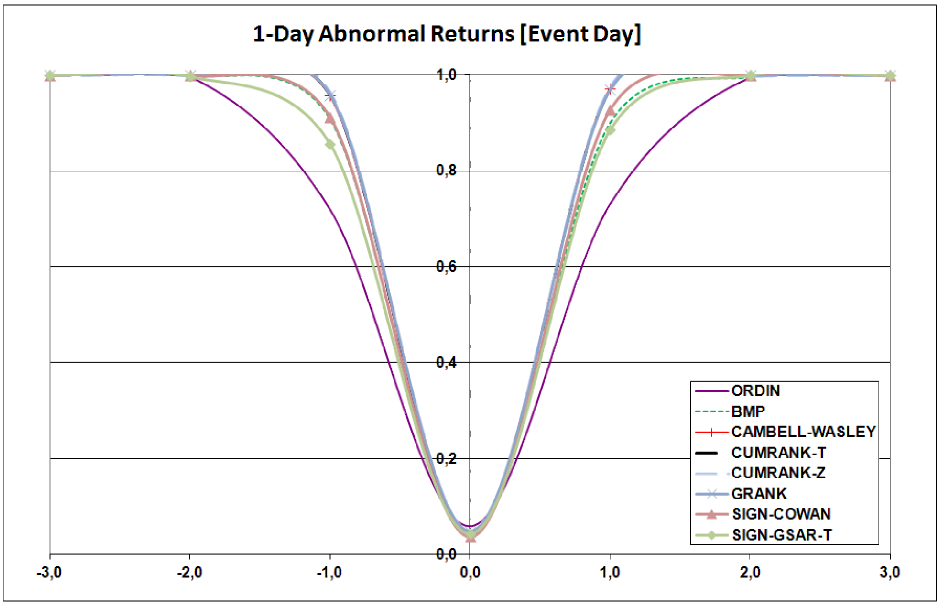


**Figure 8.** The power results of the selected test statistics for AR(0) with an abnormal return ranging from −3 percent to +3 percent and when the event days are non-clustered.

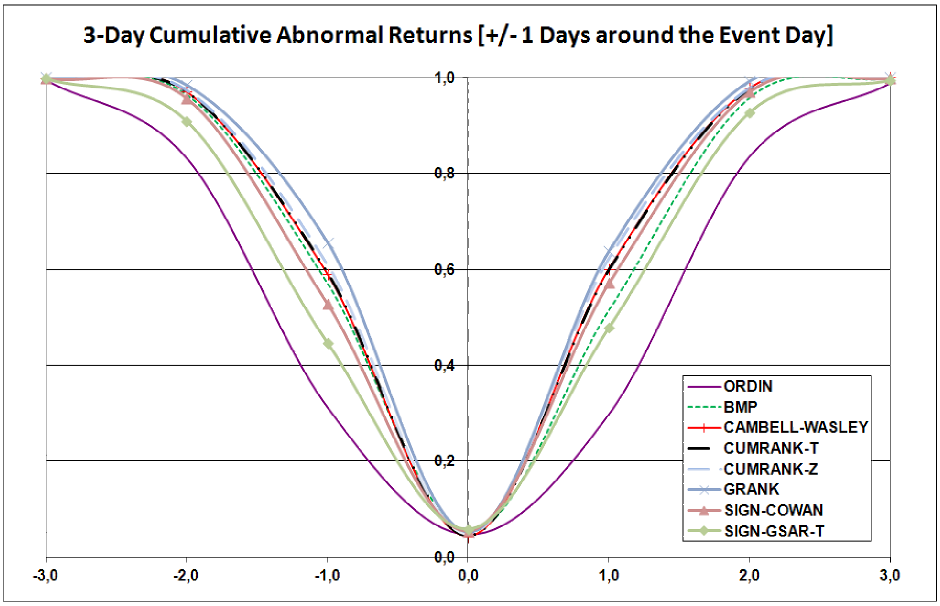


**Figure 9.** The power results of the selected test statistics for CAR$(-1,+1)$ with an abnormal return ranging from −3 percent to +3 percent and when the event days are non-clustered.

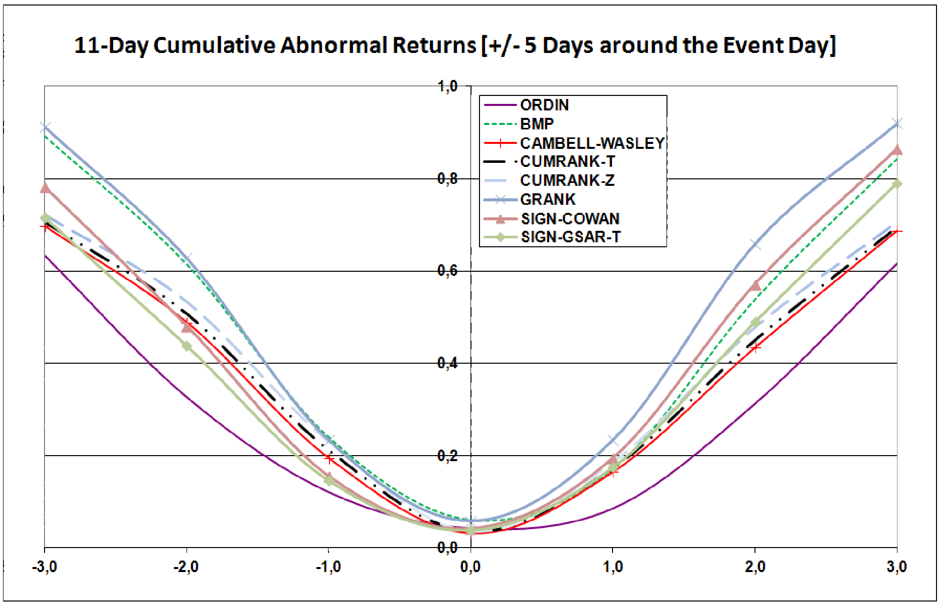


**Figure 10.** The power results of the selected test statistics for CAR$(-5,+5)$ with an abnormal return ranging from $-3$ percent to $+3$ percent and when the event days are non-clustered.

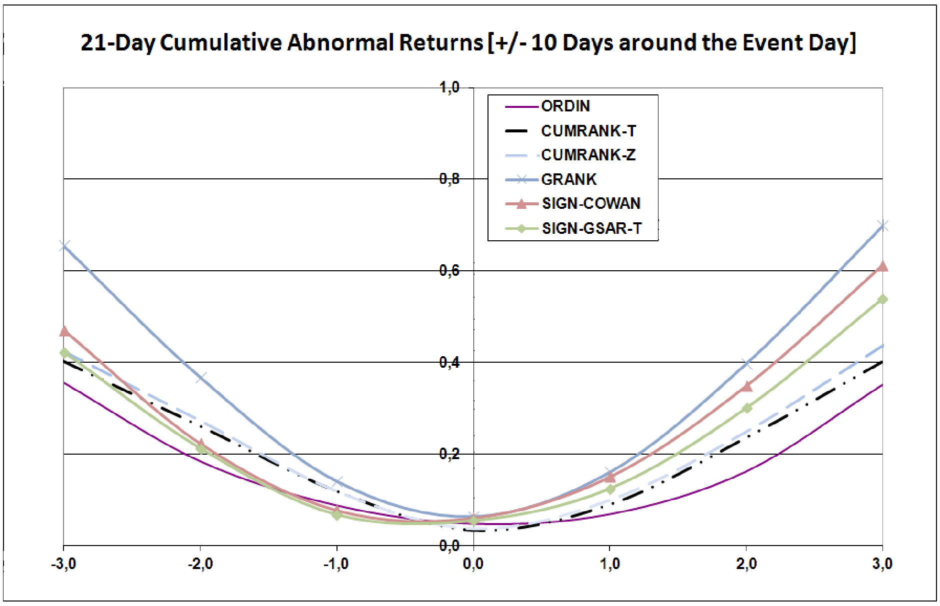


**Figure 11.** The power results of the selected test statistics for CAR$(-10,+10)$ with an abnormal return ranging from $-3$ percent to $+3$ percent and when the event days are non-clustered.

Some of the tests tend to over-reject or under-reject the null hypothesis, as also reported in previous section, which makes some power comparisons complicated or perhaps even impossible. For example, because the test statistic SIGN-GSAR-Z over-rejects the null hypothesis, it seems to be more powerful than some other statistics. However, this kind of straightforward conclusion should not be drawn. In the same way, it seems that the test statistic CAMPBELL-WASLEY looses its power for longer CAR-windows faster than it does for CUMRANK-Z and CUMRANK-T. However, according to Table 3 the test statistic CAMPBELL-WASLEY also under-rejects the null hypothesis more than the test statistics CUMRANK-Z and CUMRANK-T for longer CAR-windows, which may be the reason why it seems to be less powerful than CUMRANK-Z and CUMRANK-T for longer CAR-windows. Therefore, the conclusions on the power of the tests for longer CAR-windows should be drawn with caution, because many of the test statistics over-reject or under-reject the null hypothesis especially when the CAR-window consists of 21 days. Figures 8–11 graphically present the power properties for the test statistics, which do not over-reject or under-reject the null hypothesis when the abnormal return is zero.[6] There are four outstanding results. First, for shorter and longer CAR-windows, ORDIN, which is based on non-standardized returns is materially less powerful than the other test statistics that are based on standardized returns. Second, GRANK seems to be one of the most powerful tests for shorter CAR-windows as well as for longer CAR-windows. Third, SIGN-COWAN seems to be somewhat more powerful than SIGN-GSAR-T. Fourth, generally it seems that, in cases where power comparisons are possible to make, the nonparametric test statistics are at least as powerful as the parametric test statistics.

### *6.4.2 Clustered event days*

Table 6 reports the Type I error and power results of the test statistics with clustered event days. Hence, the event day is now the same day for each stock in the portfolio. The zero abnormal return line in each panel again indicates the Type I error rates at the five percent level under the null hypothesis of no event mean effect. Again, for $\text{AR}(0)$ the abnormal performance is artificially introduced by adding the indicated percentage (a constant) to the day-0 return of each security. While, for $\text{CAR}(-1,+1)$, $\text{CAR}(-5,+5)$ and $\text{CAR}(-10,+10)$, abnormal performance is introduced by selecting

[6] In other words, Figures 8–11 do not present the power properties for the test statistics BMP and CAMPBELL-WASLEY for the case of $\text{CAR}(-10,+10)$, and for the test statistics PATELL and SIGN-GSAR-Z for the cases of $\text{AR}(0)$, $\text{CAR}(-1,+1)$, $\text{CAR}(-5,+5)$ and $\text{CAR}(-10,+10)$.

one day of the CAR-window at random and adding the particular level of abnormal performance to that day's return.

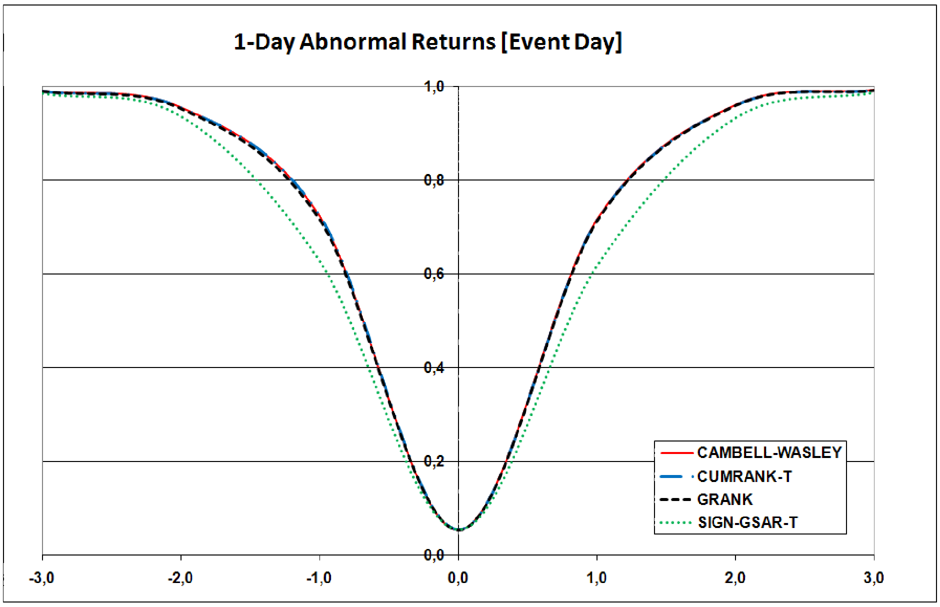


**Figure 12.** The power results of the selected test statistics for AR(0) with an abnormal return ranging from −3 percent to +3 percent and when the event days are clustered.

Consistent with Section 4.4 and earlier results [e.g. Kolari and Pynnönen (2010)], test statistics ORDIN, PATELL, BMP, CUMRANK-Z and SIGN-GSAR-Z together with SIGN-COWAN are prone to material over-rejection of the true null hypothesis of no event effect. According to Table 6 the nonparametric test statistics CAMPBELL-WASLEY, CUMRANK-T, GRANK and SIGN-GSAR-T are much more robust to cross-correlation caused by event day clustering. The power results of the test statistics CAMPBELL-WASLEY, CUMRANK-T, GRANK and SIGN-GSAR-T, in the case of AR(0), for clustered event days are graphically depicted in Figure 12. It seems that when the event days are clustered, the test statistics CAMPBELL-WASLEY, GRANK and CUMRANK-T are as powerful with each other while, the test statistic SIGN-GSAR-T seems to be somewhat less powerful, but the difference is quite small. The power comparisons of the test statistics CAMPBELL-WASLEY, CUMRANK-T and GRANK for longer CAR-windows should be done with caution when the event days are clustered, because those test statistics also somewhat over-reject the null hypothesis for longer CAR-windows. It is notable that the over-rejection is not as clear

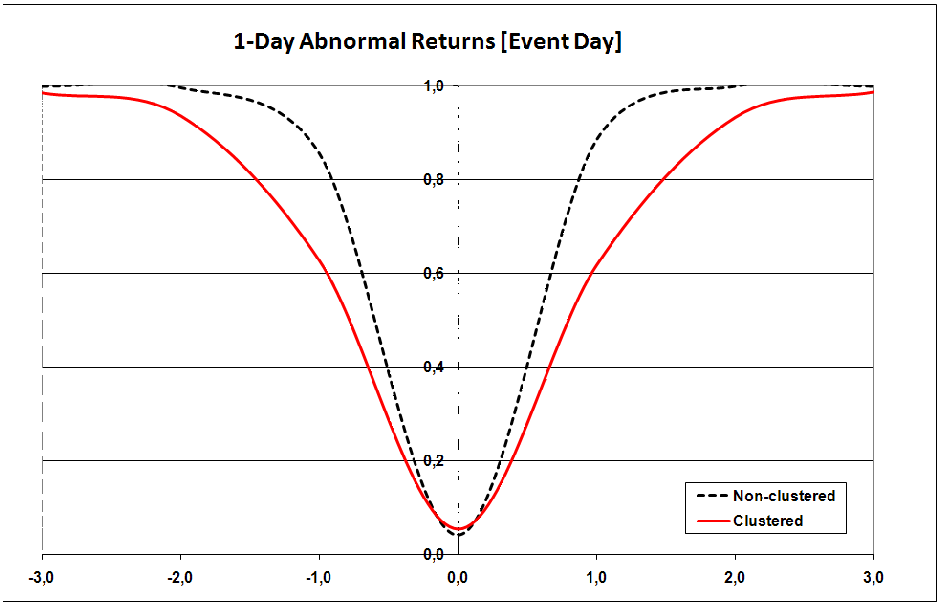


**Figure 13.** The power results of the test statistic SIGN-GSAR-T for AR(0) with an abnormal return ranging from −3 percent to +3 percent and when the event days are non-clustered and clustered.

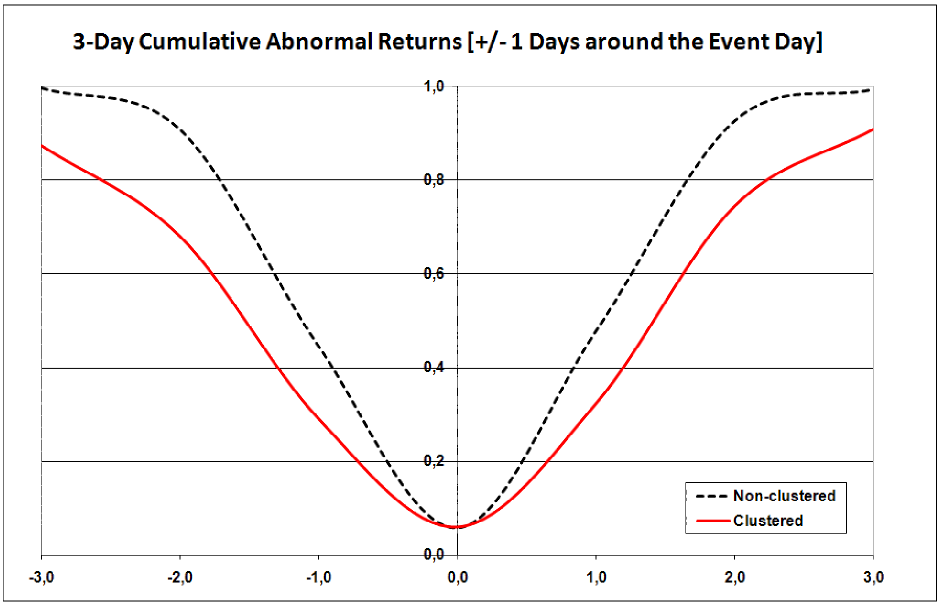


**Figure 14.** The power results of the test statistic SIGN-GSAR-T for CAR(−1, +1) with an abnormal return ranging from −3 percent to +3 percent and when the event days are non-clustered and clustered.

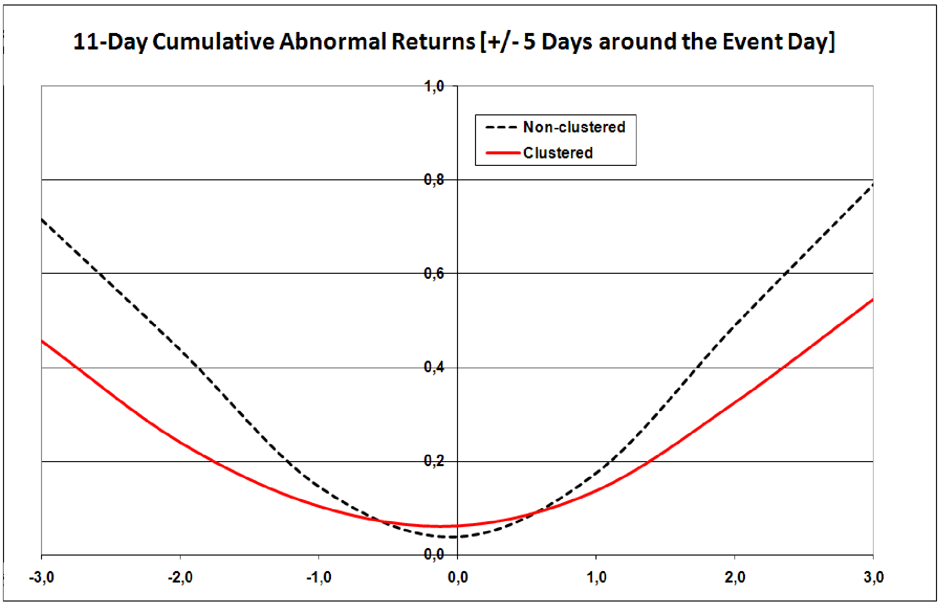


**Figure 15.** The power results of the test statistic SIGN-GSAR-T for CAR$(-5, +5)$ with an abnormal return ranging from $-3$ percent to $+3$ percent and when the event days are non-clustered and clustered.

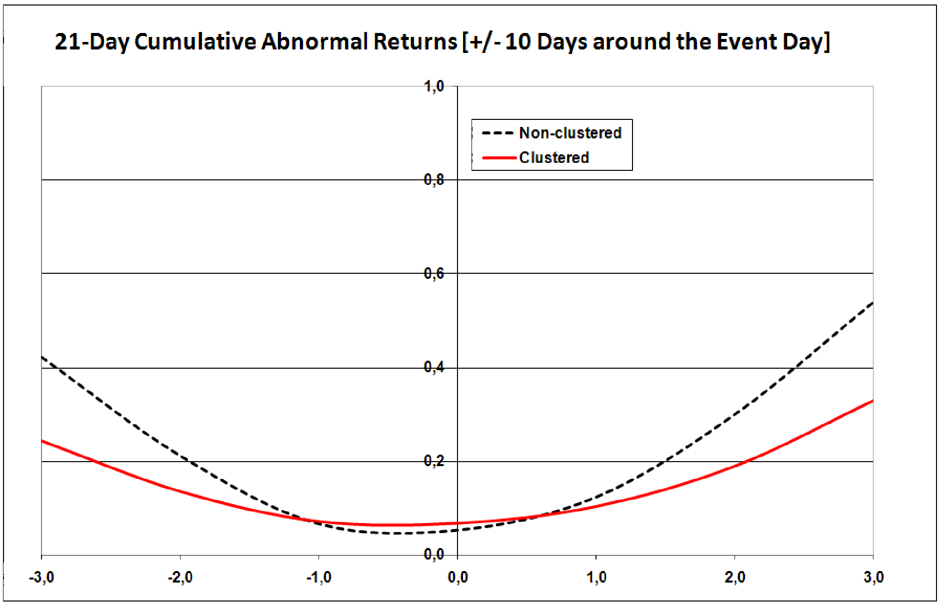


**Figure 16.** The power results of the test statistic SIGN-GSAR-T for CAR$(-10, +10)$ with an abnormal return ranging from $-3$ percent to $+3$ percent and when the event days are non-clustered and clustered.

as for the test statistics ORDIN, PATELL, BMP, CUMRANK-Z, SIGN-COWAN and SIGN-GSAR-Z. When the event days are clustered the test statistic SIGN-GSAR-T is the only one that rejects for every CAR-window length close to the nominal rate with rejection rates that are well within the approximate 99 percent confidence interval of $[0.032, 0.068]$.

The power results in Table 6 are notably distinct from those in Table 5 in that the powers tend to be discernibly lower in the clustered case. This is due to the information loss caused by cross-correlation, which is discussed more detail in Kolari and Pynnönen (2010). The non-clustered and clustered power properties of the test statistic SIGN-GSAR-T, for the cases $\text{AR}(0)$, $\text{CAR}(-1,+1)$, $\text{CAR}(-5,+5)$ and $\text{CAR}(-10,+10)$, are presented in Figures 13–16. Also these Figures reveal that the powers of the test statistic SIGN-GSAR-T tend to be somewhat lower in the clustered case.

To summarize, the derived test statistics CUMRANK-T and SIGN-GSAR-T as well the test statistics CAMPBELL-WASLEY and GRANK are quite robust faced with clustered event days. In addition, the well-established asymptotic properties of SIGN-GSAR-T and CUMRANK-T, their being robust against event-induced volatility, and highly competitive power compared to tests that do not tolerate cross-correlation make them useful robust testing procedures in event studies.

**Table 6.** Clustered event days: Powers of the selected test statistics in two-tailed testing at the 5% significance from 1,000 random portfolios of $n = 50$ securities. The zero abnormal return line in each panel of this table indicates the type I error rates. The rest of the lines of this table indicate the rejection rates for the respective abnormal returns shown in the first column.

| Panel A: AR(0) | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| AR | ORDIN | PATELL | BMP | CAMPBELL-WASLEY | CUM-RANK-T | CUM-RANK-Z | GRANK | SIGN-COWAN | SIGN-GSAR-T | SIGN-GSAR-Z |
| −3.0 | 0.995 | 0.998 | 0.996 | 0.991 | 0.991 | 0.998 | 0.990 | 0.997 | 0.985 | 0.995 |
| −2.0 | 0.968 | 0.988 | 0.982 | 0.956 | 0.956 | 0.984 | 0.953 | 0.981 | 0.936 | 0.980 |
| −1.0 | 0.692 | 0.858 | 0.844 | 0.725 | 0.725 | 0.886 | 0.716 | 0.851 | 0.627 | 0.852 |
| ±0.0 | 0.177 | 0.203 | 0.216 | 0.055 | 0.055 | 0.238 | 0.055 | 0.190 | 0.055 | 0.244 |
| +1.0 | 0.690 | 0.838 | 0.828 | 0.716 | 0.716 | 0.875 | 0.714 | 0.839 | 0.618 | 0.853 |
| +2.0 | 0.965 | 0.989 | 0.980 | 0.961 | 0.961 | 0.990 | 0.960 | 0.985 | 0.933 | 0.987 |
| +3.0 | 0.999 | 0.999 | 0.996 | 0.993 | 0.993 | 0.999 | 0.992 | 0.998 | 0.987 | 0.999 |
| Panel B: CAR(-1,+1) | | | | | | | | | | |
| AR | ORDIN | PATELL | BMP | CAMPBELL-WASLEY | CUM-RANK-T | CUM-RANK-Z | GRANK | SIGN-COWAN | SIGN-GSAR-T | SIGN-GSAR-Z |
| −3.0 | 0.928 | 0.976 | 0.956 | 0.842 | 0.845 | 0.956 | 0.923 | 0.954 | 0.874 | 0.960 |
| −2.0 | 0.750 | 0.895 | 0.880 | 0.657 | 0.663 | 0.872 | 0.779 | 0.858 | 0.680 | 0.863 |
| −1.0 | 0.397 | 0.563 | 0.572 | 0.332 | 0.332 | 0.602 | 0.380 | 0.529 | 0.290 | 0.586 |
| ±0.0 | 0.202 | 0.244 | 0.258 | 0.072 | 0.073 | 0.254 | 0.080 | 0.215 | 0.059 | 0.254 |
| +1.0 | 0.421 | 0.575 | 0.582 | 0.336 | 0.339 | 0.615 | 0.382 | 0.594 | 0.323 | 0.618 |
| +2.0 | 0.774 | 0.893 | 0.880 | 0.703 | 0.705 | 0.900 | 0.805 | 0.902 | 0.745 | 0.905 |
| +3.0 | 0.950 | 0.981 | 0.967 | 0.876 | 0.880 | 0.970 | 0.940 | 0.969 | 0.909 | 0.976 |

(*continues*)

Table 6 (continues).

Panel C: CAR(-5,+5)

| AR | ORDIN | PATELL | BMP | CAMPBELL-WASLEY | CUM-RANK-T | CUM-RANK-Z | GRANK | SIGN-COWAN | SIGN-GSAR-T | SIGN-GSAR-Z |
|---|---|---|---|---|---|---|---|---|---|---|
| −3.0 | 0.591 | 0.775 | 0.779 | 0.319 | 0.339 | 0.627 | 0.603 | 0.689 | 0.456 | 0.730 |
| −2.0 | 0.405 | 0.554 | 0.577 | 0.234 | 0.246 | 0.516 | 0.348 | 0.480 | 0.240 | 0.532 |
| −1.0 | 0.244 | 0.322 | 0.355 | 0.107 | 0.113 | 0.341 | 0.140 | 0.280 | 0.104 | 0.337 |
| ±0.0 | 0.197 | 0.221 | 0.247 | 0.053 | 0.055 | 0.247 | 0.083 | 0.204 | 0.062 | 0.257 |
| +1.0 | 0.278 | 0.350 | 0.363 | 0.123 | 0.129 | 0.355 | 0.182 | 0.331 | 0.137 | 0.392 |
| +2.0 | 0.450 | 0.586 | 0.609 | 0.248 | 0.260 | 0.567 | 0.425 | 0.594 | 0.325 | 0.623 |
| +3.0 | 0.654 | 0.797 | 0.801 | 0.363 | 0.377 | 0.721 | 0.662 | 0.810 | 0.545 | 0.806 |

Panel D: CAR(-10,+10)

| AR | ORDIN | PATELL | BMP | CAMPBELL-WASLEY | CUM-RANK-T | CUM-RANK-Z | GRANK | SIGN-COWAN | SIGN-GSAR-T | SIGN-GSAR-Z |
|---|---|---|---|---|---|---|---|---|---|---|
| −3.0 | 0.442 | 0.583 | 0.638 | 0.189 | 0.216 | 0.468 | 0.387 | 0.486 | 0.244 | 0.587 |
| −2.0 | 0.327 | 0.442 | 0.462 | 0.146 | 0.170 | 0.401 | 0.227 | 0.337 | 0.136 | 0.381 |
| −1.0 | 0.247 | 0.295 | 0.319 | 0.104 | 0.111 | 0.306 | 0.128 | 0.227 | 0.072 | 0.280 |
| ±0.0 | 0.206 | 0.239 | 0.249 | 0.064 | 0.075 | 0.243 | 0.087 | 0.214 | 0.068 | 0.257 |
| +1.0 | 0.237 | 0.287 | 0.315 | 0.083 | 0.098 | 0.299 | 0.145 | 0.300 | 0.104 | 0.339 |
| +2.0 | 0.339 | 0.451 | 0.476 | 0.130 | 0.152 | 0.410 | 0.266 | 0.447 | 0.190 | 0.469 |
| +3.0 | 0.468 | 0.614 | 0.626 | 0.187 | 0.209 | 0.488 | 0.443 | 0.619 | 0.330 | 0.639 |

# 7 DISCUSSION

The following section summarizes the findings of the previous sections on the event study test statistics. In addition, recommendations for further research are made. The previous sections presented new nonparametric testing methods for testing cumulative abnormal returns (CARs). The main focus was on the nonparametric test statistics, because they do not, for example, assume the stock returns to be normally distributed [e.g. Cowan (1992)]. Earlier studies [e.g. Fama (1976)] have found that the distributions of daily returns are fat-tailed relative to a normal distribution. Therefore, the use of nonparametric test statistics is justified. However, most of the parametric test statistics can be applied for testing CARs too. Hence, this study focused on deriving the nonparametric rank and sign test statistics for testing CARs.

In Section 2 the general background of the nonparametric testing methods was presented. First, the definitions of the nonparametric testing methods were presented and the conclusion is that the definition of nonparametric varies slightly depending on the author. Second, the general advantages and disadvantages of the nonparametric testing methods were described.

Section 3 dealt with the background of the event study testing methods. First, the overview of the history of the event study was reported. Second, the outline of the event study was presented. Third, the widely known parametric and nonparametric event study test statistics were described. Fourth, the testing for CARs was described. Section 3 concluded that the new rank and sign procedures derived in Section 4 make nonparametric tests available for general application to the mainstream of event studies.

Section 4 shows that the variance estimator in the test statistic derived by Campbell and Wasley (1993), called CAMPBELL-WASLEY, is biased and a new test statistic CUMRANK-T, based on corrected variance estimator, is suggested instead of the test statistic CAMPBELL-WASLEY. Also the test statistic CUMRANK-Z is presented as a modified version of the test statistic introduced by Corrado and Truong (2008) for scaled ranks. The generalized standardized abnormal returns (GSARs), derived by Kolari and Pynnönen (2011), are used to extend the sign test in Corrado and Zivney (1992) for testing CARs. The resulting new test statistics are called SIGN-GSAR-T and SIGN-GSAR-Z. Consequently, Section 4 presented the rank test statistics CUMRANK-T, CUM-RANK-Z and CAMPBELL-WASLEY as well as the sign test statistics SIGN-GSAR-T and SIGN-GSAR-Z.

Section 4 also included the presentation of the theoretical distributions of these nonparametric test statistics. The main focus in this study has been on the test statistics CUMRANK-T and SIGN-GSAR-T. For a fixed estimation period the asymptotic distributions of those test statistics are shown to be Student's $t$-distributions with $T-2$ degrees of freedom, where $T$ is the length of the time series considered. It was also found that for large $T$ all the null distributions of the statistics CAMPBELL-WASLEY, CUMRANK-T, CUMRANK-Z, SIGN-GSAR-T and SIGN-GSAR-Z can be approximated by the standard normal distribution. The theoretical derivations indicated that when the event-dates are clustered, the test statistics CAMPBELL-WASLEY, CUMRANK-T and SIGN-GSAR-T behave better than the test statistics CUMRANK-Z and SIGN-GSAR-Z. Hence, the theoretical derivations of Section 4 indicated that the rank test statistic CUMRANK-T seems to have better properties than the rank test statistics CAMPBELL-WASLEY and CUMRANK-Z, and the sign test statistic SIGN-GSAR-T seems to have better properties than the sign test statistic SIGN-GSAR-Z. Therefore, the theoretical derivations suggest use of the test statistics CUMRANK-T and SIGN-GSAR-T in future research especially in cases where the event dates are clustered.

The simulation design was presented in Section 5. The simulation approach is similar to that widely used in several other methodological studies. The sample constructions of the simulations study were presented together with the abnormal return model and data used in the simulation study. Section 5 also presented the other test statistics used in the simulation study, including the parametric test statistics ORDIN, PATELL and BMP, and the nonparametric test statistics SIGN-COWAN and GRANK.

Section 6 presented the empirical results of the simulation study. The results of Section 6 are promising especially regarding the new test statistics CUMRANK-T and SIGN-GSAR-T, because those statistics seem to also behave well in the empirical simulations. In those empirical simulations the main focus was on four different points. First, the sample statistics of the test statistics were studied. Second, the empirical distributions of the test statistics were investigated. Third, the rejection rates were studied. Fourth, the power properties of the test statistics were investigated.

While studying the sample statistics of the test statistics, it was noted that focusing only on single day abnormal returns, $\mathrm{AR}(0)$, the means of all the test statistics are statistically close to zero. In longer event windows the means of the test statistics, albeit small, start to deviate significantly away from the theoretical value of zero. The means of the parametric test statistics PATELL and BMP deviate more rapidly from the theoretical value of zero than their nonparametric counterparts. Nonetheless, the standard deviations of the all test statistics are quite close to unity as expected. Therefore, according to the sample statistics of the simulation study, the nonparametric test statistics seem to behave at least as well as the parametric test statistics.

The results of the empirical distributions indicated that particularly for short CAR-windows a sample size of $n = 50$ series seems to be large enough to warrant the asymptotic distributions of the nonparametric test statistics CAMPBELL-WASLEY, CUMRANK-T, CUMRANK-Z, GRANK and SIGN-GSAR-T. It is notable that the Cramer-von Mises test rejects the normality of the test statistic SIGN-GSAR-Z for both short and long CAR-windows. The shortened estimation period does not have such a strong effect on the rejection of the Cramer-von Mises test statistics for other test statistics except the parametric test statistics ORDIN.

When studying the rejection rates, it was noted that the SIGN-GSAR-Z test statistic over-rejects the null hypothesis in every case. Therefore, it seems that the tails of the test statistic SIGN-GSAR-Z are fat, which maybe the reason why the Cramer -von Mises test statistics rejected the normality of the test statistic SIGN-GSAR-Z. The CAMPBELL-WASLEY test statistic seems to under-reject the null hypothesis for longer CAR-windows for the upper tail and two-tailed tests. Therefore, consistent with the theoretical derivations, the simulation results with actual returns confirmed that in longer CAR-windows the test statistic CUMRANK-T tends to reject the null hypothesis closer to the nominal rate than the rank test based approach suggested in Campbell and Wasley (1993). The Campbell and Wasley statistic suffers from a small technical bias in the standard error of the statistic that does not harm the statistic in short CAR-windows, but causes under-rejection of the null hypothesis in longer CAR-windows. The test statistics CUMRANK-T, CUMRANK-Z, GRANK, SIGN-COWAN and SIGN-GSAR-T seem to reject close to the nominal rate in almost every case, and also to be robust against volatility increases. The simulation results also revealed that ORDIN and PATELL tests over-reject the variance increases, which is a well-known outcome. Generally the rejection rates are not very sensitive to the length of the estimation period, but the length of the estimation period seems to substantially affect the under-rejection of CAMPBELL-WASLEY in the long CAR-windows. Hence, according to the rejection rates at least, the nonparametric test statistics CUMRANK-T and SIGN-GSAR-T together with the test statistics CUMRANK-Z, GRANK and SIGN-COWAN seem to behave quite well.

Both CUMRANK-T and SIGN-GSAR-T together with the other nonparametric test statistics seem to have good empirical power properties. The test statistics CUMRANK-T and SIGN-GSAR-T, in addition to the test statistics CAMPBELL-WASLEY and GRANK, have also the advantage of being quite robust to cross-correlation (clustered event days) of the returns. Test statistics ORDIN, PATELL, BMP, CUMRANK-Z, SIGN-COWAN and SIGN-GSAR-Z seem to over-reject the null hypothesis when the event days are clustered. The power comparisons of the test statistics CAMPBELL-WASLEY, CUMRANK-T and GRANK for longer CAR-windows should be made with caution when the event days are clustered, because those test statistics also somewhat over-reject the null hypothesis for longer CAR-windows. It is notable that the over-rejection of those test statistics is not so clear as the over-rejection of the test statistics ORDIN, PATELL, BMP, CUMRANK-Z, SIGN-COWAN and SIGN-GSAR-Z. According to the empirical simulations, when the event days are clustered the test statistic SIGN-GSAR-T is the most optimal choice, because it is the only one, which rejects close to the nominal rate with rejection rates that are well within the approximate 99 percent confidence interval of $[0.032, 0.068]$.

The previous sections introduced many new results and opened up new perspectives on viewing event study tests. For practical reasons, the event studies have often focused on parametric test statistics, because the parametric test statistics are usually quite widely known, easy to use and can usually be used both for single day abnormal returns as well as for cumulative abnormal returns. As demonstrated in the previous sections, however, the parametric test statistics are not usually the optimal choices. The empirical results of this study indicated that some attention should also be paid to the data which is investigated. The simulation results indicate that, for example, if the returns are not normally distributed, there exists event induced variance or the event days are clustered, the new nonparametric test statistics CUMRANK-T and SIGN-GSAR-T would be more effective choices than the traditional parametric test statistics, such as ORDIN, PATELL and BMP. Hence, the results of this study suggest choosing the event study testing method according to the situation, and trusting that in some cases nonparametric test statistics can be a better choice than traditional parametric test statistics.

# REFERENCES


Ball, R. & Brown, P. (1968). An empirical evaluation of accounting income numbers. *Journal of Accounting Research* 6, 159–177.

Barber, B. M. & Lyon, J. D. (1997). Detecting long-run abnormal returns: The empirical power and specification test statistics. *Journal of Financial Economics* 81, 175–213.

Beaver, W. H. (1968). The information content of annual announcements. *Journal of Accounting Research Supplement* 6, 67–92.

Berry, M. A., Gallinger, G. W. & Henderson, G. V. (1990). Using daily stock returns in event studies and the choice of parametric versus nonparametric test statistics. *Quarterly Journal of Business and Economics* 29, 70–85.

Boehmer, E., Musumeci J. & Poulsen, A. B. (1991). Event-study methodology under conditions of event-induced variance. *Journal of Financial Economics* 30, 253–272.

Bradley, J. V. (1968). *Distribution-free Statistical Tests*. New Jersey: Prentice-Hall, Inc.

Brown, S. J. & Warner, J. B. (1980). Measuring security price performance. *Journal of Financial Economics* 8, 205–258.

Brown, S. J. & Warner, J. B. (1985). Using daily stock returns: The case of event studies. *Journal of Financial Economics* 14, 3–31.

Campbell, J. Y., Lo, A. W. & MacKinlay, A. C. (1997). *The Econometrics of Financial Markets*. Princeton, NJ: Princeton University Press.

Campbell, C. J. & Wasley C. E. (1993). Measuring security performance using daily NASDAQ returns. *Journal of Financial Economics* 33, 73–92.

Campbell, C. J. & Wasley C. E. (1996). Measuring abnormal trading volume for samples of NYSE/ASE and NASDAQ securities using parametric and non-parametric test statistics. *Review of Quantitative Finance and Accounting* 6, 309–326.

Charest, G. (1978). Dividend information, stock returns, and market efficiency. *Journal of Financial Economics* 6, 265–296.

Cleveland, W. S. (1993). *Visualizing Data*. USA: AT&T.

Corrado, C. J. (1989). A nonparametric test for abnormal security price performance in event studies. *Journal of Financial Economics* 23, 385–395.

Corrado, C. J. (2010). Event studies: a methodology review. *Accounting & Finance* 51, 207–234.

Corrado, C. J. & Truong, C. (2008). Conducting event studies with Asia-Pasific security market data. *Pacific-Basin Finance Journal* 16, 493–521.

Corrado, C. J. & Zivney, T. L. (1992). The specification and power of the sign test in event study hypothesis test using daily stock returns. *Journal of Financial and Quantitative Analysis* 27, 465–478.

Cowan, A. R. (1992). Nonparametric event study tests. *Review of Quantitative Finance and Accounting* 2, 343–358.

Cowan, A. R. & Sergeant, A. M. A. (1996). Trading frequency and event study test specification. *Journal of Banking & Finance* 20, 1731–1757.

Daniel, W. W. (1990). *Applied Nonparametric Statistics*. Boston: PWS-Kent.

Dechow, P., Sloan, R. & Sweeney, A. (1995). Detecting earnings management. *Accounting Review* 70, 3–42.

Dolley, J. C. (1933). Characteristics and procedure of common stock split-ups. *Harvard Business Review* 11, 316–326.

Dyckman, T., Philbrick, D. & Stephan, J. (1984). A Comparison of Event Study Methodologies. *Journal of Accounting Research* 22, 1-33.

Fama, E. F. (1976). *Foundations of Finance*. New York: Basic Books.

Fama, E. F., Fischer, L., Jensen, M. & Roll, R. (1969). The adjustment of stock prices to new information. *International Economic Review* 10, 1–21.

Gibbons, J. D. (1976). *Nonparametric Methods for Quantitative Analysis*. University of Michigan: Holt, Rinehart and Winston.

Gibbons, J. D., & Chakraborti, S. (1992). *Nonparametric Statistical Inference*. New York: Marcel Dekker Inc.

Harrington, S. E. & Shrider, D. G. (2007). All events induce variance: Analyzing abnormal returns when effects vary across firms. *Journal of Financial and Quantitative*

*Analysis* 42, 229–256.

Hettmansperger, T. P. & McKean, J. W. (2011). *Robust Nonparametric Statistical Methods*. USA: CRC press.

Kolari, J. W. & Pynnönen, S. (2010). Event study testing with cross-sectional correlation of abnormal returns. *Review of Financial Studies* 23, 3996–4025.

Kolari, J. W. & Pynnönen, S. (2011). Nonparametric Rank Tests for Event Studies. *Journal of Empirical Finance* (Forthcoming). DOI:10.1016/j.jempfin.2011.08.003

Kothari, S. P., Leone, A. & Wasley, C. (2005). Performance-matched discretionary accruals. *Journal of Accounting and Economics* 39, 163–197.

Kothari, S. P. & Warner, J. B. (2007). Econometrics of event studies, Espen Eckbo (ed.). *Handbook of Corporate Finance: Empirical Corporate Finance*. The Netherlands: North-Holland.

Lehmann, E. L. (1999). *Elements of Large-Sample Theory*, New York, Springer-Verlag.

Mann, H. B. & Whitney, D. R. (1947). On a test whether one of two random variables is stochastically larger than the other. *Annals of Statistics* 18, 50–60.

Maynes, E. & Rumsey, J. (1993). Conducting event studies with thinly traded stocks. *Journal of Banking and Finance* 17, 145–157.

Mikkelson, W. H. (1981). Convertible calls and security returns. *Journal of Financial Economics* 9, 237–264.

Patell, J. A. (1976). Corporate forecasts of earnings per share and stock price behavior: Empirical test. *Journal of Accounting Research* 14, 246–276.

Penman, S. H. (1982). Insider trading and the dissemination of firms forecast information. *Journal of Business* 55, 479–503.

Pynnönen, S. (2010). The joint distribution of a linear transformation of internally studentized least squares residuals. *Working Papers of the University of Vaasa, Department of Mathematics and Statistics* 16. Cited 27.09.2011. Available:
http://www.uwasa.fi/julkaisusarjat/luettelo/workingpapers/

Rosenstein, S. & Wyatt, J. G. (1990). Outside directors, board independence, and shareholders wealth. *Journal of Financial Economics* 26, 175–191.

Siegel, S. (1956). *Nonparametric Statistics for the Behavioral Sciences*. Tokyo: McGraw-Hill Book Company, inc.

Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics Bulletin* 1, 80–83.

# A APPENDIX

In this appendix the properties of the sign of the GSAR are shown. Therefore, the theoretical expectation and variance of $G_{it}$ as well as the theoretical covariance between $G_{it}$ and $G_{is}$, $t \neq s$, $t,s = 1,\ldots,T$, in both of the cases $T = L_1 + 1$ being even and odd, are derived. Using equations (48) to (51) it is straightforward to see that

$$E[G_{it}] \quad = \quad 0 \tag{A.1}$$

and

$$\mathrm{var}[G_{it}] = \begin{cases} 1, & \text{for even } T \\ \frac{T-1}{T}, & \text{for odd } T. \end{cases} \tag{A.2}$$

Again, if $t \neq s$, it is straightforward to verify the following probabilities

$$\Pr[G_{it}G_{is} = 1] = \begin{cases} \frac{\frac{T}{2}-1}{T-1}, & \text{for even } T \\ \frac{T-3}{2T}, & \text{for odd } T, \end{cases} \tag{A.3}$$

$$\Pr[G_{it}G_{is} = 0] = \begin{cases} 0, & \text{for even } T \\ \frac{2}{T}, & \text{for odd } T \end{cases} \tag{A.4}$$

and

$$\Pr[G_{it}G_{is} = -1] = \begin{cases} \frac{\frac{T}{2}}{T-1}, & \text{for even } T \\ \frac{T-1}{2T}, & \text{for odd } T. \end{cases} \tag{A.5}$$

Furthermore for $T$ being even

$$\mathrm{cov}[G_{it}, G_{is}] = E[G_{it}G_{is}] = -\frac{1}{T-1} \tag{A.6}$$

and for $T$ being odd

$$\mathrm{cov}[G_{it}, G_{is}] = E[G_{it}G_{is}] = -\frac{1}{T}. \tag{A.7}$$

# B APPENDIX

In this appendix the proofs for the Theorem 4 and the Theorem 5 are given. In these proofs, $T = T_2 - T_0$, when focusing on the rank test statistics ($Z_3$ and $Z_4$). While, $T = T_1 - T_0 + 1$, when focusing on the sign test statistics ($Z_6$ and $Z_5$). The following lemmas are utilized in the proofs of the Theorem 4 and the Theorem 5 and the proofs of these lemmas can be obtained as special cases from Pynnönen (2010).

**Lemma 1.** *Define*

$$x = \mathbf{Q}y,$$

*where* $\mathbf{Q}$ *is a* $T \times T$ *idempotent matrix of rank* $r \leq T$ *and* $y = (y_1, \ldots, y_T)'$ *is a vector of independent* $N(0,1)$ *random variables, such that* $y \sim N(0, \mathbf{I})$, *where* $\mathbf{I}$ *is a* $T \times T$ *identity matrix. Furthermore, let* $m$ *be a* $T$ *component column vector of real numbers such that* $m'\mathbf{Q}m > 0$. *Then*

$$z_m = \frac{m'x/\sqrt{m'\mathbf{Q}m}}{\sqrt{x'x/r}} \tag{B.1}$$

*has the distribution with density function*

$$f_{z_m}(z) = \frac{\Gamma(r/2)}{\Gamma[(r-1)/2]\sqrt{r\pi}} \left(1 - \frac{z^2}{r}\right)^{\frac{1}{2}(r-1)-1}, \tag{B.2}$$

*when* $|z| < \sqrt{r}$, *and zero otherwise, where* $\Gamma(\cdot)$ *is the gamma function.*

**Lemma 2.** *Under the assumptions of Lemma* 1

$$t_m = z_m \sqrt{\frac{r-1}{r - z_m^2}} \tag{B.3}$$

*is distributed as the Student t-distribution with* $r-1$ *degrees of freedom.*

The proofs of the following theorems are adapted from Kolari and Pynnönen (2011).

*Proof of Theorem* 4*:* In order to derive the asymptotic distribution of the modified CAMPBELL-WASLEY statistic, $Z_3$, defined in equation (41), the deviations of cross-sectional average abnormal (scaled) ranks $K_{it}$ defined in (14) from their expected values, $1/2$, are collected to a column vector $d_i = (d_{i,1}, d_{i,2}, \ldots, d_{i,T})'$, where $d_{i,t} = K_{i,T_0+t} - 1/2$, $t = 1, \ldots, T = T_2 - T_0$. The prime denotes transpose and $i = 1, \ldots, n$ with $n$ the number of series. Similarly, in order to derive the asymptotic distribution of the $Z_6$ defined in equation (57), the $G_{it}$s defined in (47) are collected to a column vector $G_i = (G_{i,T_1}, G_{i,T_2}, \ldots, G_{i,T})'$, where $T = T_1 - T_0 + 1$. The prime denotes again transpose and $i = 1, \ldots, n$ with $n$ the number of series. Then by assumption the random vectors $d_i$s and $G_i$s are independent and, by Proposition 1 and Proposition 4, identically distributed random vectors such that

$$E[d_i] = 0, \tag{B.4}$$

$$E[G_i] = 0, \tag{B.5}$$

$$\mathrm{cov}[d_i] = \frac{T-1}{12(T+1)}\left((1-\rho)\mathbf{I} + \rho\,\iota\iota'\right) \tag{B.6}$$

and

$$\mathrm{cov}[G_i] = \begin{cases} (1-\rho)\mathbf{I} + \rho\iota\iota', & \text{for even } T \\ \frac{T-1}{T}[(1-\rho)\mathbf{I} + \rho\iota\iota'], & \text{for odd } T. \end{cases} \tag{B.7}$$

Again $i = 1, \ldots, n$. Furthermore $\iota$ is a vector of $T$ ones, $\mathbf{I}$ is a $T \times T$ identity matrix, and

$$\rho = -\frac{1}{T-1}. \tag{B.8}$$

Thus, the covariance matrixes in (B.6) and (B.7) become

$$\mathrm{cov}[d_i] = \frac{T}{12(T+1)}\left(\mathbf{I} - \frac{1}{T}\iota\iota'\right) \tag{B.9}$$

and

$$\mathrm{cov}[G_i] = \begin{cases} \frac{T}{T-1}\left(\mathbf{I} - \frac{1}{T}\iota\iota'\right), & \text{for even } T \\ \left(\mathbf{I} - \frac{1}{T}\iota\iota'\right), & \text{for odd } T. \end{cases} \tag{B.10}$$

It should be noted that the matrix $\mathbf{I} - T^{-1}\iota\iota'$ is an idempotent matrix of rank $T-1$, which implies that covariance matrixes (B.9) and (B.10) are singular. However, because $d_i$s and $G_i$s are independent with zero means and finite covariance matrixes (B.9) and (B.10), the Central Limit Theorem applies such that

$$\sqrt{n}\bar{d} \xrightarrow{d} \left(\frac{T}{12(T+1)}\right)^{\frac{1}{2}} x, \tag{B.11}$$

$$\sqrt{n}\bar{G} \xrightarrow{d} \left(\frac{T}{T-1}\right)^{\frac{1}{2}} x, \tag{B.12}$$

when $T$ is even and

$$\sqrt{n}\bar{G} \xrightarrow{d} x, \tag{B.13}$$

when $T$ is odd, as $n \to \infty$, where

$$x \sim N(0, \mathbf{Q}), \tag{B.14}$$

with the (idempotent) singular covariance matrix

$$\mathbf{Q} = \mathbf{I} - \frac{1}{T}\iota\iota'. \tag{B.15}$$

In (B.11) $\bar{d} = (\bar{d}_1, \ldots \bar{d}_T)'$ with

$$\bar{d}_t = \frac{1}{n}\sum_{i=1}^{n} d_{it} = \frac{1}{n}\sum_{i=1}^{n}\left(K_{i,T_0+t} - \frac{1}{2}\right), \tag{B.16}$$

the time $t$ cross-sectional mean of the deviations $K_{i,T_0+t} - 1/2$ of the scaled ranks, $t = 1, \ldots, T$. Note that the sum of $d_{i,t}$ over the time index $t$ is zero for all $i = 1, \ldots, n$, i.e., $\iota' d_i = 0$ for all $i = 1, \ldots, n$, which implies that $\iota'\bar{d} = 0$. In (B.12) and (B.13), $\bar{G} = (\bar{G}_1, \bar{G}_2, \ldots, \bar{G}_T)'$ with

$$\bar{G}_t = \frac{1}{n}\sum_{i=1}^{n} G_{it}, \tag{B.17}$$

where again $t = 1, \ldots, T$. Note that the sum of $G_{i,t}$ over the time index $t$ is zero for all $i = 1, \ldots, n$, i.e., $\iota' G_i = 0$ for all $i = 1, \ldots, n$, which implies that $\iota'\bar{G} = 0$.

Let $\iota_{\tau_1,\tau_2}$ be a column vector of length $T$ with ones in positions in the event window from $\tau_1$ to $\tau_2$ and zeros elsewhere. Then $Z_3$ defined in equation (41) can be written as

$$Z_3 = \frac{\iota'_{\tau_1,\tau_2}\bar{d}}{\sqrt{\bar{d}'\bar{d}}}\sqrt{\frac{T(T-1)}{\tau(T-\tau)}} = \frac{\iota'_{\tau_1,\tau_2}\bar{d}/\sqrt{\tau(T-\tau)/T}}{\sqrt{\bar{d}'\bar{d}/(T-1)}}. \tag{B.18}$$

Defining in Lemma 1

$$m = \iota_{\tau_1,\tau_2} \tag{B.19}$$

and

$$\mathbf{Q} = \mathbf{I} - \frac{1}{T}\iota\iota', \tag{B.20}$$

the following can be obtained

$$m'\mathbf{Q}m = \frac{\tau(T-\tau)}{T}, \tag{B.21}$$

such that the ratio $z_m$ in (B.1) becomes

$$Z_{m1} = \frac{\iota'_{\tau_1,\tau_2}x/\sqrt{\tau(T-1)/T}}{\sqrt{x'x/(T-1)}}. \tag{B.22}$$

Similarly, let $\iota_0$ be a column vector of length $T = T_1 - T_0 + 1$ with one in position in the event day $t = 0$ and zeros elsewhere. Then $Z_6$ in equation (57) can be written as

$$Z_6 = \frac{\iota'_0\bar{G}}{\sqrt{\bar{G}'\bar{G}/T}} = \frac{\iota'_0\bar{G}/\sqrt{(T-1)/T}}{\sqrt{\bar{G}'\bar{G}/(T-1)}}. \tag{B.23}$$

Defining now in Lemma 1

$$m = \iota_0 \tag{B.24}$$

and

$$\mathbf{Q} = \mathbf{I} - \frac{1}{T}\iota\iota', \tag{B.25}$$

following can be written

$$m'\mathbf{Q}m = \frac{(T-1)}{T}, \tag{B.26}$$

such that the ratio $z_m$ in (B.1) becomes

$$Z_{m2} = \frac{\iota'_0x/\sqrt{(T-1)/T}}{\sqrt{x'x/(T-1)}}. \tag{B.27}$$

The distributions of $Z_{m1}$ and $Z_{m2}$, after arranging term, have the density function,

$$f_{Z_{m1}}(z) = f_{Z_{m2}}(z) = \frac{\Gamma[(T-1)/2]}{\Gamma[(T-2)/2]\sqrt{(T-1)\pi}}\left(1 - \frac{z^2}{T-1}\right)^{\frac{1}{2}(T-3)} \tag{B.28}$$

for $|z| < \sqrt{T-1}$ and zero elsewhere.

Because of the convergence results in (B.11), (B.12) and (B.13) and that the functions

$$h(\bar{d}) = \frac{\iota'_{\tau_1,\tau_2}\bar{d}/\sqrt{\tau(T-\tau)/T}}{\sqrt{\bar{d}'\bar{d}/(T-1)}} \tag{B.29}$$

and

$$h(\bar{G}) = \frac{\iota'_0\bar{G}/\sqrt{(T-1)/T}}{\sqrt{\bar{G}'\bar{G}/(T-1)}} \tag{B.30}$$

are continuous, the continuous mapping theorem implies $h(\bar{d}) \xrightarrow{d} h(x)$ and $h(\bar{G}) \xrightarrow{d} h(x)$. That is,

$$Z_3 = \frac{\iota'_{\tau_1,\tau_2}\bar{d}/\sqrt{\tau(T-\tau)/T}}{\sqrt{\bar{d}'\bar{d}/(T-1)}} \xrightarrow{d} \frac{\iota'_{\tau_1,\tau_2}x/\sqrt{\tau(T-1)/T}}{\sqrt{x'x/(T-1)}} = Z_{m1}, \tag{B.31}$$

and

$$Z_6 = \frac{\iota'_0\bar{G}/\sqrt{(T-1)/T}}{\sqrt{\bar{G}'\bar{G}/(T-1)}} \xrightarrow{d} \frac{\iota'_0x/\sqrt{(T-1)/T}}{\sqrt{x'x/(T-1)}} = Z_{m2}, \tag{B.32}$$

which implies that the density function of the limiting distribution of $Z_3$ and $Z_6$ for fixed $T$, as $n \to \infty$, is of the form defined in equation (B.28), completing the proof of Theorem 4.

*Proof of the Theorem* 5*:* By the proof of Theorem 4, $Z_3 \xrightarrow{d} Z_{m1}$, where $Z_{m1}$ is defined in equation (B.22) with $r = T-1$. Similarly, $Z_6 \xrightarrow{d} Z_{m2}$, where $Z_{m2}$ is defined in equation (B.27) with $r = T-1$. Again because the function $g(z) = z\sqrt{(T-2)/(T-1-z^2)}$ is continuous, for $|z| < \sqrt{T-1}$, the continuous mapping theorem implies $Z_4 = g(Z_3) \xrightarrow{d} g(Z_{m1})$ and $Z_5 = g(Z_4) \xrightarrow{d} g(Z_{m2})$. That is,

$$Z_4 \xrightarrow{d} Z_{m1}\sqrt{\frac{T-2}{T-1-Z_{m1}^2}}$$

and

$$Z_5 \xrightarrow{d} Z_{m2}\sqrt{\frac{T-2}{T-1-Z_{m2}^2}},$$

where the distributions of the right hand side expression are by Lemma 2 the $t$-distributions with $T-2$ degrees of freedom, completing the proof of Theorem 5.