ANALYZING EVENT STATISTICS IN CORPORATE FINANCE

Methodologies, Evidences, and Critiques

Jau-Lian Jeng

ANALYZING EVENT STATISTICS IN CORPORATE FINANCE

This page intentionally left blank

Analyzing Event Statistics in Corporate Finance

Methodologies, Evidences, and Critiques

Jau-Lian Jeng

palgrave
macmillan

ANALYZING EVENT STATISTICS IN CORPORATE FINANCE

Softcover reprint of the hardcover 1st edition 2015 978-1-137-39717-1

First published in 2015 by
PALGRAVE MACMILLAN®
in the United States—a division of St. Martin's Press LLC,
175 Fifth Avenue, New York, NY 10010.

Where this book is distributed in the UK, Europe and the rest of the world, this is by Palgrave Macmillan, a division of Macmillan Publishers Limited, registered in England, company number 785998, of Houndmills, Basingstoke, Hampshire RG21 6XS.

Palgrave Macmillan is the global academic imprint of the above companies and has companies and representatives throughout the world.

Palgrave® and Macmillan® are registered trademarks in the United States, the United Kingdom, Europe and other countries.

ISBN 978-1-349-48481-2 ISBN 978-1-137-49160-2 (eBook)
DOI 10.1057/9781137491602

Library of Congress Cataloging-in-Publication Data is available from the Library of Congress.

A catalogue record of the book is available from the British Library.

Design by Newgen Knowledge Works (P) Ltd., Chennai, India.

First edition: February 2015

10 9 8 7 6 5 4 3 2 1

Contents

This page intentionally left blank

Preface

The literature on the corporate finance is so abundant that, ever since the pioneering work of Fama, Fisher, Jesen and Roll (1969), many research articles are published every year for event studies. Although many new methodologies are devised, empirical corporate finance is still struggling from controversies of many issues such as proper data collection, time-horizon decisions, event period determination, and robust statistical procedures. These issues certainly are not easy tasks for any researcher or an enthusiastic graduate student in economics and finance to resolve them promptly. The aim of this book hence is not set to provide answers to all these issues, either. Instead, the intent is to reiterate the necessary cautions in analyses on the event studies in corporate finance and to provide some alternatives. Therefore, this book aims to become a reference for the researchers, graduate students, and professionals who are interested in exploring the possible impacts of corporate finance events rigorously.

Starting from data collection, researchers should be careful in identifying the issues of interest and the representativeness of data set. Before anything starts, one can't be more cautious than to investigate the intrinsic (time-series/cross-sectional) properties of the data introduced either from the sampling schemes or from the nature of the data themselves, particularly associated with financial times series. More importantly, although it is tempting to explore all different models or regularities that possibly can explain the financial data, all models or proposals for these data should provide sufficient verifications from the foundation of financial economics. Or, even more pleasingly, these models should provide some better ones in theories or rigors if possible.

The contents of the book starts from Chapter 1 that surveys (a) the issues of event-oriented constructed data and (b) the inclusion of firm-specific attributes as variables for empirical asset pricing models and leads to the importance of model search for the normal (or expected) returns. The task then is continued in Chapter 2 where diagnostic tests for the nondiversifiability of factor-oriented variables in empirical asset pricing models (for normal returns) are covered. Although various specification tests or model selection criteria have been applied to the empirical asset pricing models, only a few of them emphasize the pre-requisite that these included variables should be nondiversifiable so that separation between normal (or expected) returns versus abnormal returns can be well stated. In essence, many empirical research results on corporate event studies are questioned since the findings can either be the outcomes of events or model specification errors (in normal returns), or both. Criticism results since little justification in (model search for) empirical asset pricing models is done prior to applications of abnormal returns. Chapter 2 in this book attempts to provide some alternatives.

For many studies in corporate finance or financial economics, studies based on abnormal returns are applied in the verifications of events. Conventional studies either apply the cumulative abnormal returns or perform the regressions of abnormal returns on firm-specific attributes to obtain statistical tests. Either way, the hypothesis of interest is to verify that whether the mean of abnormal returns is nonzero within the event window or not. This in turns, shows that the hypothesis is equivalent to verification of structural changes of means particularly, in abnormal returns. If empirical asset pricing models with regressions are applied to obtain both the normal and abnormal returns, the hypothesis of interest is equivalent to a structural change of parameter(s) in regressions over the event window. Given so, it is not difficult to see that the conventional tests such as cumulative abnormal returns (CARs) (with normality or not) are similar to the CUSUM (cumulative sums) tests for parameter changes—at least, from the perspective of hypotheses.

Although the asymptotic arguments of CUSUM tests are mostly based on Brownian motion, the applications and hypotheses of CUSUM tests coincide with those of CARs tests in using the cumulative sums of abnormal returns or residuals from regressions of empirical asset pricing models. Chapter 3 argues that the deficiency of conventional CARs tests in event studies is bound to appear since financial time series almost always contain parameter changes even if there is no significant event at all. Hence, using CARs tests is not entirely robust to verify the impacts of events—even though the empirical asset pricing models for normal returns are selected properly. The same critique applies to the CUSUM-based (monitoring) tests for parameter changes of empirical asset pricing models as well. It is hardly conclusive to consider that parameter changes are solely the results of firm-specific events. Many financial time series (and the proposed models in using them) are usually subject to time-varying parametrization. These arguments show that the tests simply base on the mean changes of abnormal returns are not sufficient enough to tell whether the impacts from the events are significant or not. Some alternatives must be provided to investigate the strength of impacts from events without incurring the similarity between CARs tests and CUSUM tests for structural changes.

Chapter 4 is prepared as a prerequisite for Chapter 5. Given the model identified in model search, the possible time-varying coefficients are taken into account. The intuition is, if the capital market adjustments are successfully efficient, the updating mechanism should possibly incorporate concurrent information into the normal (or expected) returns. Likewise, in forming the normal (or expected) returns, the recursive estimation should be introduced to accommodate the time-changing expectations of the market. Notice that these updating procedures apply only the systematic (nondiversifiable) information provided from the model(s) identified earlier. Therefore, the residuals resulted from the time-varying coefficient models of normal (or expected) returns approximate closer to the abnormal returns as planned. Loosely speaking, this approach provides a better classification between normal and abnormal returns. In addition, the recursive algorithms applied do not assume some particular estimation periods or, event windows. Hence, less subjective classifications are introduced into the statistics (and the functionals) of abnormal returns.

Following from Chapter 4, Chapter 5 provides an alternative methodology where the intensity of cumulative abnormal returns is represented as the duration of the absolute values of cumulative abnormal returns in crossing certain levels of thresholds. In other words, instead of using the CARs tests (or CUSUM tests), the interest of research is on the time horizon that the absolute values of these cumulative abnormal returns may surpass the thresholds. Under the assumption of invariance principles where these cumulative abnormal returns (after normalization) may converge to the Brownian motion, these durations (of different thresholds) may converge (in distribution) to the so-called occupation time (or sojourn time) of reflected Brownian motion under the null. Hence, using the asymptotic distribution (and moments) obtained by Takacs (1998), the test statistics can be formed to identify if the duration of level crossings of these cumulative abnormal returns is significantly different from the occupation time of reflected Brownian motion. This provides an alternative to consider the event studies in corporate finance where tests over the possibly time-varying coefficients in financial time series are not in need.

Acknowledgments

This book will never be possible at all if without encouragement, assistance, and comments of many peoples. In particular, my wife Fang who was studying earnestly to finally pass all her CPA exams, dedicated to support me passionately, especially when difficulties developed along the path of writing. Help from Palgrave Macmillan is also gratefully appreciated. Editors Leila Campoli and Sarah Lawrence helped me with so much of timely advices to finish the project. The ex-editor Brian Foster provided me with this opportunity to pursue a totally different perspective on event studies in corporate finance. Many thanks are also offered to the participants and discussants in the 2013 Western Economic Association Conference in Seattle. Dr. Jack Hou's comments and advice are also gratefully acknowledged. Their comments and encouragements triggered the conception of this book. Moreover, I am also deeply grateful to the one who said, "...I am!".

JAU-LIAN JENG, Ph.D.
School of Business and Management
Azusa Pacific University

Part I

Event Study Methodology I

Chapter 1

Data Collection in Long-Run or Short-Run Format?

Introduction

In this chapter, a critical question is raised for the empirical finance of corporate event studies. That is, what kind of data set should one apply? Should the short-run data set such as daily returns (or even high-frequency data) be applied? Or, should one try with the longer horizon data? A painful browse through all related literature shows that it is easy to find that there is no definite rule applied to this issue. One question often asked is whether the short-run returns contain more updated information or, the longer horizon data that may provide more insightful views since the impacts of corporate events may be persistent over time.

For the issues of mergers and acquisitions in particular, the controversies over the choices of data frequencies and study horizons remain unresolved even after decades since the pioneering study of Fama et al. (1969) on corporate finance. Various empirical findings of event studies in corporate finance can be found in many leading finance journals and elsewhere. Yet, these issues remain mainly unsettled. Throughout this chapter, these differences in sampling and data frequencies are surveyed and the related critiques are offered while using the event studies of mergers and acquisitions as the examples to depict the issues.

Some rules of thumb for the data selection issues are provided although they remain preliminary. The intent is to enlist some possible basic criteria to these empirical finance issues together so that consistency in analyses among them can possibly be fetched. First of all, the determination of sampled data frequencies should take into account the feasibility of social-ecnonmic/financial data and the relevancy of theoretical foundation of financial economics. In other words, reasoning with economic/financial theoretical arguments must take precedence when considering the search of data and the statistical methods to apply. In addition, the modeling for normal returns (especially for robust model specifications on empirical asset pricing models) should be cautiously examined prior to the event studies in using the abnormal returns. Lastly, stochastic properties of the data sampled (or constructed) should be thoroughly investigated and checked (using diagnostic tests or else) before any attempt to elaborate the hypotheses of interest.

Not surprisingly, shovelling data series with current computational facilities (or techniques) seems trivial among the empirical finance issues. Fabrication of intended results can be obtained using skills in data manipulation. The essence of event studies of corporate finance issues therefore, is not to present some eye-catching representations in showing the startling results of empirical finance. Instead, presenting the fact-related discussions based on robust specification, and devising some sound guidance for the finance professionals (academicians or practitioners) in developing analyses and making proper decisions are essentially needed in the future for empirical corporate finance.

1.1 Samples, Data Formats and Variables Selection

For many event studies in corporate finance, data collection becomes one of the most formidable tasks to study the hypotheses of interest. For instance, it is overwhelmingly evident that in financial economics/econometrics literature on mergers and acquisitions, many works endeavor to study the stock returns from the events either for the acquirers or the targets. While it is interesting, different data sets or frequencies of data collected may cause various empirical results or conclusions. Unfortunately, there is no definite rule to determining the length of sampled periods, event windows, data frequencies, and the sampling criteria. A challenging task for empirical finance in event studies of corporate finance may start from the decision of time horizon of the studies, the frequency of data, and the collection for event study data. Specifically, two major techniques for event studies in mergers and acquisitions are commonly employed: short-horizon event studies surrounding the merger announcements and long-horizon market valuation studies that discuss the benefits and consequences of the mergers and acquisitions. However, determining the time spans for these stock returns on corporate event issues is never straightforward. Roughly classified, some clarifications of the findings in these conventional event studies (in mergers and acquisitions, for instance) can be shown as follows:

1. the determination of time horizons and data frequencies depends on the possible hypotheses on persistence of the impact(s) and the feasibility of attributes/variables that are applicable,
2. the constructed data streams (whether from simulated series or from the raw data) for these hypotheses of interest may have some particular stochastic or statistical properties that can influence the statistical results of studies,
3. regardless of the time frames of studies, a correctly-specified model for normal (or expected) returns with systematic attributes or variables (based on capital market equilibrium) should be devised so that robust assessments on abnormal returns can be obtained for corporate event studies.

Although some statistical assumptions can be applied for the normal returns (especially for the short-horizon returns), incorporation with both statistical assumptions and economic modeling provides the essential features and explanatory specification for the normal returns. Nevertheless, one common understanding is that less attributes or variables may be feasible for the short-horizon specification of normal (expected) returns in event studies. Moreover, as pointed out in Mackinlay (1997), difficulty arises in that even if there are (economic) attributes/variables available to specify the normal returns, the explanatory power of these variables (in empirical findings) is not sufficient enough as to reduce an essential proportion of variance of abnormal returns. Hence, the results in Brown and

Warner (1980, 1985) show that even the simplest constant-mean return model provides similar results for the event studies when compared with more sophisticated models.

As stated in Bremer et al. (2011), "...In essence, short-term event studies focus on the impact of new information on the current expectation of future returns while long-term event studies focus on the ultimate effect of the changes transmitted in the information release on future returns." Although inclusively stated above, most criticism on corporate finance event studies also raises the questions on the frequency of the data stream applied. The issues on the frequencies of stock returns such as daily, weekly, monthly, or else are still left unresolved. The critical fact is that in either short-horizon or long-horizon event studies, constructing the data series of stock returns contains certain stochastic properties (for instance, serial/cross-sectional dependence, heteroscedasticity, and skewness) that some robust methods must be devised to obtain verifiable results. In addition, one of the most formidable tasks in event studies is to cleanly separate the expected and unexpected components in stock returns as stated in Fama et al. (1969). Hence, the conventional methods in corporate finance event studies may (most likely) be criticized and denoted as joint tests for both model specifications (for expected components in stock returns) and the impacts of new information (presented in the unexpected abnormal components in stock returns).

The difficulty in selecting the short-run stock returns is that these returns are noisy and possibly contaminated with information or feedback irrelevant with the event(s). And that if the acquisition of abnormal returns is the interest of study, some robust high-frequency adaptive filtering (with asset-pricing models or else) must be devised to capture the essential systematic components that represent the market expectation (conditional on the available information set or event dates) in these stock returns. More dilemmas may appear if determination of event windows for the impact of event(s) is considered. On the contrary, if applications of long-run stock returns are applied, the difficulty will lie on the data construction and on especially, the available variables or methods to obtain the systematic components in the long horizon. In brief, the sampling issues in corporate event studies focus on (1) the hypothesis on effect(s)/impacts of interest, (2) the time horizon where the impact may persist, and (3) the availability of economic variables and the other (corporate) attributes for the horizon of studies.

Bremer et al. (2011) propose the usage of non-overlapping quarterly compounded stock returns for long-horizon event studies. Even though their simulations are robust, the results are based on the conventional test statistics where the above data properties still influence the results. Hence, to obtain a robust event study on corporate finance issue such as mergers and acquisitions, two decisions must be made: the horizon of event studies and data frequencies, and the model specifications of the expected rates of returns. More essentially, these two decisions are mutually related in most studies.

Following the setting of Strong (1992), the conventional event studies in corporate finance demonstrate that there exists some impacts on the share prices when some firm-specific events are disclosed. Specifically, under the null hypothesis that there is no impact on the stock returns, it is shown that

$$f(r_{it}|y_i) - f(r_{it}) = 0, \tag{1.1.1}$$

where $f(r_{it}|y_i)$ stands for the conditional distribution of r_{it} when the information signal y_i is declared, $f(r_{it})$ stands for the marginal distribution of r_{it}. If the events are essential, the above hypothesis will be rejected statistically.[1] However, given the complexity in obtaining the conditional and marginal distributions for stock returns, Strong (1992) re-denotes the hypothesis as

$$E[r_{it}|y_i] - E[r_{it}] = E[\epsilon_{it}|y_i] = 0. \tag{1.1.2}$$

The intuition for Equation (1.1.2) is that, if the signal y_i stands for the event of interest and if the signal is observable then, under the null hypothesis that the event is not essential, the signal y_i does not contribute any improvement of forecasts on rates of return r_{it}, where $\{\epsilon_{it}\}_{i=1,2,\cdots n}$ is considered the abnormal return.

Notice that the presumption of the logic (in Strong's format), however, lies on the availability of y_i and its proper representation (such as event windows, firm-specific attributes) on the corporate finance events of interest. By assuming that $E[r_{it}|y_i] = g_i(y_i)$, it is easy to see that the first half of Equation (1.1.2) implies $g_i(y_i) - E[g_i(y_i)] = 0$ since $E[r_{it}] = E[E[r_{it}|y_i]] = E[g_i(y_i)]$ by the law of iterated expectations. Yet, this part of Equation (1.1.2) does not necessarily imply that $E[r_{it}|y_i] = 0$. In other words, the signal for the event(s) of interest y_i can still have predictability for rate of return r_{it} even though the (signal of) event is not essential at all. Therefore, showing predictability of an event-related signal on excess returns may not necessarily indicate the essentiality of the corporate finance event(s) according to the first half of Equation (1.1.2) unless the $\{\epsilon_{it}\}_{i=1,2,\cdots,n}$ is explicitly related with $\{r_{it}\}_{i=1,2,\cdots,n}$.

It is under the assumption of linear unconditional asset pricing model that

$$\begin{aligned} r_{it} &= E[r_{it}] + \epsilon_{it} \\ &= \sum_{j=1}^{k} \beta_{ij}\varphi_{jt} + \epsilon_{it}, \end{aligned} \tag{1.1.3}$$

and if both the idiosyncratic risk $\{\epsilon_{it}\}_{t=1,2,\cdots}$ and the signal of event y_i are orthogonal to the systematic (and possibly unobservable) components $\{\varphi_{jt}\}_{j=1,2,\cdots,k}$, then under the null hypothesis of no essential event,

$$E[r_{it}|y_i] = \sum_{j=1}^{k} \beta_{ij} E[\varphi_{jt}|y_i] + E[\epsilon_{it}|y_i] = 0. \tag{1.1.4}$$

An even more difficult reality is, however, these signals $\{y_i\}_{i=1,2,\cdots,n}$ may either be unobservable or not entirely representative for the event(s). However, if equation (1.1.4) holds then, it would also imply that $E[r_{it}] = 0$ for all securities by the law of iterated expectations, for all $i = 1, 2, \cdots, n$. Yet, if some of $\{E[\varphi_{jt}|y_i]\}_{j=1,2,\cdots,k}$ are nonzero, then it would imply that the firm-specific signal becomes systematic component also.

On the other hand, the condition $E[\epsilon_{it}|y_i] = 0$ may have two implicit meanings as well. It implies either the signal $\{y_i\}_{i=1,2,\cdots,n}$ for the event(s) of interest are not informative (or correct), or the specific event has no impact at all. Notice that the signals $\{y_i\}_{i=1,2,\cdots,n}$ may be some indicators declaring for the events or simply considering if the data fall in the even windows or not. Since the classifications for the event windows may be presumed subjectively, these signals could be uninformative. Hence, the implications from equation (1.1.2) may require further elaboration for event studies.

Alternatively, if the conditional asset pricing model is applied and conditioning information set is denoted as the filtration Π_t at time t and the above null hypothesis can be re-denoted as $E[\epsilon_{it}|\Pi_t] = 0$. The conventional event studies emphasize the specifications as equation (1.1.5) where the return of a selected security i for a particular event interest is denoted as r_{it}, and this rate of return is decomposed into two components, namely

$$r_{it} = E[r_{it}|\Pi_t] + \epsilon_{it}, \tag{1.1.5}$$

where $E[r_{it}|\Pi_t]$ represents the normal (expected) return (or systematic component) based on available information Π_t (which includes the event date if known), $\{\epsilon_{it}\}$ is the abnormal returns for security i. Empirical event studies most likely focus on the statistics formed by the abnormal returns $\{\epsilon_{it}\}_{i=1,2,\ldots}$ around the preselected event windows across values of null hypothesis H_0. Hence, the correctness of specification in $E[r_{it}|\Pi_t]$ is crucial for the test statistics formed by $\{\epsilon_{it}\}_{i=1,2,\ldots}$. The short-run stock returns (such as daily returns), although are of higher frequency and more concurrent information, are still subject to this requirement.

Based on this setting, the conventional tests in event studies may actually be joint tests for hypotheses that the mean of abnormal returns is identical to zero over time and whether the assumed model for expected rate of return $E[r_{it}|\Pi_t]$ is correctly specified or not. Two issues hence are essential here in developing the tests on event-study issues. One is the correctness of model specification for $E[r_{it}|\Pi_t]$ in terms of functional forms, parametrization, and/or distributional properties. The other one is on the inclusion of essential explanatory variables for the specification of pricing kernel or core of empirical asset pricing models. Specifically, the second issue is on what variables in Π_t should be considered as essential.

Neither one of these issues is easy for corporate finance event studies.

Therefore, proper model specification of the normal (or expected) returns determines the plausibility of statistical inferences using abnormal returns.[2] In other words, regardless of long- or short-horizon event studies, model specification errors and inclusion of irrelevant explanatory variables in the expected rate of return will directly influence the validity of event studies using abnormal returns. Various event studies using short-run stock returns are provided through many financial literature and publications. Morse (1984) discussed the choice of monthly and daily stock returns for the event studies. He suggested that the application of daily returns is superior to the monthly returns when there is no uncertainty over the precise announcement date for the event information. This seems to support that applications of shorter measurement intervals is more robust for the event studies. As stated in Fama (1998), the short-horizon studies may be subject to less model specification errors since the event windows are short and the daily expected returns are closer to zero. Nevertheless, conventional model verification on the normal (or expected) returns when selecting the relevant variables for empirical asset pricing models is still based on statistical relevancy (or predictability) alone. Yet, it is not surprising to see that neither the predictability nor statistical relevancy may sustain or remain stable over different time periods.

Specifically, empirical evidence surrounding mergers and acquisitions using short-run stock returns are somewhat inconclusive, for instance. While many studies find that the bidder's returns are indistinguishable from zero (Dodd and Ruback 1977, Asquith et al. 1983, Dennis and McConnell 1986), many find that the shareholders of the bidding firm lose significant value upon the announcement (Dodd 1980, Draper and Paudyal 1999). The preselection preferences on the sample periods and time spans for event windows may cause the results of hypothesis tests less reliable as planned. Fama's (1998) criticism on the long-horizon event studies such as "… Splashy results get more attention, and this creates an incentive to find them" may apply to the short-run event studies as well if the (time series) properties of selected data are not carefully studied. In the worst scenario, the sampling or event horizon (and windows) are preselected so that the statistical results may be in favor of the preferred hypotheses.

As the overreaction and underreaction hypotheses in the long-horizon event studies may prescribe, the observed short-horizon market reaction will not necessarily reflect the true nature of the change in value when there are noisy information or rumors surrounding the events. This seems plausible if there is market inefficiency and the capital market cannot suddenly become fully informed. One plausible argument is that the acquiring management has yet full access on private information of the target firm on the announcement date. However, as declared by Fama (1998), the findings of long-horizon event studies may not show market inefficiency if the overreaction and underreaction reversals are equally likely for pre-event and post-event periods. In addition, these long-horizon results may disappear when different methodologies applied. Hence, implied by the comments of Fama (1998), choices on the horizons in event studies and different model specifications (and their errors) may be possible causes for various and controversial empirical results in mergers and acquisitions, for instance. More discussions on the event window selection will be covered in Chapter 3.

Providing various techniques dealing with the abovementioned issues, many short-run studies in mergers and acquisitions such as Dodd and Ruback (1977), Bradley (1980), Jarrell and Poulsen (1989) point out that target firms see a significant and positive market reaction to the news, while the bidding firm typically offers a premium to the target's shareholders. The evidence concerning the acquiring firm's market performance is less clear. Some simulation methodology of short-horizon event studies with daily stock returns was provided in Brown and Warner (1985). Similar studies are also presented in Dann (1981) and Masulis (1980). Whatever the objectives that the corporate finance event studies may focus on, the findings of Brown and Warner (1985) indicate that both serial dependence and changing variance of the data series prevail in these short-run abnormal stock returns. Modifications in statistical procedures for these time-series properties are necessary for more powerful conclusions. In addition, the cross-sectional dependence due to possibly industry clusters may also be influential to the power of the statistical tests. Although cross-sectional dependence can be reduced in (weighted) portfolio returns for the selected securities, it was criticized that the statistical significance of the event may be overstated if the historical variability of portfolio returns is applied for the test statistics. (Brown and Warner 1980, 1985, Collins and Dent 1984).

The main controversy of the short-horizon event studies is that they assume that the market quickly becomes aware of information and seemingly had heretofore adapted to the new information thoroughly. These shortcomings in short-horizon research lead to the investigation of the long-horizon performance of the merged firms. Several proofs are also provided using the long-run abnormal stock returns or buy-and-hold returns. For instance, it is shown that acquiring firms experience long-horizon underperformance against the market (Loughran and Vijh 1997) and against matched samples (Rau and Vermaelen 1998). Short of dismissing the control method used in this strand of research, one must realize that the research is perhaps more fruitful if one can be more ascertain on the model specifications, data construction, and statistical methodologies. Therefore, while attractive as an alternative, the method of using long-horizon assessment does not necessarily improve the tractability of the mergers/acquisition or market inefficiency. Many empirical studies with simulations provide some evidence for this perspective.

According to Kothari and Warner (2007), their simulations indicate that the long-horizon event studies are prone to higher sensitivity of model specifications on (conditional) expected returns. This is quite apparent that as the time span expands, more influential schemes may affect the normal (or expected) returns in addition to the events themselves. In other words, as the time horizons for event studies expand, less controllable model specification on (systematic) expectation may result. This, in turn, may cause the abnormal returns to contain the significant model specification errors so that the underlying statistics for event studies become unreliable. This shows another issue that the difficulty of determining event window is linked with the data to be collected.

As stated in Kothari and Warner (2007), the short-horizon study can be powerful only if the abnormal performance is concentrated in the event window. In other words, if the impact of the events are short and immediate where no spill-overs are shown after the event windows selected, the short-horizon event studies seem more conclusive. However, the conventional or robust estimates on the return variability (such as the standard deviation of cross-sectional average abnormal returns) are either based on pre-event historical time series or those of post-event periods. This possibly makes the test statistics to reject the null hypothesis more likely. On the other hand, the possible increase in return variability during the presumed event window was considered as a significant event even though the increase may also be considered as normal since the uncertainty increases due to noisy information or rumors. Furthermore, as stated in Kothari and Warner (2007), the long-horizon event studies may suffer from (1) poor specification on the systematic components of stock returns and (2) sensitivity of test statistics versus the assumption on return generating process. Hence, the long-run event studies (either on mergers and acquisitions or else) will have lower power to detect abnormal performance.

As a result, in verifying the importance of mergers and acquisitions for the corporate finance perspectives, several decision rules must be made. First of all, the decisions must be made on the frequency of sampled data such as stock returns. Should the short-run or the long-run stock returns be sampled in the studies? Is the interest of study on the short-horizon or long-horizon impact? Many empirical studies have been provided when both the short-run and long-run returns are applied. Unfortunately, no unified decision rules are devised to investigate the corporate event issues. Although the event studies in using long-run returns are more consistent in different settings than the short-term returns, the difficulty lies in the controversies of necessary information or explanatory variables for the specifications of fitted regressions or mechanisms in forming the (conditional) expectations of stock returns over the event periods.

It is not difficult to envision that, over the long-run horizon (across the pre-event onto the after-event periods), the presumed attributes in forming the systematic components for the stock returns may not be sufficient due to various contingencies of the capital markets or business cycles of economy. However, it is critical for the correctness of the specifications of systematic components in stock returns to enable robust conclusions of event studies in using the so-called abnormal returns. The arguments on using long-run stock returns may assert that if certain corporate information such as mergers and acquisitions have initiated essential impacts, the effect could be persistent and hence, resulted more vividly in the long-run abnormal returns instead of the short-run noisy abnormal returns. Even so, two additional questions should be considered in performing event studies in corporate finance: (1) what is the availability and frequency for the (economic) attributes that may be applied to approximate the (systematic) expected rate of return with acceptable accuracy? (2) what are the robust statistical procedures that can be devised when considering the possible (built-in) properties in the collected data?

Although applications of short-run stock returns enjoy more frequent observations in capital markets and hence receive more concurrent information that may be considered more informative, such noisy information may contain less tractability since the short-run approximation of the expected rate of return such as asset pricing models usually suffers from lower explanatory power. As such, it seems that the choices of short-run or long-run stock returns also lie crucially on the availability (and frequency) of economic/corporate data or any sensible attributes in forming the (conditional) expectations to specify the systematic components of stock returns.

In order to obtain the critical approximation for (conditional) expectations that represent the systematic component in stock returns, some filtering mechanism must be devised for either short- or long-horizon event studies. Yet, for the short-horizon studies, the filters must be robust to filter out the critical components since the data are quite noisy. In addition, the filters should be adaptive enough (in data frequency, dynamic structures, or else) so that the concurrent market information is taken into account efficiently. However, for the long-horizon studies, robust filters may also be more difficult to fetch since (a) the underlying systematic components are not necessarily stable and the corporate-event related information could be endogenous to the system that causes the systematic components and structures to alternate over time, (b) the included list of (economic) attributes may not be enough to capture the underlying system across expanded time horizons. As in Fama et al. (1969) where the separation of expected and unexpected components in stock returns is critical for the soundness of testing empirical hypothesis in corporate finance, any prudent attempt to search for the critical approximation on the expected components (through some sensible asset-pricing models or else) is essential in corporate finance event studies. In other words, discussions of corporate event studies must start with some thorough model specification search (or filtering) for the expectations in stock returns for either short- or long-horizon studies.

Nonetheless, since approximating expected components of stock returns is subject to limited information in most cases, even the best model (in statistical criteria or else) may still contain approximation errors. Some robust devices or statistical methods that may undertake certain acceptable approximation errors and be capable of providing reliable hypothesis testings are needed for the event studies in corporate finance. Hence, a well-established event study in corporate finance based on stock returns should at least contain two particular ingredients: (1) a thorough model search and filtering for the (conditional) expected rates of returns relating to the included (economic) attributes and (2) a robust statistical procedure that allows acceptable approximation errors and reduces the subjectivity in sampling, choices of event windows, or time horizons of interest. This is discussed more extensively in Part II of the book.

Barber and Lyon (1997) and Kothari and Warner (1997) both discovered that the long-run stock returns may lead to misspecified test statistics due to biases such as new listing, re-balancing, and skewness of the distribution of long-term abnormal stock returns. Lyon et al. (1999) provided some improved methods to correct those biases. However, as stated in their research, "...The most serious problem with inference in studies of long-run abnormal stock returns is the reliance on a model of asset pricing. All tests of the null hypothesis that long-run abnormal stock returns are zero are implicitly a joint test of (i) long-run abnormal returns are zero and (ii) the asset pricing model used to estimate abnormal returns is valid." Specifically, the validity of statistical arguments on long-term abnormal returns can be well-established only when the underlying systematic components are thoroughly clarified.

In Barber and Lyon (1997), for instance, the calculation of (post-event) buy-and-hold long-run abnormal returns is formed as $\prod_{t=1}^{T}[1+r_{it}]-\prod_{t=1}^{T}[1+E[r_{it}]]$, where the first term represents the buy-and-hold return on the security i, and the second term represents expected buy-and-hold return on some bottom-line benchmarks or attributes, where $\{r_{it}\}_{t=1,2,...,}$ are the short-run (say, monthly) returns for security i, and $E[r_{it}]$ represents the expectation based on the benchmarks or attributes. Hence, without the correct specification of $E[r_{it}]$, the constructed buy-and-hold abnormal returns will possibly contain the specification errors from the models. Another important issue is that serial dependence may appear when these buy-and-hold abnormal returns are constructed as the presumed time span (say, a year or more) T expands, even though the expected rates of returns $E[r_{it}]$ are correctly formed.

Showing the deficiency in using the long-run abnormal returns for corporate finance event studies is straightforward. The following example assumes continuous compounding for simplicity. The same conclusion will also hold even if discrete compounding such as $\prod_{t=1}^{T}[1+r_{it}]$ in the works of Jakobsen and Voetmann (2005), Barber and Lyon (1997), and Lyon et al. (1999) is applied.[3]

Let the buy-and-hold return for a selected security from the announcement day with the pre-determined time span T be determined as follows. For an initial investment W_{it} for security i, and with the stochastic (monthly) returns r_{it} for any starting date t, the terminal wealth of this investment is equal to

$$W_{i,T}^{t}=W_{i,t}e^{\sum_{t}^{t+T}r_{it}},\tag{1.1.6}$$

Hence, the buy-and-hold return for this security is equal to $\sum_{t}^{t+T}r_{it}$.

Let the benchmark attributes (or so-called characteristic-based matching firm's returns) that can be applied to describe the systematic components of all security returns be denoted as $\{r_{mt}\}_{t=1,...,T}$ for simplicity. [4]

Assuming the similar accumulated wealth relative as in Ritter (1991) and Loughran and Ritter (1995), or as the market-adjusted returns denoted as long-run abnormal returns in

Kothari and Warner (1997), we may find that if $W^t_{m,T} = W_{m,t}e^{\sum_t^{t+T} r_{mt}}$,

$$\left(W^t_{i,T}/W^t_{m,T}\right) = (W_{i,t}/W_{m,t})e^{\sum_t^{t+T}(r_{it}-r_{mt})}. \quad (1.1.7)$$

In other words, the buy-and-hold abnormal returns with the presumed long-run time span T, can be denoted as

$$ln\left(W^t_{i,T}/W^t_{m,T}\right) = ln(W_{i,t}/W_{m,t}) + \sum_t^{t+T}(r_{it} - r_{mt}). \quad (1.1.8)$$

Let $\varepsilon_{it} = r_{it} - r_{mt}$ be denoted as the interim (monthly) abnormal returns for security i, and denoting the long-run wealth-relative (abnormal) return (or so-called buy-and-hold return) for security i as $ar^i_{t,T} = ln\left(W^t_{i,T}/W^t_{m,T}\right)$ such that that $ar^i_{t,T} = ln(W_{i,t}/W_{m,t}) + \sum_t^{t+T}\varepsilon_{it}$. Let $\mu_i = ln(W_{i,t}/W_{m,t})$ for all $t's$ for simplicity, it is easy to see that $\left\{ar^i_{t,T}\right\}_{t=1,2,...}$ will be subject to serial dependence across $t's$, since $E[(ar^i_{t,T} - \mu)(ar^i_{t+1,T} - \mu)] \neq 0$, in general.

More specifically, this serial dependence will possibly increase with more lags when the presumed time span T becomes larger. As a result, such built-in serial dependence in time series of long-run abnormal returns can not be easily eliminated even with the cross-sectionally weighted portfolio returns unless the built-in serial dependence vanishes as the number of included securities becomes relatively large. This is easy to verify if we obtain the weighted average of the individual firm's abnormal returns or buy-and-hold returns,

$$ar^p_{t,T} = \frac{1}{n_t}\sum_{i=1}^{n_t} ar^i_{t,T} = \frac{1}{n_t}\sum_{i=1}^{n_t}\left[\sum_t^{t+T}\varepsilon_{it}\right]. \quad (1.1.9)$$

We may express the long-run abnormal return for security i in terms of an autoregressive time series model such that

$$A(L)ar^p_{t,T} = \mu_p + \nu_{pt}, \quad (1.1.10)$$

where $A(L)$ is a polynomial of lag operator L, ν_{pt} is a zero-mean random noise. For an over-simplified setting, we may simply assume that there is only one lag in the $A(L)$, so that $A(L) = 1 - \rho L$, where $\rho > 0$, for instance. Hence, the (long-run) mean of the $\left\{ar_{t,T}^{p}\right\}_{t=1,2,\ldots,T}$ will be equal to $\mu_p/(1-\rho_p)$. Since the serial dependence for $\left\{ar_{t,T}^{p}\right\}_{t=1,2,\ldots,}$ will become stronger, it can be assumes that $|\rho|$ will increase as T expands. This shows that as $\mu_p \neq 0$, the absolute value of the mean of long-run abnormal returns will become larger when the serial dependence in $\left\{ar_{t,T}^{p}\right\}_{t=1,2,\ldots,T}$ increases. As a result, if the presumed long-run time span T increases where the lag coefficient $|\rho|$ in the autoregression $A_p(L)$ increases, it is highly likely that the mean of long-run abnormal returns for weighted portfolio will become relatively large. This phenomenon is confirmed in the simulation study of Kothari and Warner (1997).

This serial dependence is also shown in the empirical study by Ritter (1991) for the long-run performance of initial public offerings. In addition, Debont and Thaler (1985, 1987) also presented that there is a negative relation between the past and subsequent abnormal returns on individual securities using holding periods of a year or more. Although their interpretation considers this relation as the evidence of market overreaction, this serial dependence can possibly be a result of data construction based on the above reasoning. Hence, in using the long-run abnormal returns for corporate finance event studies, one needs to be cautious on the possible consequence for the built-in serial dependence from data construction. Similar arguments can be found in Jegadesh and Karceski (2009) that they also discover the long-horizon returns may exhibit positive autocorrelation due to overlapping return data.

In other words, the attempt in using long-run abnormal returns to avoid possibly autocorrelated and noisy short-horizon abnormal returns will actually introduce serial dependence due to the constructed data series. Hence, applications in using the long-run cumulative abnormal returns over event periods, the test statistics ought to take this property into account. This built-in serial dependence, for example, represents an issue of statistical properties of abnormal returns that must be taken into account in forming test statistics for empirical studies (Khotari and Warner 2007). More specifically, Jegadeesh and Karceski (2009) modify the conventional test on long-run abnormal returns by using robust estimates for variance regarding the aforementioned serial correlation and possible heteroskedasticity.

To consider the possible constructed serial dependence, an example with Exxon-Mobil returns is included for the above claim. The data set includes the Exxon-Mobil monthly (dividend-adjusted) stock returns from January 1988 to December 2013 accordingly. The time period covers the period prior to merger of these two firms in 1989, and the after-merge period. Following the construction in Equation (1.1.8) (and use the market index returns for normal or benchmark returns), the cumulative abnormal returns are formed. The holding-time horizon T is set to be 3, 6, 12, and 18 months, respectively. In addition, for the company Pfizer, the monthly returns (dividend-adjusted) are collected from June 1989 unti December 2013 for the sme holding-time horizon T. A Table 1.1 is included in the following where the first-lag coefficients of autoregressive time series model and Ljung-Box Q-statistics are both shown.

Not surprisingly, the data show some significant serial dependence across different time lags assumed. In other words, if constructed as the sample allows, the long-run serial dependence will result. Therefore, claiming the long-run serial dependence may not confirm or refute the hypotheses of interest. It may simply result from the data constructed. Or, more technically instead, various choices over the benchmark returns in equation (8) may cause the serial dependence, too. Notice that the claim is not to say, there is no such serial dependence in the long-run abnoraml returns at all. However, caution must be applied to the data stream first, prior to further elaborations or modellings. Table 1.1 provides the Ljung-Box Q test statistics for the null hypothesis of white noise.

This shows that the constructed series are subject to significant serial dependence. In particular, when the time span expands, the serial dependence turns stronger. In other words, whether the time horizon is long or short for the event studies, additional cautions must be applied to the possible time series properties of constructed data. Hence, one should be cautious

Table 1.1 Ljung-Box Q-statistics and AR(1) coefficients for Exxon Mobil and Pfizer Abnormal Returns

		Exxon-Mobil		
Lags	3 months	6 months	12 months	18 months
1	109.858*	185.032*	233.290*	248.555*
3	131.082*	329.788*	540.199*	625.898*
5	134.027*	346.582*	711.598*	873.989*
AR(1)	0.609*	0.794*	0.898*	0.920*
		Pfizer		
Lags	3 months	6 months	12 months	18 months
1	125.194*	192.796*	235.120*	250.423*
3	149.558*	384.452*	585.988*	663.766*
5	155.968*	428.392*	826.447*	986.029*
AR(1)	0.649*	0.810*	0.903*	0.944*

All the asterisk signs "*" in the table indicate that the statistics are significant at 1 percent level.

on whether the serial dependence is actually from the data set constructed or is from the population of the sampled data.

As in Kothari and Warner (1997), and Lyon et al. (1999), both studies applied (block) resampling bootstrapping methods for the test statistics to reduce the biases of these long-run abnormal returns. However, according to Corollary 2.1 of Lahiri (2003), such serial dependence (for individual security) can make the independent bootstrap method fail so that the bootstrap estimator for the distribution of sample statistics (such as the mean of individual security) is not consistent. Hence, a more general resampling schemes (such as block bootstrap method) on the long-run abnormal returns should be applied for these test statistics instead.

Recent study of Bremer et al. (2011) applies a similar block bootstrapping simulation for the long-term event studies. They discover that the application of non-overlapping quarterly abnormal returns (which are obtained based on the characteristic-based matching firm's returns) may provide better power for the tests in long-run event studies. In particular, they also discover that in generating abnormal returns of buy-and-hold returns, the application of equally weighted portfolio of characteristic-based matching firm's returns as the benchmark returns is more superior than the individual characteristic-based matching firm. In fact, the result precisely indicates that the choices of model specifications in forming the expected components of the stock returns are essential for the robustness of event studies in empirical corporate finance. A better model specification for the expected components of the stock returns (conditional on the aforementioned information set Π_t), regardless of the frequency of data, will provide a more robust statistical result in hypothesis testing.

For using high frequency news and data in event studies, Bohn et al. (2013) provide an alternative method for corporate finance issues. Instead of following conventional approach that starts with identification for corporate event dates and news, the alternative approach starts from verifications on the "jumps" of stock returns. Namely, the approach applying the statistical methods such as (bi)power variation to consider the jumps of stock returns without reference to some particular news or announcements. After identifying the jumps, the method is then to associate the news data with the already-identified jumps of stock returns. The advantage is that there is no need to identify the event time in advance. In other words, the approach avoids the subjective determination on the event period, event dates through the researcher's a priori judgment, especially on the unscheduled news or announcements. In particular, high-frequency data such as intraday observations are collected for the studies.

Unfortunately, even though the approach is promising, the high-frequency data may contain some microstructural noises, which may not be associated with any event (or news) of interest. Another difficulty is that, although identification of jumps of stock returns is carried out, the association between the news and the jumps seems arbitrary. In other words, even though some jumps may be related with the news by associating two time series, there is no particular robust statistical method applied to verify these results.

Another issue is that even though the association is legitimate, there could be so many different news or events that were identified and some of them may be mutually related also. Thus, it is difficult to see (1) which news is the main cause for the events or jumps since some jumps may be caused by non-corporate-finance issues, (2) what theoretical/economical reason(s) may be identified to explain the results. In addition, not all jumps can be explained by the feasible news data sources or news time series. A question that should be clearly considered, however, is whether identification of jumps alone fulfills the purpose of event studies or not. Regardless of the frequency of data applied, a meaningful event study is to consider certain social-economic/corporate finance issues where stock returns data are applied to verify the hypotheses. In turns, these verified hypotheses help to understand the functioning of capital market equilibrium and the impacts of events on the corporates or related institutions. While essential verification is provided by robust methods, identifying jumps or else (such as cumulative abnormal returns) is only applied as a tool for such purpose.

1.2 Variable Selection in Assessing the Probability of an Event?

For a corporate event, both the impact of the event and the informativeness of pre-event signals are essential in corporate finance. Likewise, it is equally important to determine the probability (of success) of the events. Many research works have discussed the related issues, for instance, on the mergers and acquisitions. Two areas of research can be shown as: (1) the prediction of likelihood of being the target firm and (2) the assessment of probability of successful takeovers. Either approach requires assessments for the implicit information (from stock returns or else) to infer the hypotheses of interest. The implication is that for the issues such as mergers and acquisitions, the management of both target and acquiring firms may possess more informative sources in comparison with the investors in public. The process of negotiation for the deals and others may not be fully disclosed. The success or failure of the events such as mergers and acquisitions requires inferences of the hidden nondisclosed information to improve the assessments. In particular, the emphasis is on how probable the events may occur. However, the issues are much more complicated than just model specification on how the market may infer the underlying events to happen or not. A more essential issue is that variable selection for the models of interest remains challenging for the empirical finance.

For instance, Brown and Raymond (1986) investigate the so-called risk arbitrage and the probability of successful corporate takeovers. Their arguments claim that although risk arbitrage may cause investor to search for the best price spread between the offered price and the market price for the target firms, the competition in the market will eventually show that the success of the deal can be inferred from the prices in the post-announcement period. For example, let P_{mt} stand for the prevailing market price for a target firm when there is a tender offer at the price P_{Tl}. The purchase of the stock will have a return as $\{P_{Tl} - P_{mt}\}/P_{mt}$ if the merger is completed successfully. However, there is a chance of unsuccessful merger where the target's stock price may fall to the "fallback" price P_F, where the return may fall to $\{P_F - P_{mt}\}/P_{mt}$. Assuming that the expected risk arbitrage payoff is zero, it can be shown that at time t,

$$E[\pi_t] = x_t \left[\frac{P_{Tl} - P_{mt}}{P_{mt}}\right] + (1 - x_t)\left[\frac{P_F - P_{mt}}{P_{mt}}\right] = 0, \quad (1.2.1)$$

where x_t stands for the period t merger probability, π_t represents the risk arbitrage payoff. Thus,

$$x_t = \frac{1 - \left(\frac{P_F}{P_{mt}}\right)}{\left(\frac{P_{Tl}}{P_{mt}}\right) - \left(\frac{P_F}{P_{mt}}\right)}. \quad (1.2.2)$$

Equivalently, the probability can be rewritten as

$$x_t = (P_{mt} - P_F) \div (P_{Tl} - P_F). \quad (1.2.3)$$

In order to ensure Equation (1.2.3) represents a proper probability measure, Brown and Raymond (1986) introduce

$$x_t = Min\{[Max\{(P_{mt} - P_F), 0\} \div (P_{Tl} - P_F)], 1\}. \quad (1.2.4)$$

To measure the "fallback" price, Brown and Raymond (1986) apply the target firm's stock price four weeks prior to the announcement date to approximate it. However, there are two difficulties for this method of obtaining the probability of event's success of failure. There is no theoretical reasoning that this measurement will represent the "failure" price closely. In addition, the framework does not consider the occasions that some private firms could be the targets whose stock prices are unavailable in public. Hutson (2000) extends their analysis to Australian market. In particular, the "fallback" price is replaced with the initial price P_I, which is the initial price of the target firm one month before the announcement of the bid. However, there is no theoretical justification that such a sampling scheme is well-suited for the analyses.

Alternatively, Acharya (1993) introduces the latent variable model to infer the hidden information (and the probability) for corporate events. The analysis also starts with the model such as

$$r_{it} = \beta_i' W_{it} + \epsilon_{it}, \tag{1.2.5}$$

where W_{it} includes a vector of firm-specific attributes as of time $t-1$ and market-wide factors as of time t, β_i is a vector of associated coefficients. [5] Notice that the condition $E[\epsilon_{it}|W_{it}] = 0$ holds in the period in which the event is improbable.[6]Accordingly, the conventional event studies obtain the event period residuals as

$$\hat{\epsilon}_{it} = r_{it} - \hat{\beta}_i' W_{it}, \tag{1.2.6}$$

provided that time t lies in event period. In short, for any time period that contains probable events (whether announced or not), the model of Equation (1.2.5) can be rewritten in to an alternative form such as

$$r_{it} = \beta_i' W_{it} + \delta_{ei} I_{eit} + \delta_{ni} I_{nit} + \eta_{it}, \tag{1.2.7}$$

where time t can be either in event or non-event periods, I_{eit} is a dummy variable, which is equal to one when t is in the event period, and is equal to zero otherwise, and $I_{nit} = 1 - I_{eit}$. $\{\delta_{ei}\}_{i=1,2,\cdots,n}$ and $\{\delta_{ni}\}_{i=1,2,\cdots,n}$ are expected returns conditional on the event or the nonevent. For further analysis, Acharya (1993) introduces a latent variable model such that

$$I_{eit} = \begin{cases} 1, & \theta' z_{it-1} + \xi_{it} > 0 \\ 0, & \theta' z_{it-1} + \xi_{it} \leq 0, \end{cases} \tag{1.2.8}$$

where $x_{it} = \theta' z_{it-1} + \xi_{it}$, and x_{it} stands for normalized increment of net present value of a firm when announcing an event versus not announcing an event at time t, ξ_{it} is the standardized latent information at time t for firm i, z_{it-1} is a vector of firm-specific attributes included in the vector W_{it}. In addition, a cross-sectional relationship is assumed that

$$\epsilon_{it} = q\xi_{it} + \eta_{it}. \tag{1.2.9}$$

After substitution, it is feasible to express the conditional expectation

$$\begin{aligned} E[\epsilon_{it}|I_{eit}=1, W_{it}] &= qE[\xi_{it}|z_{it-1}, \xi_{it} > -\theta' z_{it-1}] \\ &= q\frac{\phi\left(\theta' z_{it-1}\right)}{\Phi(\theta' z_{it-1})}, \end{aligned} \tag{1.2.10}$$

and accordingly,

$$E[\epsilon_{it}|I_{nit}=1, W_{it}] = -q\frac{\phi\left(\theta' z_{it-1}\right)}{1-\Phi(\theta' z_{it-1})}, \tag{1.2.11}$$

if normality is assumed, where $\phi(.)$ and $\Phi(.)$ are the standard normal density and cumulative distribution function, respectively. Using the notions of conditional expectations, Equation (1.2.7) can be re-written into

$$r_{it} = \beta_i' W_{it} + q\left[\frac{\phi\left(\theta' z_{it-1}\right)}{\Phi(\theta' z_{it-1})} I_{eit}\right] - q\left[\frac{\phi\left(\theta' z_{it-1}\right)}{1-\Phi(\theta' z_{it-1})} I_{nit}\right] + \nu_{it}. \tag{1.2.12}$$

The difficulty of this model is that, similar to the conditional expectations of the abnormal returns (denoted as $\{\epsilon_{it}\}_{i=1,2,\cdots,n}$), the conditional expectations in the "beta's" incurs firm-specific attributes as well. In other words, the abnormal returns are then also related with the normal (expected) returns where investors can infer both from the firm-specific attributes in the past and concurrent information sets. In addition, the model (in Equation (1.2.12) for instance) also assumes that these past firm-specific attributes will contribute to the specification of expected abnormal returns when events happen. For instance, Equation (1.2.12) for the time period where events are probable is actually expressed as

$$r_{it} = E[r_{it}|W_{it}] + \nu_{it}, \tag{1.2.13}$$

where here W_{it} includes a vector of firm-specific attributes as of time $t-1$ and market-wide factors as of time t. In this case, is the specification of $E[r_{it}|W_{it}]$ in Equations (1.2.12) and (1.2.13) the correct model for normal (expected) returns or not? If the answer is yes, that would imply these firm-specific attributes also contribute to the formation of market expectations on both the market-wise fluctuations and corporate finance events, even though the corporate finance events are considered idiosyncratic or diversifiable risks. For instance, it seems straightforward to rewrite Equation (1.2.12) when considering $W^{\star}_{it}$ of marketwise factors as a subset of W_{it} that excludes the sub-vector $\left\{z_{it-1}\right\}_{i=1,2,\cdots,n}$ where $W_{it} = \begin{bmatrix} W^{\star}_{it} \\ z_{it-1} \end{bmatrix}$, and

$$\begin{aligned} r_{it} &= E[r_{it}|W^{\star}_{it}] + g(z_{it-1}) + \nu_{it} \\ &= E[r_{it}|W^{\star}_{it}] + \epsilon^{\star}_{it} \\ &= \widetilde{\beta}'_i W^{\star}_{it} + \epsilon^{\star}_{it}, \end{aligned} \tag{1.2.14}$$

by setting $g(z_{it-1}) = \overline{\beta}'_i z_{it-1} + q\left[\frac{\phi(\theta' z_{it-1})}{\Phi(\theta' z_{it-1})} I_{eit}\right] - q\left[\frac{\phi(\theta' z_{it-1})}{1-\Phi(\theta' z_{it-1})} I_{nit}\right]$, $\overline{\beta}_i$ and $\widetilde{\beta}_i$ are sub-vectors of β_i such that $\beta_i = \begin{bmatrix} \widetilde{\beta}_i \\ \overline{\beta}_i \end{bmatrix}$.

Notice that in Equation (1.2.13), it is pre-supposed that the essential variables for model specification have been identified. In empirical finance, this is far more than a question of model/variable selections or statistical tests of significance. Additional justification must be provided to consider the distinction between normal (expected) returns and abnormal returns, accordingly. Furthermore, in Equation (1.2.14), it shows that if these firm-specific attributes actually contribute to the identification of probable corporate events of interest, it seems more ideal to simply construct the normal (expected) returns using the market-wise information set $W^{\star}_{it}$ and leave the rest to the abnormal returns $\epsilon^{\star}_{it}$ *per se*. The abnormal returns when properly constructed, will be informative enough to tell the stories. Given so, the variable selection in the model of Equation (1.2.7) is critical. In empirical finance, the choices of variables are not immediately straightforward using model specification tests or selection criteria such as AIC, BIC or else. Instead, specific economic properties of the normal and abnormal returns should be considered first.

Although filtering out all possible non-related noises to isolate the signal of event(s) explicitly seems straightforward, it remains dubious for the presumption that inclusion of additional firm-specific attributes (in the expression of normal returns) may improve the tests for event studies. A critical issue then, is whether it is optimal to include all perceived (or possibly relevant) variables in the model(s) of normal (expected) returns to analyze the corporate finance events or not. Furthermore, as indicated in Branch and Yang (2003), some firm-specific attributes are useful in identifying the likelihood of successful takeovers for different payment methods/merger types. Since mergers and acquisitions are considered as firm-specific events, it seems more ideal to confine these variables to the model specification for abnormal returns. This, in turn, leads to the discussions in Chapter 2 where model search for the normal (or expected) returns is essential prior to the studies on abnormal returns in all cases whether the time periods are of probable events or not. The event studies in empirical corporate finance may become more reliable if the dichotomy of normal (or expected) returns and abnormal returns can be better assumed when variable selection is processed rigorously.

Chapter 2

Model Specifications for Normal (or Expected) Returns

Introduction

For corporate finance event studies that look into abnormal returns, robust model specifications for normal (expected) returns are needed. However, to verify the model specification, one needs to be cautious about the included explanatory variables. Although many candidate variables seem useful in forecasting the returns, they are not necessarily genuine systematic variables that explain the capital market equilibrium. Common-sense reasoning may be considered for filtering the returns thoroughly with all seemingly significant variables to provide cleaner abnormal returns. Yet, inclusion of nonsystematic firm-specific variables in the expected rates of returns may, in fact, result in incorrect conclusion due to possible over-rejection in statistics applied. This chapter introduces some new arguments for specification of normal returns.

In particular, it is shown that inclusion of redundant variables for the specification of normal (or expected) returns may become detrimental for the event studies especially when the included (firm-specific) variables are associated with the underlying events. The nondiversifiability of any included variable is emphasized so that strong cross-sectional dependence will result if the variable is indeed a systematic one. New test statistics are introduced to verify this necessary condition for nondiversifiability of included variables. The guidance for model specification is simple. That is, (1) to verify the nondiversifiability of included variables with the cross-sectional strong dependence, and (2) to avoid including the possibly diversifiable (and known) variables or attributes in the expected returns even though they may improve forecasts. Namely, the purpose is to make the abnormal returns less "scrambled" with unnecessary ad hoc information.

2.1 Model Search

As mentioned in Chapter 1 that model specification for the normal (or expected) components in stock returns is essential in the tests for event studies, this difficult task has been discussed in many event studies of corporate finance literature. Strong (1992), for instance, surveyed different methods of modeling abnormal returns using various models for normal (or expected) returns. One of the major issues in the model specifications for expected components in stock returns is the selection of (economic) attributes — either for short- or long-horizon event studies. The main question is "what determines the normal (expected) returns?" Or specifically, "is it legitimate to include all explanatory variables in the empirical asset pricing models as long as they have statistical/time series predictability?"

When short-horizon event studies are of interest, for instance, the selection of (economic) attributes and determinants is usually limited due to the availability and frequencies of influential data for expected components of stock returns. For the long-horizon event studies, the difficulty of data availability may become less severe, depending on the corporate issues under studies. Nevertheless, the longer the horizon of interest in the long-horizon studies, the more likely the (unknown) external factors or else may become influential. In other words, model specifications (based on the current and past information of presumed variables) may become insufficient. Even if all relevant variables are available, a device for variable selection must be provided to perform model search for modeling the (systematic) expected components. One issue then is that inclusion of all statistically relevant variables in forming the expected returns may not necessarily be ideal at all.

Notice that the dichotomy of stock returns into expected and unexpected components facilitates the analysis for both expectations subject to market (systematic) risks and the abnormal returns (which are considered idiosyncratic). While it is interesting enough to include all publicly available (economic or firm-specific) attributes in modeling the expected components, some of these variables are not necessarily of systematic risks borne by the entire capital market. For instance, some accounting measurements or corporate finance attributes such as corporate controls or anti-takeover actions although essential in corporate finance may not be identified as systematic attributes for the entire market unless these variables are hidden components for systematic risks. Not surprisingly, these abnormal returns are denoted as "abnormal" is simply because they are unexplained by the systematic components (or systematic risks) of stock returns. Hence, determination on what kind of financial/economic attributes are considered "systematic" is critical for the classification of normal and abnormal returns and, more importantly, on the validity of event studies in corporate finance.

Although ideally the data processing of corporate event studies may attempt to obtain the abnormal returns as "thorough" as possible when stock returns are filtered through all relevant variables, the search for the most representative model that provides the genuinely nondiversifiable (economic) attributes for normal (or expected) returns is necessary. A critical question then, is what the normal (or expected) returns should be. If the "normal" returns are considered as systematically expected for the benchmarks when the capital market equilibrium is achieved (or assumed), then concerns over the declared firm-specific attributes may not be essential for their specifications in normal returns. A well-known reason is, these firm-specific attributes are only for the selected firms and not for the systematic components of the entire capital market expectations. However, on the other hand, a dilemma also starts when one needs to exhaust feasible model search as thoroughly as possible so that no essential systematic information is dismissed.

For instance, due to the noisy data in the short run, some robust and efficient filters must be devised to screen out the systematic components of stock returns. More specifically, the filtering should thoroughly identify the normal (or expected) components of stock returns that may represent the benchmarks in capital market equilibrium (at least, ideally or theoretically). Although there might be concerns that these filterings may filter out the essential information including those impacts from the event(s), it is conceivable to consider that (for both short- or long-horizon event studies) the impact(s) of event(s) may become endogenous to the system since the market may digest the information or rumor(s) prior to or surrounding the event date(s). In other words, if such filtering in applications can filter out the essential event impacts from the data already, it is feasible to consider that the investors (individuals or institutions) in the entire capital market can trace out the same regularity with similar technicalities in forming their expectations or forecasts as well. Therefore, applications of robust and efficient filters (when thoroughly applied), in turn, will result with some abnormal returns honestly representing the unexpected components of stock returns.

In short, the model specifications for normal (or expected) returns need to consider two steps:(1) to identify the essential attributes (or factors) in expected returns via model search, and (2) to apply possible filtering devices (algorithms, or statistical methodologies) in filtering the expected returns from stock returns. For brevity, the applications of robust filters for the expected returns will be discussed in Part 2 of this book.

Many research papers had discussed the applications of different model specifications for expected returns. For instance, Thompson (1988) claimed that little difference is found in using different methods for the benchmarks of stock returns. In other words, the industry indices, return forms, have little impact on the event studies when compared to the conventional market model. However, it is also essential that the dichotomy of rates of return into normal (or expected) returns and the unexpected (abnormal) returns requires the risks incurred to be considered as systematic (nondiversifiable) or else.

In other words, if there is some firm-specific information that is relevant to the presumed event(s) to study, this information (or variable) if not systematic should not be included in the expected components of stock returns. Hence, despite that some so-called asset-pricing models may entail the statistically significant variables or factors, the verification on the nondiversifiable property of these included or presumed variables is essential for the construction of normal (expected) returns. On the other hand, if these information sets of presumed variables (in forming the normal (or expected) returns) include the statistically significant yet firm-specific variables that are diversifiable, the abnormal returns thus-wise may be polluted with additional noises.

For instance, let the true structural equation of the stock returns i =1,2,...,N be shown as follows

$$r_{it} = E[r_{it}|\Pi_t] + \epsilon_{it}, \tag{2.1.1}$$

where Π_t represents the information set of nondiversifiable (or systematic) economic attributes or factors (including their past history), ϵ_{it} is the genuine idiosyncratic risk that may contain all firm-specific variables or signals and especially, the impact(s) from the event(s) of interest. For simplicity, let $E[r_{it}|\Pi_t] = \beta_i'\widetilde{X}_t$, where $\widetilde{X}_t$ is a k-by-1 vector of economic attributes that are considered non-diversifiable, β_i is a k-by-1 vector of parameters. In addition, let

$$\epsilon_{it} = \theta' y_{it} + v_{it}, \tag{2.1.2}$$

where y_{it} is a p-by-1 (say, $p > 1$) vector of firm-specific diversifiable risks that represent the impacts of event(s), θ is a p-by-1 non-null vector of unknown coefficients, v_{it} is a pure random noise such that $E[v_{it}|y_{it}] = 0$, and that $\mathrm{Var}(\epsilon_{it}) = E[\epsilon_{it}^2] = \sigma_{it}^2 = \theta' E[y_{it}y_{it}']\theta + E[v_{it}^2]$ is the variance of the idiosyncratic risk.

Let $E[y_{it}] = 0$, $E[y_{it}y_{it}'] = \Sigma$, and Σ is a p-by-p positive-definite matrix under the null hypothesis of no essential impact(s) from the event(s). Now suppose the empirical study proposes to include some firm-specific diversifiable variables (which may be a subvector of y_{it}) and are denoted in a q-by-1 ($q < p$) subvector z_{it} where $E[\widetilde{X}_t|z_{it}] = 0$, for simplicity. Then, the extended model will become $E[r_{it}|\Pi_t, z_{it}] = \beta_i'\widetilde{X}_t + \theta' E[y_{it}|z_{it}]$, where $E[y_{it}|z_{it}] \neq \underline{0}$, and it is easy to see (under iterated law of expectations) that $E[r_{it}|z_{it}] = \beta_i' E[\widetilde{X}_t|z_{it}] + \theta' E[y_{it}|z_{it}] = \theta' E[y_{it}|z_{it}] \neq 0$. Now that if equation (2.1.1) is the correct specification for systematic component of r_{it}, it follows that

$$r_{it} = \{\beta_i' \widetilde{X}_t + \theta' E[y_{it}|z_{it}]\} + \theta' \{y_{it} - E[y_{it}|z_{it}]\} + v_{it}. \quad (2.1.3)$$

In other words, even though z_{it} is not a vector of systematic components in equilibrium asset pricing kernel, it may appear that these variables are useful to specify or predict the asset returns. Hence, if Equation (2.1.3) is applied for the normal (expected) returns, the presumed abnormal returns will become $\tilde{\eta}_{it} = \theta' \{y_{it} - E[y_{it}|z_{it}]\} + v_{it} = \xi_{it} + v_{it}$. Namely, the abnormal returns are no longer the original idiosyncratic risks. In fact, they now contain the projection errors where the impact(s) of event(s) and other unobservable components are projected onto the presumed firm-specific variables.

The event studies based on these presumed abnormal returns will not necessarily represent the impacts of the events. Instead, the studies are based on the residual effects of the event(s) after projecting the stock returns on these firm-specific variables. In other words, various empirical results may occur once different firm-specific diversifiable variables are applied.

Although screening out irrelevant noises to purify the abnormal returns may provide more reliable results for the hypothesized impacts from event(s), filtering the stock returns to obtain the abnormal returns must be handled with cares. Applications of more extended models (with factors or else) may not be more ideal unless the systematic components of stock returns are correctly identified. Inclusion of variables or attributes for normal (or expected) returns ought to take the nondiversifiability of these included variables into account. Overfitting the models with various firm-specific attributes for expected returns may not improve the power of statistical tests in the event studies.

In addition, there are numbers of research articles emphasizing either optimality or robustness in model selection criteria or other statistical procedures. Yet, few of the studies focuses on testing the essential nondiversifiable characteristics of the selected attributes or proxies for factors. Specifically, identification of statistically significant attributes or variables does not necessarily imply the selected ones are nondiversifiable or systematic.

More specifically, it is straightforward to find that even though $E[\tilde{\eta}_{it}] = \theta' E[y_{it} - E[y_{it}|z_{it}]] + E[v_{it}] = 0$,

$$\mathrm{Var}[\tilde{\eta}_{it}] = E[\tilde{\eta}_{it}^2] = E\{\theta'[y_{it} - E[y_{it}|z_{it}]][y_{it} - E[y_{it}|z_{it}]'\theta\} + E[v_{it}^2], \quad (2.1.4)$$

where $E[y_{it} - E[y_{it}|z_{it}]][y_{it} - E[y_{it}|z_{it}]]' = E[y_{it}y_{it}' - 2y_{it}E[y_{it}|z_{it}]' + (E[y_{it}|z_{it}])(E[y_{it}|z_{it}])']$ provided that $E[y_{it}E[y_{it}|z_{it}]'] = E[E[y_{it}\,(E[y_{it}|z_{it}])'|z_{it}] = E(E[y_{it}|z_{it}](E[y_{it}|z_{it}])'$.

This implies that $\sigma_{it}^2 \geq \mathrm{Var}[\tilde{\eta}_{it}]$ as long as $E[y_{it}|z_{it}] \neq \underline{0}$. In other words, the variance of the presumed abnormal returns will be less than the variance of genuine abnormal returns if a redundant event-related variable is included in the empirical asset pricing model for normal (expected) returns.

Hence, even if the null hypothesis is to claim that $E[\tilde{\eta}_{it}] = 0$, the reduced variance may make the test statistics tend to over-reject the null. In fact, as shown in Equation (2.1.3), it is easy to see that when $E[r_{it}|z_{it}] = \beta_i' E[\widetilde{X}_t|z_{it}] + \theta' E[y_{it}|z_{it}] = \theta' E[y_{it}|z_{it}] \neq 0$, there exists a specification error that some statistically significant variables (or attributes) are omitted in the (conditionally) expected returns if one wants to build a comprehensive model for stock returns. Namely, the variables z_{it} contribute some (in-sample) predictability for the stock returns, indeed. And, in particular, in statistical sense, the model simply based on the systematic attributes alone may become insufficient since it omits some significant variables.

However, the myth is "are these variables essential for asset pricing for capital market equilibrium?" May be not. The reason is that these variables (although informative) may become diversifiable by portfolios constructed with sufficiently large number of assets when these variables or phenomena are well-known to the investors. In other words, they may not be the pricing kernel for stock returns in capital market equilibrium. Therefore, model specification for asset pricing models should not simply identify statistically significant variables that may explain (or predict) the asset returns. Namely, verification of empirical asset pricing models simply bases on (statistical) predictability is not enough.

It is also common to see that the predictability (with finite samples collected) of these statistically significant variables may vary over different time periods. Some variables included may perform better than the others in certain time periods indeed. Yet, this superiority does not prevail on all occasions. Otherwise, the discussions over empirical asset pricing models should end up with a prespecified set (or category) of explanatory variables already. This however, is not the case. Variables that plausibly explain stock returns may change over time, especially in financial time series. Therefore, verifications of empirical asset pricing models should not limit the scope to statistical (or time-series) predictability alone. Further verification (perhaps, recursively if possible) on the nondiversifiabilty or systematic essentiality of these included variables is needed.

On the other hand, if assuming that the corporate finance event is relevant and within the time period (say, $[T_1, T_2]$), the mean of y_{it} is nonzero. That is, $E[y_{it}] \neq 0$, when $t \in [T_1, T_2]$, $t \leq T_1 < T_2 < T$, $[0, T]$ is the entire time period sampled. The property that $E[\tilde{\eta}_{it}] = \theta' E[y_{it} - E[y_{it}|z_{it}]] + E[v_{it}] = 0$ still holds. Yet, the statistics (for event studies) based on $\{\tilde{\eta}_{it}\}_{t=1,2,\cdots}$will have a zero mean uniformly when the variable $\{z_{it}\}_{t=1,2,\cdots}$ is included in the model of normal (expected) returns where $E[\epsilon_{it}] = \theta' E[y_{it}] \neq 0$, when $t \in [T_1, T_2]$. In that case, the statistics based on $\{\tilde{\eta}_{it}\}_{t=1,2,\cdots}$ will have a bias (for corporate finance events) within the time period $[T_1, T_2]$. Based on the above arguments, a proposition can be easily stated in the following.

Proposition 2.1.1: Given the setting in the equations (2.1.1) and (2.1.2), $\sigma_{it}^2 \geq \text{Var}[\tilde{\eta}_{it}]$ as long as $E[y_{it}|z_{it}] \neq \underline{0}$, whether $E[y_{it}] = 0$ or not. If, on the other hand, $E[y_{it}] \neq 0$, when $t \in [T_1, T_2]$, $t \leq T_1 < T_2 < T$, the abnormal returns $\tilde{\eta}_{it}$ (when additional explanatory variable z_{it} is included in the model for normal (expected) returns) will have a bias for the actual abnormal returns of corporate-finance events within the event period $[T_1, T_2]$.

From Proposition 2.1.1., it noticeable that inclusion of redundant event-related explanatory variable(s) in the empirical asset pricing model for normal (or expected) returns will not only influence the power of hypothesis tests, but also introduce bias in abnormal returns for the event period. For empirical finance, it is easy to see that almost all model builders will endeavor to search for some explanatory power from perceived information. For instance, Saporoschenko (2011) studied the effect of Santa Ana wind condition on the southern California stock returns. Fortunately, there is no statistically significant evidence in showing that the wind condition can predict the individual stock returns. Otherwise, if there is such a predictability, should one also include this variable as a systematic component in the specification of expected returns? Predictability of included variables does not necessarily imply nondiversifiability.

Hence, conventional applications with model selection criteria such as information criteria with prediction errors do not necessarily identify the systematic components for the asset pricing entirely unless additional criteria for nondiversifiability of included variable(s) are attached. Identification for empirical asset pricing models for excess returns is much more than just the statistical significance or predictability of the plausible explanatory variables. Although verifications of predictability (with model selection or else) are still considered essential for dynamic modeling of stock returns, the task for identification of (system-wise) asset-pricing factors remain undone unless nondiversifiability is also investigated.

This consequence of overfitting is confirmed by the empirical study of Ang and Zhang (2004) when a Fama-French three-factor model is applied with an additional momentum-related factor to approximate the normal returns. The inclusion of this additional (possibly firm-specific) variable tends to cause over-rejection of the test statistics in event study. Furthermore, the inclusion of additional factor(s) for the Fama-French three-factor model was unnecessary as confirmed in Jeng and Liu (2012) that no additional nondiversifiable risk (or factor) is needed when recursive forecast errors from the Fama-French three-factor model are applied in diagnostic tests for hidden nondiversifiable factors. Although many analyses (such as Collins and Dent 1984; Bernard 1987) had discussed that cross-sectional dependence or event clustering among the firm's abnormal returns may provide reduced standard deviation that causes over-rejecting *t* statistics, the above result shows that overfitting the asset pricing models for normal (expected) returns may also be equally distressful on empirical event studies.

Kothari and Warner (2007) state that "...The search in the Fama and French (1993) three-factor model, was further modified by Carhart (1997) to incorporate the momentum factor. However, absent a sound economic rationale motivating the inclusion of the size, book-to-market and momentum factors, whether these factors represent equilibrium compensation for risk or they are indication of market inefficiency has not been satisfactorily resolved in the literature." In fact, it is clear that if one includes some factors or attributes that are genuinely unsystematic or diversifiable in empirical asset pricing models, the consequence will very possibly ruin the soundness of the test statistics in verifying the hypotheses of event studies.

While it is promising to increase the dimensionality of explanatory variables in forming asset pricing models, the fact is that more complicated models are not always better. For instance, Cable and Holland (1999) apply the general-to-specific model selection framework to search for alternative models of normal returns in event studies. Their findings show that the market model is preferred to the capital asset pricing model. Brown and Warner (1980) state that the market- and mean-adjusted returns are more robust than the more complex models. Pettengill and Clark (2001) instead consider that market model provides biased results. Thompson (1988) shows that the expansion of market model to market-industry model, which may include industry indices, return form, and individual firm extraneous events, has very little impact on event study results. Yet, it is obvious that the ultimate determinant for the inclusion of variables should rely on the "systematic" or "nondiversifiable" property of the presumed explanatory variables.

Justification on which model is more appropriate for the normal returns should take us back to the concept or perspective of systematic risk. The included variables should not be based merely on statistical significance in either model selection criteria or tests. More importantly, the verification should be deeper in finding the nondiversifiability of these included variables or attributes. In all these cases of model specification for normal returns, the arguments may be more succinctly solved by further investigation on the nondiversifiable or systematic features of the included variables or attributes in empirical asset pricing models for normal returns.

Alternatively, Thompson (1989) discusses the so-called difference-in-return (DIR) model to obtain the abnormal returns. Specifically, the DIR model for an individual firm is shown as

$$r_{it} - r_{ct} = \varepsilon_{it}, \tag{2.1.5}$$

where r_{it} is the asset return for firm i and r_{ct} represents the control firm's rate of return. The control firm is selected as the compatibility of systematic risk, features of operations and industry, and in particular, their accounting variables (such as book-to-market ratio, earning per share, total capitalization, etc.) and other information. According to Thompson (1989), the t-statistics for the DIR model the conventional simple linear regression using the control firm's rate of return (denoted as t_{DIR} and t_m, respectively) can be shown for t in the event window and $[1, T]$ as the estimation period,

$$t_{DIR} = \frac{(r_{it} - r_{ct})}{\sqrt{\frac{1}{T}\sum_{\tau=1}^{T}(r_{i\tau} - r_{c\tau})^2}}, \tag{2.1.6}$$

$$t_m = \frac{(r_{it} - \hat{\alpha}_i - \hat{\beta}_i r_{ct})}{\hat{\sigma}_c}, \tag{2.1.7}$$

where $\hat{\sigma}_c = \hat{\sigma}_\varepsilon \left(1 + \frac{1}{T} + \frac{(r_{ct} - \overline{r}_c)^2}{\sum_{\tau=1}^{T}(r_{c\tau} - \overline{r}_c)^2}\right)^{\frac{1}{2}}$, $\hat{\sigma}_\varepsilon$ is the root mean square error of the simple linear regression, $\hat{\alpha}_i$ and $\hat{\beta}_i$ are estimates based on estimation period. The simulation result of Thompson (1989) shows that the DIR model does not improve the power of the tests. In other words, the DIR model in obtaining the abnormal returns may not be more ideal than simple linear regression (such as market model) even though the control firms are well chosen to fit the characteristics of the firms of interest.

One of the difficulties is that the control firms must be selected properly to reflect the essential (systematic) components for the firms of interest. Selection of these firms may not necessarily be easier than the model selection among all alternatives for normal (or expected) returns. Nevertheless, the finding in Thompson (1989) actually indicates that controversies in obtaining the abnormal returns through different methods of specifying normal (or expected) returns must be resolved so that certain principles are devised for consistency in empirical findings of the corporate events. More specifically, although a vast amount of literature and publications have discussed the corporate finance events such as mergers and acquisitions, confusion may still persist if the methods applied to obtain the correct and systematic specification of normal returns remain unclarified.

Some theoretical papers have discussed the factor structure of asset returns. Bai (2003) provides an inferential theory of factor structure of large dimension where both the dimensions of cross-sectional data and time series can tend to infinity. In his framework, the large-dimensional panel data can be expressed as

$$X_{it} = \lambda_i' F_t + e_{it}, \tag{2.1.8}$$

where $i = 1,2,...,N$, $t = 1,2,...T$, X_{it} is the ith dependent variable, λ_i is also a r-by-1 vector of factor loadings, F_t is a r-by-1 vector of *true* factors (observable or not). More compactly, it can be expressed as

$$X_t = \Lambda F_t + e_t, \tag{2.1.9}$$

where $X_t = (X_{1t}, X_{2t}, ...X_{Nt})'$ is a N-by-1 vector of dependent variables, $\Lambda = (\lambda_1, \lambda_2, ...\lambda_N)'$ is a N-by-r matrix of (nonstochastic) factor loadings, and $e_t = (e_{1t}, e_{2t}, ..., e_{nt})'$ is a N-by-1 vector of idiosyncratic risks. To identify that these factors are fundamentally essential, Bai (2003) imposes the following conditions for factors and factor loadings:

$$E||F_t||^4 \leq M < \infty, \qquad T^{-1}\sum_{t=1}^{T} F_t F_t' \xrightarrow{p} \Sigma_F$$

$$||\lambda_i|| \leq \bar{\lambda} < \infty, \qquad ||\frac{1}{N}\Lambda'\Lambda - \Sigma_\Lambda|| \longrightarrow 0,$$

where $||A|| = [tr(A'A)]^{\frac{1}{2}}$ denotes the (Frobenius) norm of the matrix A, Σ_F and Σ_Λ are both r-by-r positive-definite matrices for factors and factor loadings, respectively.

Notice that these conditions are simply to ensure that the factors are not degenerating over time and the factor loadings are somewhat converging toward a positive-definite full rank matrix when equally weighted by N. Alternatively (although not entirely equivalent), since the condition on factor loadings is to ensure that the matrix $\frac{1}{N}\Lambda'\Lambda$ will converge to a positive definite matrix Σ_Λ (in Frobenius norm), a necessary condition of these r essential factors ($r \geq 1$) can be shown that for all factor loadings $\{\beta_{ij}\}_{i=1,2,\ldots N}$, of the j-th factor in asset return i, $j = 1,2,\ldots,r$, where $\lambda_i = (\beta_{i1},\beta_{i2},\ldots,\beta_{ir})'$

$$\frac{1}{N}(\sum_{i=1}^{N}\beta_{ij}^2) \longrightarrow \sigma_r > 0.$$

In other words, the squared sum of factor loadings of each factor is growing with number of assets N and with a growth rate $O(1)$. That is, these squared sums of factor loadings will be growing with the same rate as number of assets N. Loosely speaking, the factor loadings are not squared-summable. Instead, they are (proportionally) growing with the number of assets N.

This shows that the factor loadings of these essential factors or attributes in the factor models of normal (or expected) returns must exhibit certain properties so that they are not degenerated as number of assets increases. Namely, the identification of variables or attributes (especially in the empirical asset-pricing models) that bases merely on statistical significance of parameters of these variables is not enough. Further justification of the above essentiality should be included also. Hence, statistical inferences on the empirical asset pricing models when some observable (economic) variables or attributes are applied should include the verification on the “systematic” property of the variables or attributes. Otherwise, the empirical asset pricing models may not represent genuine specification for the normal (or expected) returns. In turn, these modeling errors will cause incorrect assessments of abnormal returns and hence reduce the power of statistical tests for the event studies. In other words, choosing to have a conservative perspective to verify the inclusion of these variables (in addition to statistical significance in parametric or nonparametric factor analyses) will enhance reliability of event studies in corporate finance when using the abnormal returns.

One difficulty in Bai's (2003) work is that even though the factors and their factor loadings are estimable, there is no economic interpretation on these factors estimated. In particular, the factors are usually unobservable and hence are not applicable to filter the normal (nondiversifiable) components empirically from stock returns. However, many other applications of empirical asset pricing models also failed to justify that the included attributes satisfy the condition of factor loadings as shown above.

For instance, it is prevalent to discover that sizes, industry indices, or so-called own-volatility measures that are related with individual stock returns. However, they are not essentially systematic so that the entire capital market will flutter on the variation of these variables. Undoubtedly, these variables are well-known to the public—unless these institutions attempt to conceal them. If available in the public domain, why are the investors not making use of the firm-specific information to diversify their portfolios? If they are called "firm-specific," why is it they are still systematic components of normal (or expected) returns (in capital market equilibrium)? Or, is it not true that these firm-specific signals are orthogonal to the so called "systematic" variables (such as market index return or variables included in earlier Fama-French three-factor model)? If proper interpretation is attempted for these extended asset pricing models, these models should be furthermore verified with their nondiversifiable characteristics or factor-loading conditions for these included attributes. Otherwise, any empirical finding with statistically significant explanatory variables can be applied to explain the normal (or expected) returns, which in turn, must be based on equilibrium condition of capital market.

Certainly, another alternative for model search is to consider the model selection criteria such as AIC, or BIC, among others. Unfortunately, model selection criteria are only to satisfy the role as selecting statistically significant explanatory variables for the regression models of interest. For model search in asset pricing models or normal returns in stock returns, additional properties with concerns on nondiversifiability for explanatory variables must be considered. Otherwise, the dichotomy of stock returns into systematic (nondiversifiable) components and nonsystematic firm-specific abnormal returns will not hold empirically since some firm-specific variables may be identified as statistically significant predictors included in the systematic (nondiversifiable) components. In other words, this shows a contradictory remark that the so-called firm-specific diversifiable variables are considered as necessary explanatory variables for the systematic components, which are instead nondiversifiable. If so, the event studies that are based on the abnormal returns will not be robust enough to provide convincing evidence for the empirical findings.[1]

In the following, the analysis for model adequacy of empirical asset pricing models with pre-specified (economic) variables or attributes is discussed. The notion is that the identification of nondiversifiability on included or hidden factor(s) will result in cross-sectional strong dependence where long-memory models may be applied. The purpose in the following analysis is basically, diagnostic for the model specification when the need for additional (economic) variables is considered.

Chudik et al. (2011) consider the cross-sectional weak and strong dependence of large panels. Their framework defines the weak and strong cross-sectional dependence using weight sequences that satisfy the granularity condition where the vector norms of these weights are bounded by the growth rate $N^{1/2}$, where N represents the number of cross-sectional items in panel data. In addition, the weak and strong cross-sectional dependence is defined as whether the conditional variance of the weighted sequence of panel data observations at a certain point in time will converge to zero or not.

Although similar, the framework here differs from that of Chudik et al. (2011) in the sense that the (diversifying) weight sequences are subject to less restrictive conditions and applied to the entire opportunity set of well-diversified (efficient) portfolios. In contrast to their work, a stronger condition for nondiversifiability (for factor(s)) is introduced. A formal proof is also shown that the nondiversifiability of the (hidden) factor will lead to the cross-sectional long dependence. In particular, the long memory tests such as cross-sectional KPSS test, cross-sectional rescaled variance test, and assessments on Hurst exponent are also introduced.

While Chudik et al. (2011) apply the conditional variance assumption for cross-sectional dependence to consider the dynamic factors, the analysis here focuses on the static factor structure similar to that of Bai's (2003). Even though differing from these previous contributions, the intent here is to develop an alternative method in searching for expansion of dimensionality of model specifications for normal returns in empirical asset-pricing models where interpretable economic attributes are included. In addition, to avoid overfitting the model, a sequential model search starting from the most primitive stage is suggested. In other words, the model search is a forward-looking, specific-to-general approach. At each stage (for expansion of model dimensionality), a diagnostic test is applied to see if there is a hidden nondiversifiable factor in the presumed idiosyncratic risks that renders the need for model expansion.

To develop the analysis, there are a few prerequisites. Definition 2.D.1 describes a Hilbert space of (real) squared-integrable random variables defined on the probability space . The excess returns of assets forms a subset of this Hilbert space. Definition 2.D.2. is to define the set of diversifying weights. Definition 2.D.3. is for defining diversification. Assumption 2.A.1 shows the conditions of factors or pre-specified reference variables applied as factors. The assumption allows different choice sets of instrumental or reference variables applied to specify the risk premium. These choice sets may be evolving over time or different sample sizes. In particular, the reference variables may be generated by the innovations from the multivariate time series models of prespecified economic variables in a conditional expectation approach. Assumption 2.A.2 is to describe the presumed multifactor model of the asset returns. And, in particular, the assumption states that the initial stage of model search (including the most primitive case where only a drift is included) is already provided. Hence, the analysis is to identify whether there is a need to expand the dimension of the model for normal returns.

In fact, Assumption 2.A.2 may also apply to the studies that some similar works such as Bai (2003), Chudik et al. (2011), or Bailey et al. (2012) where their analyses are applied for the initial stage of exploratory data analysis to identify cross-sectional dependence among stock returns. Specifically, the following framework is to devise diagnostic statistics for model search in explaining the expected returns and to overcome overfittings in empirical finance. Assumption 2.A.3 shows the possible hidden factor and idiosyncratic risk in the excess return after projecting on the presumed explanatory variables. Assumption 2.A.4 provides a condition for diversifying weights for the infinite dimensional optimization problem. In addition, the sample size N of all firms is growing asymptotically and much larger than the sample size for selected firms of event studies. More explicitly, the sample size N includes both event-related and event-unrelated firms in identifying the model search for systematic components of asset returns.

Definition 2.D.1: Let $\mathscr{H} = L^2(\Omega, \mathscr{F}, P)$ be a Hilbert space of squared-integrable real-valued random variables (with respect to probability measure P on a complete probability space $(\Omega, \mathscr{F}, P)$, where $\mathscr{H}$ is endowed with L^2- norm$\|\|$ such that $\|x\| = (\int |x|^2 dx)^{1/2}$ for $x \in \mathscr{H}$. Let the inner product of $\mathscr{H}$ be denoted as $< x, y >= E[xy]$ for all $x, y \in \mathscr{H}$. Let $\{r_{it}\}_{i=1,2,\cdots,N,,t=1,2,\cdots T,}$ represent a sequence of all assets' excess returns contained in $\mathscr{H}$.

Assumption 2.A.1: Let $f_t = (f_{1t}, f_{2t}, \cdots, f_{pt})'$ be a vector of p proxies for factors in the information set Π_t, $p \geq 1$, and $E|f_{jt}| = 0$, where $E|f_{jt}|^2 = \sigma^2_{jt}$ (defined in $\mathscr{H}$) at time t for the presumed multi-factor pricing model, f_{jt} is the j-th proxy for factor at time t, for all $j = 1, 2, \cdots, p$, $t = 1, 2, \cdots, T$.

Assumption 2.A.2: Let the excess return r_{it} of each asset i at time t be regressed on the fitted $k-$factor structure ($k \geq 0$) with non-stochastic factor loadings as

$$r_{it} = E[r_{it}|\Pi_t] + \epsilon_{it} = \mu_i + \sum_{j=1}^{k} \beta_{ij} f_t^{(j)} + \epsilon_{it}, \qquad (2.1.10)$$

where $i = 1, 2, \cdots, N$, as randomly assigned subindices for asset returns, and $t = 1, 2, \cdots T$, where Π_t represents the information filtration (including lagged dependent variables and the proxies for factors) up to time t, $0 \leq k \leq p$, $\mu_i + \sum_{j=1}^{k} \beta_{ij} f_t^{(j)}$ stands for the conditional expected excess return $E[r_{it}|\Pi_t]$, where the selected factor $\left\{f_t^{(j)}\right\}_{j=1,\cdots,k} \subseteq (f_{1t}, f_{2t}, \cdots, f_{pt})'$ is a subset of known proxies for factors, and $\left\{f_t^{(j)}\right\}_{j=1,\cdots,k}$ don't have to follow the same order of designated sequence or indices for proxies in $f_t = (f_{1t}, f_{2t}, \cdots, f_{pt})'$, ϵ_{it} stands for the projection error (or so-called "presumed" idiosyncratic risk) for asset i at time t with the assumed multifactor pricing model.[2]

Assumption 2.A.3: Let the projection error (if contains a hidden factor) be expressed as a linear model such that $\epsilon_{it} = \eta_{it} + \nu_{it} = \beta_i^h f_{ht} + \nu_{it}$, where f_{ht} represents a stationary stochastic hidden factor with a nondegenerated distribution, $E[f_{ht}] = 0$, $f_{ht} \in \mathscr{H}$, such that f_{ht} is orthogonal to all selected proxies or factors $\left\{f_t^{(j)}\right\}_{j=1,\cdots,k}$, and f_{ht} is cross-sectional stationary for all assets and inter-temporal independent over time.[3] The β_i^h represents the real-valued nonstochastic unobservable factor loading for asset i on the hidden factor f_{ht} for all $i = 1, 2, \cdots N$. ν_{it} is a mean-zero (diversifiable) random noise with finite moments and independent of $\{\eta_{it}\}_{i=1,2,\cdots,N,t=1,2,\cdots,T}$ at any time t such that $E[\nu_{it}\eta_{it}] = 0$ where $E[\epsilon_{it}|f_t] = 0$, $E|\nu_{it}|^2 < \infty$ for all $i = 1, 2, \cdots N$. In addition, let $\sup_i |\beta_i^h|^2 < \infty$, $i = 1, 2, \cdots N$, such that $\{\beta_i^h\}_{i=1,2,\cdots} \in B$, where $B \subseteq \ell_\infty$ is a proper factor-loading subspace of ℓ_∞-space, and ℓ_∞-space contains all sequences $\{x_i\}_{i=1,2,\cdots}$, such that $\sup_i |x_i| < \infty$.

Notice that Assumption 2.A.2 claims that the model expressed in equation (2.1.10) is a close approximation for the underlying data generating mechanism. Linearity here only represents a presumed multifactor pricing model that may closely approximate the conditional expectation for excess returns for information set Π_t. In addition, Assumption 2.A.3 imposes an asymptotic condition for the absolute factor loading(s) on the hidden factor. This is to ensure that as the size of portfolio expands, the factor loadings won't be exploding—provided that the excess returns are in the Hilbert space $\mathscr{H}$. To apply the ideas of diversification, a few definitions on the feasible weights in the factor pricing models are introduced in the followings. Assumption 2.A.4 is to formulate the diversification in an infinite dimensional opportunity set for sequence of feasible weights applied to each asset. The tail condition is given to prevent the diversifying weight from collapsing too soon. The definition of diversification functional is provided in 2.D.2. The notation $N \to \infty$ represents the numbers of assets N will grow sufficiently large.

Assumption 2.A.4: Let W be a compact subspace of ℓ_2 space endowed with the ℓ_2 norm that for any $y = (y_1, y_2, \cdots) \in \ell_2$, $\sum_{i=1}^{\infty} y_i^2 < \infty$, where W consists of all real bounded sequences of nondegenerated deterministic feasible weights $\underline{\omega}$, such that $\underline{\omega} \equiv \{\omega_i\}_{i=1,2,\cdots} \in \ell_2$ and $\underline{\omega} \notin \ell_2^o$, where ℓ_2^o contains sequences $\{\omega_i\}_{i=1,2,\cdots}$with finitely many nonzero weights. In particular, for all $\underline{\omega} \in W$, and for arbitrarily small $\epsilon > 0$, $\epsilon \in R$, there exists some $\gamma > 0$ (which may be distinctive for different $\underline{\omega}$'s) such that

$$\liminf_{\substack{N \to \infty \\ \sum_{i=1}^{N} \omega_i^2 < \infty}} \{|\omega_i|\}_{i=1,2,\cdots,N} L(N) N^{\gamma} = \sup_{\substack{N \to \infty \\ \sum_{i=1}^{N} \omega_i^2 < \infty}} \{ \inf_{m \leq N} (|\omega_i|)_{i=1,2,\cdots,m}\} L(m) m^{\gamma} \geq \epsilon, \tag{2.1.11}$$

where $L(j)$ is slowly varying function of j such that $\lim_{j \to \infty} \frac{L(\lambda j)}{L(j)} = 1$, for $\lambda > 0$, $\lambda \in R$. Also let $\limsup_{N \to \infty} [N(\varpi_N^2)] = 0$, where $\varpi_N = \sup(|\omega_1|, |\omega_2|, \cdots, |\omega_N|)$. Let the factor loadings of a hidden factor be given as $\underline{\beta} = \left(\beta_1^h, \beta_2^h, \cdots, \beta_N^h, \cdots\right)$, $\underline{\beta} \in B$, where $B \subseteq \ell_\infty$ is a proper factor-loading subspace endowed with ℓ_∞-norm such that $\sup_{i=1,2,\cdots,\infty} |\beta_i^h| < \infty$.

In the following, the Assumption 2.A.5 is to provide the objective functional for the diversification problem in the infinite dimensional setting. Furthermore, the definition 2.D5 is to consider the well-diversified portfolios such that the weighted sums of diversifiable random noises will converge to zero. Notice that the weight sequences in Assumption 2.A.5 does not require the weights to follow or be bounded with a particular exponential growth such as $N^{-\alpha}$, $0 < \alpha \leq 1$ uniformly for all N. Notice also that Assumption 2.A.4 does not state that these weight sequences are selected weights for the efficient, well-diversified portfolios yet. Instead, Assumption 2.A.4 only states the "tail" condition for the weight sequences and Assumption 2.A.5 then confines the weight sequences to those that can provide optimization for the diversification functional.

Assumption 2.A.5: There exists a W-continuous diversification functional $f\left(\underline{\omega},\underline{\beta}\right) : W \otimes B \rightarrow R$, defined as $f\left(\underline{\omega},\underline{\beta}\right) = E\left(\sum_{i=1}^{\infty} \omega_i \beta_i^h f_{ht}\right)^2$ for any given $\underline{\beta}$ in B, where the non-null solution for portfolio optimization as $\inf_{\omega \in W^\star} f(\underline{\omega},\underline{\beta})$ exists, in which $W^\star$is a closed non-null subset of W, for $\underline{\omega} \in W^\star$, $\sum_{i=1}^{\infty} \omega_i = 1$, $\sum_{i=1}^{\infty} \omega_i^2 < \infty$.

Definition 2.D.2: A portfolio with weight $\underline{\omega}$, $\underline{\omega} \equiv \{\omega_i\}_{i=1,2,\cdots} \in W^\star$ in the diversification problem as $\inf_{\omega \in W^\star} f(\underline{\omega},\underline{\beta})$ is denoted as efficient (or well-diversified) if the weighted sum for all idiosyncratic risks converges such that $\sum_{i=1}^{\infty} \omega_i \nu_i \xrightarrow{L_2} 0$, $\underline{\omega} \in W^\star$ as $N \rightarrow \infty$, where $\xrightarrow{L_2}$ stands for convergence in L^2-norm.

Given definition 2.D.2 and Assumption 2.A.5, and similar to Chamberlain (1983), it is straightforward to verify that $\sum_{i=1}^{\infty} \omega_i^2 = \lim_{N\rightarrow\infty} \sum_{i=1}^{N} \omega_i^2 = 0$ for all well-defined portfolios in $W^\star$. For any given time t, the idiosyncratic variance of a weighted portfolio with n numbers of assets and a hidden factor such that $E\left[f_{ht}^2\right] = \sigma_{ht}^2$ can be expressed as

$$\sigma_{pt}^2 = \sum_{i=1}^{N} \left(\omega_i^2 E\left[\beta_i^{h^2}\right]\right) \sigma_{ht}^2 + 2\sum_{i=1}^{N} \sum_{i\neq j}^{N} \omega_i \omega_j E\left[\beta_i^h \beta_j^h\right] \sigma_{ht}^2 + \sum_{i=1}^{N} \omega_i^2 \sigma_{\nu_{it}}^2. \tag{2.1.12}$$

If $\{\nu_{it}\}_{i=1,2,\cdots}$ are diversifiable, a well diversified portfolio has $\lim_{N\rightarrow\infty} \sum_{i=1}^{N} \omega_i^2 \sigma_{\nu_{it}}^2 = 0$. In other words, if $\inf_{n\rightarrow\infty} \left\{\sigma_{\nu_{1t}}^2, \sigma_{\nu_{2t}}^2, \cdots \sigma_{\nu_{Nt}}^2\right\} > 0$, then $\sum_{i=1}^{\infty} \omega_i^2 = \lim_{N\rightarrow\infty} \sum_{i=1}^{N} \omega_i^2 = 0$. Suppose not. Let $\lim_{N\rightarrow\infty} \sum_{i=1}^{N} \omega_i^2 = \delta > 0$, and $\inf_{n\rightarrow\infty} \left\{\sigma_{\nu_{1t}}^2, \sigma_{\nu_{2t}}^2, \cdots \sigma_{\nu_{Nt}}^2\right\} = \theta > 0$.

Then, $\lim_{N\to\infty}\sum_{i=1}^{N}\omega_i^2\sigma_{\nu_{it}}^2 \geq \theta\sum_{i=1}^{\infty}\omega_i^2 > 0$. This will violate the assumption that $\{\nu_{1t},\nu_{2t},\cdots\}$ are diversifiable.

That is, the sum of infinite series of $\left\{\omega_i^2\right\}_{i=1,2,\cdots}$ in $W^\star$ should be nil. For instance, if $\omega_i = \frac{1}{N}$, N is the number of assets, then $\lim_{N\to\infty}\sum_{i=1}^{N}\omega_i^2 = \lim_{N\to\infty}\frac{N}{N^2} = 0 = \sum_{i=1}^{\infty}\omega_i^2$. This does not imply that all weights $\{\omega_i\}_{i=1,2,\cdots}$ are all identically equal to zero for all $i = 1,2,\cdots,$.

Therefore, the idea of diversifiable/nondiversifiable factor focuses on the entire opportunity set of well-diversified portfolios for the infinite dimensional optimization problem. For the nondiversifiable factor, its hidden factor loadings should not be eliminated by any possible sequences of weights of the well-diversified portfolios in $W^\star$ as $N \to \infty$. Otherwise, there exists a possibility that certain sequences of weights of some well-diversified portfolios in $W^\star$ will eliminate that risk entirely.

In contrast, if the factor is diversifiable, then the factor loadings of this factor should be eliminated by all possible sequences of weights in the well-diversified portfolios. This is straightforward since $\epsilon_{it} = \eta_{it} + \nu_{it} = \beta_i^h f_{ht} + \nu_{it}$ and if f_{ht} is diversifiable, then the entire ϵ_{it} can be denoted just as another diversifiable idiosyncratic risk $\nu_{it}^\star$ as well even though there is a factor structure. If so, then any well-diversified portfolio should diversify it away since the impact of this hidden factor will become negligible as number of asset increases. Otherwise, there is a contradiction to the idea of well-diversified portfolios in Definition 2.D.2 since a well-diversified portfolio should eliminate the idiosyncratic risks asymptotically.

Hence, definition of nondiversifiable factors is not based on all possible weights for portfolios or on any weight sequence that follows (or is bounded by) a particular (exponentially) decaying rate(s) such as in Chudik et al. (2011). The logic is simple. If a hidden factor is diversifiable then, any efficient (well-diversified) portfolio (including the equally-weighted one) will diversify it away. Yet, when the equally weighted sequence (or a particular weight sequence) fails to eliminate the hidden factor (and its factor loadings), it does not ensure there exists a nondiversifiable factor in the idiosyncratic risks.

Caution must be applied since a nondiversifiable factor should not be eliminated by any weight sequence of all well-diversified portfolios, not the weights that are simply either bounded by or specified with a particular exponential rule such as $N^{-\alpha}$, $0 < \alpha \leq 1$ uniformly. In brief, diversifiability of the factor(s) should be based only on all well-diversified (efficient) portfolios, not known (or specific rules of) diversifying weights.

Definition 2.D.3: The hidden factor f_{ht} for the factor loading $\underline{\beta} = \left(\beta_1^h, \beta_2^h, \cdots, \beta_n^h, \cdots\right)$, $\underline{\beta} \in B$, where $B \subseteq \ell_\infty$ is denoted as $W^\star$-diversifiable if and only if $\inf_{\omega \in W^\star} f(\underline{\omega}, \underline{\beta}) = 0$ for all well-diversified portfolios $\underline{\omega}$, $\underline{\omega} \equiv \{\omega_i\}_{i=1,2,\cdots} \in W^\star$. On the other hand, the hidden factor f_{ht} is nondiversifiable in $W^\star$ if and only if $\inf_{\omega \in W^\star} f(\underline{\omega}, \underline{\beta}) \neq 0$ for all well-diversified portfolios in $W^\star$.

In Jeng and Tobing (2012), a similar framework is also established with two panel CUSUM-based tests developed. The study emphasizes the model search approach in applying the tests to the empirical data. The present study differs slightly from the previous one by focusing on the development on intensity of diversifiability, which also leads to (cross-sectional) memory conditions of the idiosyncratic risks. The intensity of diversifiability emphasizes on the convergence condition of the partial sums of CUSUMs of idiosyncratic risks—although incidentally, it can be shown that if this exponent (or, intensity of diversifiability) is set to zero, these partial sums will simply conform with (the cross-sectional version of) KPSS and rescaled variance test statistics (Giraitis et al. 2003).

In contrast, Chudik et al. (2011) define the intensity of cross-sectional dependence in panel data with a stronger condition by using conditional variance of weighted sums of random variables. To identify the difference between the weak and strong cross-sectional dependence in the panel data for their study, the definition applies the "granularity condition" that all weights are uniformly bounded by $N^{1/2}$ for the convergence in conditional variance of weighted cross-sectional sums of double-indexed processes. In particular, a strong dependence is to consider that there exists a sequence of weights that will not drive the conditional variance of weighted cross-sectional sums of double-indexed processes to zero. A strong or weak factor in their study in particular is defined by whether the weighted sums of factor loadings when using equal/exponential weights will converge to zero or not.

In that case, a strong factor identified in their study may not be a $W^{\star}$-nondiversifiable factor as defined in our study since (1) the $W^{\star}$-nondiversifiable factors must not be diversified away by weights of *all* well-diversified portfolios including equal/exponential weights or else; (2) unlike the "granularity condition" or bounds which applied uniformly to all weights in Chudik et al. (2011), only the asymptotic tail conditions of infinite sequences of diversifying weights are subject to the convergence rate conditions stated in equation (2.1.11) of Assumption 2.A.4.

The issue as whether the factor is nondiversifiable or not must refer to all weight sequences for all well-diversified portfolios. Specifically, the factor loadings of a nondiversifiable or systematic factor must not be diversified away with different weights available in the set of all well-diversified portfolios, not with only one sequence of diversifying weights. In contrast, the strong cross-sectional dependence (which may result from the strong factor) in Chudik et al. (2011) only refers to the existence of one sequence of weights that can not drive the conditional variance of partial sums of weighted variables to zero while this weight sequence may or may not be in the set of well-diversified portfolios. Especially, their definition of strong factor(s) refers only to "non-null" convergence of equally (or exponentially) weighted sums of factor loadings such as $\lim_{N\to\infty} \frac{1}{N^{\alpha}} \sum_{i=1}^{N} |\beta_i^h| = \kappa > 0$, where $0 < \alpha \leq 1$.

In brief, a hidden nondiversifiable factor certainly may become a strong factor in the framework of Chudik et al. (2011). But, the converse does not necessarily hold. Hence, the inference for nondiversifiable factor requires further elaborations. Theorem 2.1.2. provides a condition that a hidden nondiversifiable factor may cause cross-sectional long dependence in the projection error of the fitted regression when excess returns are regressed on a set of pre-selected proxies for factors.

Notice that Zhou and Taqqu (2006) show that a completely random reordering on the data will not influence the sum of covariances for the long-dependent time series. In other words, the results from Theorem 2.1.2. will not change even with a completely random orderings of excess returns cross-sectionally. However, this does not imply that any reordering will not change the dependence. Instead, it is only the complete randomization that will not change the dependence. For that matter, Theorem 2.1.2. and for the diagnostic tests in section 2.2, it is assumed that the excess returns of all selected firms are *of completely random samples cross-sectionally.*

Theorem 2.1.2: If there exists a f_{ht} as $W^{\star}$-nondiversifiable hidden factor (or component) for all well-diversified portfolios $\underline{\omega} \equiv \{\omega_i\}_{i=1,2,\ldots,} \in W^{\star}$ in $\epsilon_{it} = \eta_{it}^{h} + \nu_{it} = \beta_i^h f_{ht} + \nu_{it}$, then $\{\epsilon_{it}\}_{i=1,2,\ldots,}$is a cross-sectional long dependent series at any given time t such that $\sum_{j=0}^{\infty} |\sigma_{it}(j)| = \infty$, for each $i = 1, 2, \cdots$, $j = 0, 1, 2, \cdots$, where $\sigma_{it}^{\epsilon}(j) = E[\epsilon_{it}\epsilon_{(i+j),t}]$ is the cross-sectional covariance of ϵ_{it} and $\epsilon_{i+j,t}$ at time t. Conversely, if the idiosyncratic risks are covariance-stationary and have a cross-sectional long dependence such that $\sum_{j=0}^{N} \sigma_{it}^{\epsilon}(j) = N^{2H} L(N)$, $H > \frac{1}{2}$, uniformly for all well-diversified portfolios $\underline{\omega}$, $\underline{\omega} \equiv \{\omega_i\}_{i=1,2,\ldots,} \in W^{\star}$, where $L(N)$ is slowly varying function of N such that $\lim_{N \to \infty} \frac{L(\lambda N)}{L(N)} = 1$, for $\lambda > 0$, $\lambda \in R$, then there exists a nondiversifiable hidden factor in $\{\epsilon_{it}\}_{i=1,2,\ldots,}$.

Proof: The proof requires the following Lemma 1.a.

Lemma 1.a: Given the definition 2.D.1–2.D.3, there exists a local optimum for the constrained minimization problem of the diversification functional $\inf_{\omega \in W^{\star}} f(\underline{\omega}, \underline{\beta})$.

Proof of Lemma 1.a: It is straightforward to verify that the functional $f\left(\underline{\omega}, \underline{\beta}\right) : W \otimes B \to R$ is weakly lower semicontinuous such that for every sequence $\underline{\omega} \equiv \{\omega_i\}_{i=1,2,\cdots,}$that converge weakly to some $\overline{\omega} \in \mathscr{W}^{\star}$, where $\liminf_{n \to \infty} f\left(\underline{\omega}, \underline{\beta}\right) \geq f\left(\overline{\omega}, \underline{\beta}\right)$. Since the constrained set is closed such that every sequence in $\mathscr{W}^{\star}$contains a weakly convergent subsequence with limit belongs to $\mathscr{W}^{\star}$, $\mathscr{W}^{\star}$is also weakly sequentially compact. Hence, by Theorem 2.3 in Jahn (2007), there exists a local minimal solution for the constrained optimization problem.

Proof of Theorem 2.1.2: Given the model and Lemma 1.a, for sufficiently large $M < \infty$ and $N \to \infty$, it can be shown that for any well-diversified portfolio with N numbers of assets

$$
\begin{aligned}
\sigma_{pt}^2 &= \sum_{i=1}^{N} \left(\omega_i^2 \left[\beta_i^{h^2}\right]\right) \sigma_{ht}^2 + 2 \sum_{i=1,}^{N} \sum_{i \neq j, j=1}^{N} \omega_i \omega_j (\beta_i^h \beta_j^h) \sigma_{ht}^2 + o_p(1), \\
&\leq \sum_{i=1}^{N} \left(\omega_i^2 \left[\beta_i^{h^2}\right]\right) \sigma_{ht}^2 + 2 \sum_{i=1,}^{N} \sum_{i \neq j, j=1}^{N} |\omega_i||\omega_j||(\beta_i^h \beta_j^h)|\sigma_{ht}^2, \\
&\leq M \sigma_{ht}^2 \sum_{i=1}^{N} \omega_i^2 + 2 \sum_{i=1,}^{N} \sum_{i \neq j, j=1}^{N} |\omega_i||\omega_j||(\beta_i^h \beta_j^h)|\sigma_{ht}^2, \\
&\leq M \sigma_{ht}^2 \sum_{i=1}^{N} \omega_i^2 + 2 \overline{\omega}_N^2 \sum_{k=1}^{N} (N-k)\, |(\beta_i^h \beta_j^h)|\sigma_{ht}^2, \\
&\approx o(1) + 2 \overline{\omega}_N^2 \sum_{k=1}^{N} (N-k)\, |(\beta_i^h \beta_j^h)|\sigma_{ht}^2, \\
&\leq o(1) + 2 n \overline{\omega}_N^2 \sum_{k=1}^{N} |(\beta_i^h \beta_j^h)|\sigma_{ht}^2, \\
&\leq o(1) + 2 (\limsup_{N \to \infty} [N \overline{\omega}_N^2]) \sum_{k=1}^{N} |(\beta_i^h \beta_j^h)|\sigma_{ht}^2.
\end{aligned}
$$

If the hidden factor is nondiversifiable, it is conceivable to have $\lim_{N \to \infty} \sigma_p^2 > 0$. However, since $\limsup_{n \to \infty} [N \overline{\omega}_{n(o)}^2] = o(1)$, if $\lim_{N \to \infty} \sum_{k=1}^{N} |(\beta_i^h \beta_j^h)| < \infty$, it implies that $\sigma_p^2 = o(1)$ also. This will violate the assumption that the hidden factor is nondiversifiable. Hence, $\lim_{N \to \infty} \sum_{k=1}^{N} |(\beta_i^h \beta_j^h)|$ can not be finite if the hidden factor is non-diversifiable. In fact, $\lim_{N \to \infty} \sum_{k=1}^{N} |(\beta_i^h \beta_j^h)|$ must grow with a rate no less than n. Thus, $\lim_{N \to \infty} \sum_{k=1}^{N} |(\beta_i^h \beta_j^h)| \to \infty$. This implies that, if $E|f_{ht}|^2 <$

∞. $\{\beta_i^h f_{ht}\}_{i=1,2,\cdots}$ are of strong cross-sectional dependence and the idiosyncratic risks will have cross-sectional long memory. Conversely, if idiosyncratic risks are of cross-sectional long memory, they are $\mathscr{W}^\star$-nondiversifiable.

Conversely, suppose the hidden factor is $\mathscr{W}^\star$-diversifiable when cross-sectional long memory prevails. Since the definition of diversifiable factor(s) refers to all well-diversified weights in $\mathscr{W}^\star$, it follows that $\lim_{N\to\infty} |\sum_{i=1}^N \omega_i \beta_i^h|^2 = 0$ for all well-diversified portfolios in $\mathscr{W}^\star$. Choose a well-diversified portfolio $\underline{\omega} \in \mathscr{W}^\star$ with all positive weights for included assets, where $0 < a(N) \le \omega_i$ for all $i = 1, 2, \cdots N, \cdots$ and let $\lim_{N\to\infty} a(N) = 0$ such that $a(N) \approx o(1)$. Then, for any N, given that $\{\epsilon_{it}\}_{i=1,2\cdots}$ are of cross-sectional long memory, we have

$$\begin{aligned} |\sum_{i=1}^{N} \omega_i \beta_i^h|^2 \ge a(N)^2 |\sum_{i=1}^{n} \beta_i^h|^2 \\ \ge a(N)^2 N^{2H} L(N)^2 \\ \approx o(N^o) O(N^{2H}) \approx o(N^{2H}). \end{aligned}$$

That is to say, the lower bound of the variance of weighted hidden factor loadings does not converge to zero unless it was multiplied by N^{-2H}. This violates that $\lim_{N\to\infty} |\sum_{i=1}^N \omega_i \beta_i^h|^2 = 0$. Hence, the hidden factor with assumed cross-sectional long memory should be $\mathscr{W}^\star$-nondiversifiable.

Corollary 2.1.3: If the selected k proxies or explanatory variables $\left\{f_t^{(j)}\right\}_{j=1,\ldots,k} \subseteq \underline{f_t}$, $k \ge 1$, $\underline{f_t} = (f_{1t}, f_{2t}, \cdots, f_{pt})'$ in the linear approximation $\zeta_{it} = \mu_i + \sum_{j=1}^k \beta_{ij} f_t^{(j)}$ of equation (2.1.1) for the excess returns $\{r_{it}\}_{i=1,2,\ldots,N}$ are of nondiversifiable loadings, then $\{\zeta_{it}\}_{i=1,2,\ldots,}$ will also have cross-sectional long dependence as stated in Theorem 2.1.2.

Proof of Corollary 2.1.3: Applying the result in Theorem 2.1.2, it is straightforward to see that for each included non-diversifiable explanatory variables $\left\{f_t^{(j)}\right\}_{j=1,\ldots,k}$, the series $\left\{\beta_{ij} f_t^{(j)}\right\}_{j=1,2,\cdots,k}$ are also of cross-sectional long dependence.

Since $\zeta_{it} = \mu_i + \sum_{j=1}^{k} \beta_{ij} f_t^{(j)}$ is a linear combination of $\left\{f_t^{(j)}\right\}_{j=1,\ldots,k}$ with non-null factor loadings $\left\{\beta_{ij}\right\}_{i=1,2,\cdots,N,}$ for $j = 1, 2, \cdots, k$, it will also be of cross-sectional long dependence.

Notice that in Bai's (2003) assumption of factor loadings, the Forbenius norm of factor loading matrix is shown as $||\frac{\Lambda_o'\Lambda_o}{N} - \Sigma_N|| \to 0$ when $N \to \infty$, for some $r \times r$ positive definite matrix Σ_N, where Λ_o is an $N \times r$ matrix of factor loadings for N asset returns of r factors. This implies that the factor loading matrix $\frac{\Lambda_o'\Lambda_o}{N}$ converges to a positive definite matrix coordinate-wise when $N \to \infty$. Accordingly, if the number of hidden nondiversifiable factor is equal to one, the convergence of Frobenius norm for factor loadings shows $|\frac{\sum_{j=1}^{N} \beta_j^{h2}}{N} - \tilde{\sigma}^2| \to 0$, for $\tilde{\sigma}^2 > 0$ as $N \to \infty$. This, in turn, shows that $\sum_{j=1}^{N} \beta_j^{h2}$ is growing with the same rate as N and $\sum_{j=1}^{N} \beta_j^{h2} \to \infty$ when $N \to \infty$. Since $\left(\sum_{j=1}^{N} |\beta_j^h|\right) \geqq \left(\sum_{j=1}^{N} \beta_j^{h2}\right)^{\frac{1}{2}}$ for all N, and $\ell_1 \subset \ell_2$ where the space of squared-summable real sequences encompasses the space of absolute-summable sequences of real numbers, this implies $\lim_{N\to\infty} \sum_{j=1}^{N} |\beta_j^h| \to \infty$.

Pick an arbitrary asset and denote its factor loading as β_o^h, where $\beta_o^h \neq 0$. It can be shown that $\lim_{N\to\infty} \sum_{j=1}^{N} |\beta_o^h||\beta_j^h| = \sum_{j=1}^{\infty} |\beta_o^h \beta_j^h| \to \infty$. Applying the definition of covariance in Theorem 2.1.2, and denoting $\rho_{ot}^{\epsilon}(j) = E\left[\epsilon_{ot}\epsilon_{jt}\right] = \beta_o^h \beta_j^h \sigma_h^2$, we can see that Bai's (2003) assumption on factor loadings will lead to the cross-sectional long memory such that $\sum_{j=1}^{\infty} |\rho_{ot}^{\epsilon}(j)| = \sigma_h^2 \sum_{j=1}^{\infty} |\beta_o^h \beta_j^h| \to \infty$. Notice that our condition for nondiversifiable factor loadings $\sum_{j=1}^{N} \beta_j^{h2} \to \infty$ is more general than Bai's assumption since there is no assumption for the growth rate of factor loadings to be $O(1)$.

Corollary 2.1.3 shows that the verification of nondiversifiabilty for included variables (or proxies) starts from the initial stage of model search where $k = 1$ and on. This implies that the model search should be sequentially implemented from $k = 1$

and look for further expansions of dimensionality. Detailed arguments are discussed in the latter sections after provision of statistical tests. Based on the above results, it is conceivable to apply the cross-sectional dependence to verify the existence of some nondiversifiable (hidden) factor(s) in the empirical asset pricing models.

Notice that the cross-sectional dependence here only shows that the ℓ^1–norm of the (cross-sectional) covariances will grow with the number of included assets at time t. Since the above claim is based on cross-sectional dependence for any given time t, this result holds even if the factor loadings are time-varying and stochastic — as long as the factor loadings and factors (included or hidden) are mutually independent. Although there are many different definitions for long memory in time series literature, the result here applies Parzen's (1981) definition. In addition, given various definitions of long memory (Guégan 2005, Heyde and Yang 1997), the existence of long memory does not necessarily imply there exists a unit root (similar to the time series models) for the cross-sectional dependence.

In brief, the conventional statistical tests or model selection criteria when applied in expansion for model dimensionality may not be robust enough to verify the asset pricing models if no further identification on the nondiversifiability is considered. Recall the earlier setting in equation (2.1.10), it is not difficult to find that (based on the samples selected) some (economic) attributes may be associated with the asset returns in empirical findings as long as $E[y_{it}|z_{it}] \neq \underline{0}$. Or, more specifically, it may indicate that the variable z_{it} contains certain predictability for the excess returns. Yet, such a finding does not necessarily justify that the inclusion of the variable(s) z_{it} for the empirical asset pricing models across all assets in the equity markets. Diagnostic tests for the empirical asset pricing models with the verification of nondiversifiability should be devised to obtain normal (or expected) returns more explicitly so that the robust abnormal returns may result for further applications in event studies.

Namely, predictability of some identified social-economic variables or attributes for stock returns does not naturally guarantee that these variables are the asset-pricing factors or kernels. Instead, for each selected variable in empirical asset pricing models, diagnostic tests to verify the needs of inclusion should be applied together with model selection procedures.

Ouysse (2006) extends Bai's (2003) model selection approach of asset pricing models to the case when factors are observable. Based on the same framework in Bai (2003), the generalization includes two steps: the estimation of number of factors, and the identification of order when different factors are chosen sequentially. In other words, the framework considers the identification of essential factors needed for the asset returns, and with the preorderings to determine the selection of variables. However, assumption on factor loadings in Chudik et al. (2011), and Bai's (2003) is only a special case for nondiversifiability in normal (or expected) returns according to Assumption 2.A.5 and Definition 2.D.3.

In the following, a section is provided to consider an alternative method to identify the intensity of diversifiability of the (hidden) factors when empirical asset pricing models are provided. And in addition, a new model selection framework with sequential hypothesis testing is shown. As a result, the framework shows that the existence of nondiversifiable hidden factor can be discovered in using the cross-sectional long memory tests. Given the above results, and the vast amount of literature discussing the long dependence, there are many different approaches to analyze the cross-sectional dependence due to the diversifiability of (hidden) factor(s).

A relative advantage of the cross-sectional long dependence, however, is that it incurs fewer controversial issues such as the distinction between long-memory and trend-stationary time series. Nor is there a particular need to distinguish between the long memory and the regime switching since the factor loadings for the factor(s) are allowed to be different cross-sectionally. Tentatively, for the interest of study, the discussions here focus on the variance-type analyses and the associated tests. More specifically, for simplicity, the factor loadings for the hidden factor are assumed as nonstochastic and time-invariant to develop the test statistics.

2.2 Diagnostic Tests and Model Search

Various statistical specification tests or model selection criteria have been applied in many empirical asset pricing models to verify the proposed models. Nevertheless, most research works emphasize on the verifications of statistical significance of included variables, time series predictability, or practical implications. Unfortunately, to pursue further event studies based on abnormal returns, it is necessary (if not too demanding) that the empirical asset pricing models must possess the utmost pricing kernel (or core) particularly for the system-wise normal (or expected) returns when the capital market equilibrium is presupposed. Noticeably, if this is emphasized, the usual statistical properties such as out-of-sample predictability, time-varying volatility or many others are not sufficient enough to justify the presumed models for normal (or expected) returns. Thus, studies based on abnormal returns may turn out to be controversial given that the statistics provided may become the results of joint hypotheses are from either the model specification errors or the events of interest.

In this section, we apply the framework of cross-sectional CUSUMs for the (hidden) factor of presumed idiosyncratic risk in fitted multifactor pricing models. The analysis is to provide some alternative methods to identify the essential and nondiversifiable attributes or variables for empirical asset pricing models. Differing from conventional specification tests or model selection criteria applied on the empirical asset pricing models, the alternative methods provide the sequential model search with cross-sectional hypothesis testings. Following from Theorem 2.1.2, the statistical tests are mainly to verify the cross-sectional strong dependence for included and possibly hidden nondiversifiable factor(s).

Assuming that model search starts with a preliminary specification (for instance, only a drift is included), the model search procedure will assess if there is a need to expand the dimensionality by some diagnostic test statistics. The analysis states that partial sums of the cross-sectional CUSUMs of equally weighted idiosyncratic risks may become finite depending on the underlying (hidden) factor's diversifiability. Specifically, these partial sums when multiplied with certain exponentially decaying weight sequences will turn finite according to the diversifiability of the hidden factor.

In particular, these partial sums (of equally weighted idiosyncratic risks) will converge to a functional of fractional Brownian motion where these decaying rates (of these weight sequences) are related with the Hurst exponent in the fractional

Brownian motion if fractional invariance principle is applied for these CUSUMs. In brief, these exponents that represent the intensity of diversifiability of (hidden) factor for equally weighted portfolios may influence the cross-sectional memory condition of idiosyncratic risks.

Although the finding confirms that the slow convergence rates for these partial sums will lead to cross-sectional long dependence, it may only demonstrates that certain decaying rates for the equally weighted portfolios becomes a sufficient (not necessary) condition for the cross-sectional long dependence for presumed idiosyncratic risks. In other words, with some specific assumptions for the decaying rates of these sequences, the result shows that they will lead to cross-sectional long dependence. Yet, the converse does not necessarily hold.

Therefore, the verification of cross-sectional long dependence (while not necessarily base on the decaying rates of equally weighted sequences) for hidden factor or included variables is only a diagnostic test to identify the need for expansion of dimensionality in empirical asset pricing models.

Assumption 2.A.6: Let $\lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} X_t\epsilon_{it} = \tilde{o}_p(1)$, for each asset i, $i = 1,2,\ldots,N,\ldots$, as $T \to \infty$, where $\tilde{o}_p(1)$ is a $(k+1)\times 1$ vector of zero's, $X_t = \left(1, f_t^{(1)}, f_t^{(2)}, \cdots, f_t^{(k)}\right)'$ represents the $(k+1)\times 1$ vector of included known common factors at time t. In addition, let $\lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} X_t X_t' = \widetilde{M}$, where $\widetilde{M}$ is a $(k+1)\times(k+1)$ nonsingular matrix. Also let $\lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \epsilon_{it} = o_p(1)$, $\lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} f_{ht}^2 < \infty$, and $\lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} f_{ht}v_{it} = o_p(1)$ for each asset i, $i = 1,2,\ldots N,\ldots$.

Notice that in Assumption 2.A.6, we consider the idiosyncratic risks (including the hidden factor) are orthogonal to the preselected common factors $X_t = \left(1, f_t^{(1)}, f_t^{(2)}, \ldots, f_t^{(k)}\right)'$ for simplicity. For further extensions, we may allow the chosen factors be correlated with the idiosyncratic risk (or hidden factor). However, in that case, the usual least-squared estimates will be biased and inconsistent for the factor loadings for each asset.

In particular, it is easy to show that even the application of common correlated effects (CCE) estimator proposed by Pesaran (2006) may not provide consistent estimates for factor loadings when the hidden factor is nondiversifiable and the orthogonality condition fails. Hence, to reduce the complexity in discussions, we assume the hidden factor is orthogonal to all preselected factors. Notice that the assumption does not assume the other proxies excluded from X_t are orthogonal to X_t, all these proxies for factors can be mutually correlated. However, the hidden or additional factor f_{ht} is assumed to be orthogonal to X_t.

Given that the equally weighted portfolio also satisfies the properties of well-diversified portfolios, it suffices to provide an example for hidden factor using equally weighted portfolio, according to Definition 2.D.2. Since under the null hypothesis that there is no nondiversifiable hidden factor in $\{\epsilon_{it}\}_{i=1,2,\cdots}$, the equally weighted portfolio will also diversify all hidden factors away.

Hence, for $t = 1,2,...,T$, let σ_{it}^2 be the variance of ϵ_{it}, where $E\left[\frac{1}{N}\sum_{i=1}^{N}\epsilon_{it}\right]^2 \to 0$ as $N \to \infty$, it is straightforward to see that

$$\lim_{N\to\infty} E\left[\frac{1}{N}\sum_{i=1}^{n}\epsilon_{it}\right]^2 = \lim_{N\to\infty}\frac{1}{N^2}\left(\sum_{i=1}^{N}\sigma_{it}^2 + 2\sum_{i=1}^{N}\sum_{\substack{j=1\\ i\neq j}}^{N} Cov(\epsilon_{it},\epsilon_{jt})\right) \longrightarrow 0. \tag{2.2.1}$$

In other words, under the null hypothesis, a diversifiable factor implies that $E\left[\frac{1}{N}\sum_{i=1}^{N}\epsilon_{it}\right]^2$ (which is the L^2-norm of $\frac{1}{N}\sum_{i=1}^{N}\epsilon_{it}$) will converge to zero as the number of assets N grows.

However, the convergence of the partial sums over these m subseries (or so-called partial sums of cumulative sums) in the equation (2.2.3) may depend on the growth rate(s) of the cumulative sums of these covariances as $N \to \infty$. [4] Likewise, it is feasible to define an intensity of diversifiability for equally weighted portfolios using the growth rate of the partial sums over these cumulative sums of m subseries.

We may use a simple example to demonstrate this intuition. Let $\sigma_{it}^2 = \sigma_t^2$, for all $i = 1, 2, \cdots, N$, without loss of generality. For N assets, the L^2-norm of each m subseries, $1 \leq m \leq N$ will be

$$E\left[\frac{1}{N}\sum_{i=1}^{m}\epsilon_{it}\right]^2 = \frac{1}{N^2}\left(m\sigma_t^2 + 2\sum_{i=1}^{m}\sum_{j=i+1}^{m}\sigma_{ij}\right). \tag{2.2.2}$$

Since the hidden factor is diversifiable, we have $\lim_{N\to\infty}\frac{1}{N^2}\left(m\sigma_t^2 + 2\sum_{i=1}^{m}\sum_{j=i+1}^{m}\sigma_{ij}\right) = 0$, which implies that $\lim_{N\to\infty}\frac{2}{N^2}\sum_{i=1}^{m}\sum_{j=i+1}^{m}\sigma_{ij} = 0$. However, the sums over these m subseries across all m's or so-called cumulative sums will become

$$\sum_{m=1}^{N}E\left[\frac{1}{N}\sum_{i=1}^{m}\epsilon_{it}\right]^2 = \frac{N(N+1)}{2N^2}\sigma_t^2 + \frac{2}{N^2}\sum_{m=2}^{N}\left(\sum_{i=1}^{m-1}\sum_{j=i+1}^{m}\sigma_{ij}\right). \tag{2.2.3}$$

It is easy to see that the firm term in equation (2.2.3) will be finite as $N \longrightarrow \infty$. Hence, we may expand the second term in equation (2.2.3) and show that

$$\frac{2}{N^2}\sum_{m=2}^{N}\left(\sum_{i=1}^{m-1}\sum_{j=i+1}^{m}\sigma_{ij}\right) = \frac{2}{N^2}\left\{\sum_{i=1}^{N}(N-i)\sigma_{1,i+1} + \sum_{i=2}^{N}(N-i)\sigma_{2,i+1} + \cdots + \sigma_{N-1,N}\right\}. \tag{2.2.4}$$

Notice that in equation (2.2.3), if the hidden factor is (strongly) diversifiable, partial sums of these covariances $\{\sigma_{ij}\}_{i\neq j, i,j=1,2,\ldots}$ in equation (2.2.4) should decay fast enough so that partial sum of the L^2–norm of equally weighted cumulative sums $\frac{1}{N}\sum_{i=1}^{m}\epsilon_{it}$ will become either zero or finite in (2.2.3) as $N \to \infty$. [5]On the other hand, if the entire weighted sum in (2.2.4) is growing with N also, the factor is less diversifiable. Specifically, it could be growing with an exponential growth rate such as N^{α}, where $0 \leq \alpha < 1$. In other words, even though the factor is diversifiable with equally-weighted portfolios, the intensity of diversifiability may differ, according to the unknown factor loadings of hidden factor.

Therefore, the magnitude of growth rate α can be applied as a regularity condition for diversifiability of the hidden factor under *equally weighted* scheme. Specifically, if one identifies that (with equally weighted sequence) there is no persistent cross-sectional dependence among these presumed idiosyncratic risks, it is safe to say that there is no hidden nondiversifiable factor for normal returns in empirical asset pricing models. The reason is that a diversifiable hidden factor should be diversified away in weight sequence(s) applied in every efficient portfolio by definition. Since equally weighted portfolio is also efficient according to Definition 2.D.3, it is conceivable to accept that the presumed idiosyncratic risks are diversifiable. Hence, no further expansion of dimensionality is needed.

However, if the test statistic based solely on the exponent α on the equally weighted portfolios rejects the null hypothesis such as $\alpha = 0$, the rejection of null hypothesis does not necessarily guarantee that there is a nondiversifiable hidden factor either—because the result is only verified by the equally weighted portfolio(s). Specifically, if the test rejects the exponent $\alpha = 0$ in the above example, there is a modeling risk that one may jump on the conclusion to claim there exists a nondiversifiable factor in the presumed idiosyncratic risks.

In fact, the statistical test based on the exponent may just verify that using only equally weighted portfolio can not diversify the presumed idiosyncratic risks away, and nothing more. Namely, the applications for the cross-sectional exponents (on factor loadings) are for convenience to generate feasible statistics. Yet, to the best extent, and for the investigation of (hidden) non-diversifiable factor(s), these test statistics with exponents can only be used for diagnostic purpose. They are not for the model specifications in empirical asset pricing models—if one would emphasize the equilibrium concept in asset pricing of capital markets.

The following analysis shows that the cross-sectional exponent $0 \leq \alpha < 1$ on the partial sum of CUSUMs of presumed idiosyncratic risks and provides a sufficient condition for the cross-sectional long dependence and certain statistics in long memory time series modeling. In other words, with certain assumptions for these cross-sectional exponents (on factor loadings or on the partial of their CUSUMs), it will still lead to the same weak convergence toward the functionals of fractional Brownian motion.

For simplicity, the following analysis assumes there exists only one hidden factor in the presume idiosyncratic risks. This is to demonstrate the necessity of identification on the nondiversifiability of included variables and the search for further expansion of dimensionality must be cautiously pursued. Namely, the expansion of the model(s) should start from the primitive model(s) and expand it to identify the (systematic) nondiversifiable components carefully.

If $\epsilon_{it} = \eta_{it} + \nu_{it} = \beta_i^h f_{ht} + \nu_{it}$, for finite samples we have

$$
\begin{aligned}
&\frac{1}{N^2}\sum_{m=1}^{N}\left[\frac{1}{T}\sum_{t=1}^{T}\left[\left(\sum_{i=1}^{m}\epsilon_{it}\right)^2 - \left(E\left(\sum_{i=1}^{m}\epsilon_{it}\right)\right)^2\right]\right],\\
&= \sum_{m=1}^{N}\left[\frac{1}{T}\sum_{t=1}^{T}\left(\left(\frac{1}{N}\sum_{i=1}^{m}\epsilon_{it}\right)^2 - \left(E\left(\frac{1}{N}\sum_{i=1}^{m}\epsilon_{it}\right)\right)^2\right)\right],\\
&= \frac{1}{T}\sum_{t=1}^{T}\sum_{m=1}^{N}\left[\frac{1}{N^2}\left(\sum_{i=1}^{m}\beta_i^h\right)^2 [f_{ht}^2 - \left(E[f_{ht}^2]\right)]\right] + o_p(1),\\
&= \left(\sum_{m=1}^{N}\left(\frac{1}{N}\sum_{i=1}^{m}\beta_i^h\right)^2\right)\left(\frac{1}{T}\sum_{t=1}^{T}[f_{ht}^2 - \left(E[f_{ht}^2]\right)]\right), \quad (2.2.5)
\end{aligned}
$$

uniformly for all N,T $N,T = 1,2,\cdots$. These equations hold even if the hidden factor is nondiversifiable.

For instance, if $\sum_{m=1}^{N}\left(\frac{1}{N}\sum_{i=1}^{m}\beta_i^h\right)^2 = O(N^{\alpha})$, $0 < \alpha < 1$, then $\sum_{m=1}^{N}\left(\frac{1}{N}\sum_{i=1}^{m}\beta_i^h\right)^2$ may not converge to a finite number as $N \to \infty$. Instead, we may need to multiply another $N^{-\alpha}$ to make it converge. In particular, following from the equations (2.2.3) and (2.2.4), this condition only applies to the partial sums of squared cumulative sums of equally weighted factor loadings, not directly on the factor loadings. Hence, no particular growth rate for the factor loadings is assumed. In fact, this growth rate is related to the cumulative sums of covariances in the equation (2.2.4).

Furthermore, due to the factor structure, the partial sums of cumulative weighted sums for factor loadings is factored out from the term, $\frac{1}{T}\sum_{t=1}^{T}(f_{ht}^2 - (E[f_{ht}^2]))$. In other words, the cross-sectional sample size N and the time series horizon T may grow jointly, with or without any further conditions. Specifically, these arguments still hold since both cross-sectional sample size and that of the time series can increase jointly.

We may also see that these partial sums of the cross-sectional CUSUMs of the equally weighted idiosyncratic risks are subject to different exponential growth rate with respect to N, such as N^{α}, depending on the factor loadings associated with the hidden factor. Notice that even though the weak law of large numbers holds such that $\frac{1}{T}\sum_{t=1}^{T}(f_{ht}^2 - (E[f_{ht}])^2) = o_p(1)$, if $\sum_{m=1}^{N}\left(\frac{1}{N}\sum_{i=1}^{m}\beta_i^h\right)^2 = O(N^{\alpha})$, it is obvious that $\left(\sum_{m=1}^{N}\left(\frac{1}{N}\sum_{i=1}^{m}\beta_i^h\right)^2\right)\left(\frac{1}{T}\sum_{t=1}^{T}E(f_{ht}^2 - (E[f_{ht}])^2)\right) = o_p(N^{\alpha})$.

In other words, the product of partial sum of squared CUSUMs of equally weighted (hidden) factor loadings $\{\beta_i^h\}_{i=1,2,\ldots,N}$ and $\frac{1}{T}\sum_{t=1}^{T}(f_{ht}^2 - (E[f_{ht}^2]))$ will become finite only if multiplied with $N^{-\alpha}$ as $N \longrightarrow \infty$ $T \longrightarrow \infty$. The same argument can be established if the factor loadings for the (hidden) factor are stochastic. In that case, the framework only needs to state that $\left(\sum_{m=1}^{N}E\left(\frac{1}{N}\sum_{i=1}^{m}\beta_i^h\right)^2\right) = O(N^{\alpha})$ and the factor loadings and the hidden factor are mutually independent.

For instance, as $\alpha > 0$, and $\frac{1}{T}\sum_{t=1}^{T}(f_{ht}^2 - (E[f_{ht}^2])) = o_p(1)$, these weighted CUSUMs do not converge since $\frac{1}{T}\sum_{t=1}^{T}\left(\sum_{m=1}^{N}E\left[\left(\frac{1}{N}\quad\sum_{i=1}^{m}\epsilon_{it}\right)^2 - \left(E\left(\frac{1}{N}\sum_{i=1}^{m}\epsilon_{it}\right)\right)^2\right]\right) \neq o_p(1)$ even when $N \to \infty$, $T \to \infty$. In particular, when $\alpha = 0$, then the partial sums of weighted CUSUMs of factor loadings in the hidden factor are finite when $N \to \infty$, without any additional weighting. That is, $\sum_{m=1}^{N}\left(\frac{1}{N}\sum_{i=1}^{m}\beta_i^h\right)^2 = O(1)$. This will imply that the hidden factor is well-diversifiable since the partial sum of squared sub-series (or CUSUMs) of equally weighted factor loadings converges.

In the following analysis, it is assumed that $E[\epsilon_{it}] = 0$ for all $i = 1,2,...,N$, for simplicity and hence, $E\left(\frac{1}{N}\sum_{i=1}^{m}\epsilon_{it}\right) = \frac{1}{N}\sum_{i=1}^{m}E[\epsilon_{it}] = 0$. Therefore, the second term in the equation (2.2.3) can be ignored. Notice that this condition is different from the "granularity" condition in Chudik et al. (2011). The condition here is based on the sum of cumulative sums of factor loadings, not partial sums of factor loadings. Also, the setting is different from Bailey et al. (2012) where they apply certain convergent rates for the cross-sectional factor loadings with respect to the cross-sectional sample size N. Although related, the current condition is weaker than the granularity condition. 6

Definition 2.D.4: Given the partial sum of cumulative sums for equally weighted idiosyncratic risks denoted as $\sum_{m=1}^{N}E\left(\frac{1}{N}\sum_{i=1}^{m}\epsilon_{it}\right)^2$, for any date $t = 1,2,...,T$, there exists a coefficient $\alpha \in R$, $0 \leq \alpha < 1$, such that $N^{-\alpha}\sum_{m=1}^{N}E\left(\frac{1}{N}\sum_{i=1}^{m}\epsilon_{it}\right)^2 < \infty$, as $N \to \infty$. The coefficient α is denoted as the intensity of diversifiability for equally weighted portfolios on the hidden factor in idiosyncratic risks, $\{\epsilon_{it}\}_{i=1,2,}$.

Definition 2.D.5: A random process $X(\tau)$ is called "self-similar" if and only if

$$X(\lambda\tau) \stackrel{d}{\equiv} \lambda^H X(\tau), \tag{2.2.6}$$

for $\lambda > 0$, $\tau \in R$, $\tau \geq 0$, where $H \geq 0$, is the Hurst exponent, the notation $\stackrel{d}{\equiv}$ stands for the equivalence of distributions. In addition, without loss of generality, we set the initial condition $X(0) = 0$ almost surely. According to Embrechts and Maejima (2002), for any self-similar process, $H = 0$ if and only if $X(\tau) = X(0), \forall\tau \geq 0$ almost surely. If $H = 1$, it implies that all autocorrelations of $X(\tau)$ are equal to one.

Definition 2.D.6: Let $0 < H \leq 1$, a zero mean Gaussian process $\left\{B_H(\tau), \tau > 0\right\}$ is called "fractional Brownian motion" if

$$E[B_H(\tau)\,B_H(s)] = \frac{1}{2}\left\{\tau^{2H} + s^{2H} - |\tau - s|^{2H}\right\} E[B_H(1)]^2. \tag{2.2.7}$$

In particular, when $\frac{1}{2} < H < 1$, the process will have long dependence. When $H = \frac{1}{2}$, $\left\{B_H(\tau), \tau > 0\right\}$ will become the usual Brownian motion with independent identically distributed increments. If $0 < H < \frac{1}{2}$, the process is called "anti-persistent" with sum of auto-covariances being finite.

More extensively, we also discover that the intensity of diversifiability of hidden factor (or convergence rate of the partial sums of cross-sectional CUSUMs of idiosyncratic risks) for equally weighted portfolios is then, related to the Hurst exponent of a fractional Brownian motion if the following invariance principle holds. Notice that the following invariance principle actually follows from the cross-sectional long dependence conclusion from Theorem 2.1.2 (even without the specific assumption for decaying rate α of equally weighted portfolios) and it includes the case when the cumulative sums converge to a standard Brownian motion under various mixing conditions when the Hurst exponent is equal to $\frac{1}{2}$. The following assumption is to consider that the partial sums of idiosyncratic risk will converge in distribution of a fractional Brownian motion, where the Hurst exponent H will depend on the cross-sectional memory condition of the $\{\epsilon_{it}\}_{i=1,2,\ldots}$.

Assumption 2.A.7: Assuming that the idiosyncratic risks $\{\epsilon_{it}\}_{i=1,2,\ldots}$ for each time t (for sufficiently large sample size N) follow the fractional invariance principle such that

$$\frac{1}{N^H \sigma_{H,t}} \sum_{i=1}^{[Nz]} \epsilon_{it} \xrightarrow{d} B_H(z), \tag{2.2.8}$$

where $0 < z \leq 1$, $[x]$ stands for the largest integer that is less than x, $x \in R^{\dagger}$, $B_H(z)$ is a fractional Brownian motion with the

Hurst exponent H, $0 \leq H < 1$, $\sigma^2_{H,t} = \lim_{N\to\infty} \frac{1}{N} E\left(\sum_{i=1}^{N} \epsilon_{it}\right)^2$, and the notation $\xrightarrow{d}$ stands for the convergence in distribution. In particular, with the definition of fractional Brownian motion, when $0 < H < \frac{1}{2}$, the stochastic process is anti-persistent, when $H = \frac{1}{2}$, it becomes the standard Brownian motion, and when $H > \frac{1}{2}$, the stochastic process will have long dependence.

Notice that Assumption 2.A.7 does not assume any particular convergence rate for the weight sequences for any portfolios or growth rates for factor loadings of interest. Assumption 2.A.7 only shows the weak convergence (or so-called convergence in distribution) for the tests of cross-sectional long dependence where the fractional Brownian motion is only an approximation for the partial sums of idiosyncratic risks. In particular, the assumption also includes the weak convergence in distribution to conventional Brownian motion.

In addition, the equation $\sigma^2_{H,t} = \lim_{N\to\infty} \frac{1}{N} E\left(\sum_{i=1}^{N} \epsilon_{it}\right)^2$, where $0 \leq H < 1$, shows the long-run variance of $\{\epsilon_{it}\}_{i=1,2,\cdots}$ for each time t, as $N \to \infty$. More specifically, this long-run variance incurs the Hurst exponent H. Hence, the conventional heteroskedasticity and autocorrelation consistent (HAC) estimate for the cross-sectional asymptotic variance can not apply. The Hurst exponent can also be time-varying if more specific conditions are provided.

Given Theorem 2.1.2, it is easy to see that the cross-sectional long dependence condition will result if the hidden factor is nondiversifiable with all weight sequences for well-diversified portfolios. This in turns, shows that the diagnostic test for nondiversifiability when there is no specific decaying rate or exponent (assumed by the equally weighted portfolios) is more in need. The result is shown in the Theorem 2.2.2.

Proposition 2.2.1 simply relates the statistics $\sum_{m=1}^{N} \left(\frac{1}{N} \sum_{i=1}^{m} (\epsilon_{it} - \overline{\epsilon_t})\right)^2$ with the functionals of fractional Brownian motion with the Hurst exponent. The arguments thereby shows that the magnitude of intensity of diversifiability of equally weighted portfolios will lead to the Hurst exponent of fractional

Brownian motion. In other words, the intensity of diversifiability (for equally weighted portfolios) on hidden factor will lead to the cross-sectional memory of idiosyncratic risk.

Proposition 2.2.1: Given Theorem 2.2.1, Definitions 2.D.5, 2.D.6, Assumption 2.A.7, and $\epsilon_{it} = \eta_{it} + \nu_{it} = \beta_i^h f_{ht} + \nu_{it}$, the partial sum $\sum_{m=1}^{N}\left(\frac{1}{N}\sum_{i=1}^{m}(\epsilon_{it} - \overline{\epsilon_t})\right)^2$ will converge to the integral of squared fractional Brownian bridge $\overline{B}_H(z)$ in distribution such that for $0 \leq H < 1$,

$$N^{-(2H-1)}\sum_{m=1}^{N}\left(\frac{1}{N}\sum_{i=1}^{m}(\epsilon_{it} - \overline{\epsilon_t})\right)^2 \xrightarrow{d} \sigma_{H,t}^2 \int_0^1 (\overline{B}_H(z))^2 dz, \tag{2.2.9}$$

as $N \to \infty$, for any date t, $t = 1,2,...,T$. If in addition, $N^{-\alpha}\sum_{m=1}^{N} E\left(\frac{1}{N}\sum_{i=1}^{m}\epsilon_{it}\right)^2 < \infty$, $0 \leq \alpha < 1$, $H = \frac{\alpha+1}{2}$,

$$N^{-\alpha}\sum_{m=1}^{N}\left(\frac{1}{N}\sum_{i=1}^{m}(\epsilon_{it} - \overline{\epsilon_t})\right)^2 \xrightarrow{d} \sigma_{H,t}^2 \int_0^1 (\overline{B}_H(z))^2 dz, \tag{2.2.10}$$

if $N \to \infty$, where $\overline{\epsilon_t} = \frac{1}{N}\sum_{i=1}^{N}\epsilon_{it}$, $\overline{B}_H(z) = B_H(z) - zB_H(1)$.

Proof of Proposition 2.2.1: It is easy to see that

$$\begin{aligned}
&N^{-(2H-1)}\sum_{m=1}^{N}\left(\frac{1}{N}\sum_{i=1}^{m}(\epsilon_{it} - \overline{\epsilon_t})\right)^2 \\
&= \frac{1}{N}\sum_{m=1}^{N}\left(N^{-H+\frac{1}{2}}\frac{1}{\sqrt{N}}\sum_{i=1}^{m}(\epsilon_{it} - \overline{\epsilon_t})\right)^2 \\
&= \frac{1}{N}\sum_{m=1}^{N}\left(N^{-H+\frac{1}{2}}\frac{1}{\sqrt{N}}\sum_{i=1}^{m}(\epsilon_{it} - \overline{\epsilon_t})\right)^2 \\
&= \frac{1}{N}\sum_{m=1}^{N}\left(N^{-H}\sum_{i=1}^{m}(\epsilon_{it} - \overline{\epsilon_t})\right)^2
\end{aligned}$$

$$= \frac{1}{N} \sum_{m=1}^{N} \left(N^{-H} (\sum_{i=1}^{m} \epsilon_{it} - \frac{m}{N} \sum_{i=1}^{N} \epsilon_{it}) \right)^2 .$$

Apply Assumption 2.A.7, and set $m = [Nz]$, where $0 \leq z \leq 1$, $[x]$ stands for the largest integer less than x, and the continuous mapping theorem of Pötscher (2004), it can be shown that as $N \to \infty$,

$$N^{-(2H-1)} \sum_{m=1}^{N} \left(\frac{1}{N} \sum_{i=1}^{m} (\epsilon_{it} - \overline{\epsilon_t}) \right)^2 \xrightarrow{d} \sigma_{H,t}^2 \int_0^1 (\overline{B}_H(z))^2 dz.$$

Given that $E\left[\left(\frac{1}{N}\sum_{i=1}^{m} \epsilon_{it}\right)\right] = \left(\frac{1}{N}\sum_{i=1}^{m} E[\epsilon_{it}]\right) = 0$, it is easy to set that

$$N^{-\alpha} \sum_{m=1}^{N} E\left(\frac{1}{N} \sum_{i=1}^{m} \epsilon_{it} \right)^2 = N^{-\alpha} \left(\sum_{m=1}^{N} Var\left(\frac{1}{N} \sum_{i=1}^{m} \epsilon_{it} \right) \right).$$

Set $H = \frac{\alpha+1}{2}$, the above result will give

$$N^{-\alpha} \sum_{m=1}^{N} \left(\frac{1}{N} \sum_{i=1}^{m} (\epsilon_{it} - \overline{\epsilon_t}) \right)^2 \xrightarrow{d} \sigma_{H,t}^2 \int_0^1 (\overline{B}_H(z))^2 dz.$$

Notice that the result in equation (20) does not require any assumption for the cross-sectional exponent or (bounds for) decaying rate of (hidden) factor loadings. Nevertheless, it is easy to see that the intensity of diversifiability for hidden factor is related to the Hurst exponent of a fractional Brownian motion asymptotically when the number of included securities increases. Apparently, the discussions of intensity of diversifiability of the hidden factor (with equally weighted portfolios) may lead to cross-sectional long dependence among the idiosyncratic risks also. However, the discussions are only limited to the equally weighted portfolios, not all efficient, well-diversified portfolios.

More specifically, when the intensity of diversifiability $\alpha > 0$, the Hurst exponent H will be greater than $\frac{1}{2}$, which implies that the idiosyncratic risks $\{\epsilon_{it}\}_{i=1,2,\ldots}$ will be subject to cross-sectional long dependence. In other words, the less equally weighted diversifiable the hidden factor is, the stronger the cross-sectional dependence in the idiosyncratic risks $\{\epsilon_{it}\}_{i=1,2,\ldots}$ as stated in Proposition 2.2.1. And thus, as the the number of included securities increases, identification of the extent of Hurst exponent of the asymptotic distribution of the cumulative sums of idiosyncratic risks $\{\epsilon_{it}\}_{i=1,2,\ldots}$ will encompass the study of nondiversifiability of hidden factor(s) with equally weighted portfolios. If $\alpha = 0$, the statistic $N^{-\alpha}\sum_{m=1}^{N}\left(\frac{1}{N\sigma_{H,t}}\sum_{i=1}^{m}(\epsilon_{it}-\overline{\epsilon_t})\right)^2$ will become $\frac{1}{N^2}\sum_{m=1}^{N}\left(\frac{1}{\sigma_{H,t}}\sum_{i=1}^{m}(\epsilon_{it}-\overline{\epsilon_t})\right)^2$, which is similar to a cross-sectional version of KPSS test statistics at any given date t, where σ_{Ht}^2 represents the long run variance of $\{\epsilon_{it}\}_{i=1,2,\ldots}$.

In other words, if the null hypothesis for the (hidden) factor is diversifiable, Equation (2.2.9) will be similar to the KPSS test for persistent long memory, and it can be applied to detect the strong cross-sectional dependence in these asset pricing models. By the same token, the weak convergence can also be applied to developed the cross-sectional version of rescaled variance test (Giraitis et al. 2003) whether the decaying rate or cross-sectional exponent is assumed or not. More explicitly, if cross-sectional exponent (for the factor loadings) is assumed such that $0 \le \alpha < 1$, then it follows that

$$N^{-\alpha}\left\{\frac{1}{N^2\sigma_{H,t}^2}\left[\sum_{m=1}^{N}\left(\sum_{i=1}^{m}(\epsilon_{it}-\overline{\epsilon_t}\right)^2-\left(\frac{1}{N}\right)\left(\sum_{m=1}^{N}\sum_{i=1}^{m}(\epsilon_{it}-\overline{\epsilon_t})\right)^2\right]\right\}$$

$$\xrightarrow{d}\int_0^1\left(\overline{B}_H(z)\right)^2dz-\left(\int_0^1\overline{B}_H(z)dz\right)^2. \qquad (2.2.11)$$

Therefore, if the underlying factor loadings for included variables or hidden factor are diversifiable with equally weighted sequences where $\alpha = 0$, the above statistic will converge to the well-known rescaled variance test under the null hypothesis of short memory. In other words, under the null hypothesis that the underlying (hidden) factor is diversifiable with equally weighted sequences, the statistic in Equation (2.2.11) will become the the cross-sectional version of rescaled variance test at time t. This, in turn, may lead to similar empirical statistics applied to verify the result of Proposition 2.2.1 that the nondiversifiability of the (hidden) factor will exhibit in their cross-sectional long dependence. However, these test statistics (in using cross-sectional exponents α) are based on the Assumption 2.A.7 and only related to the arguments in Equation (2.2.1) of equally weighted portfolio(s).

Independently, Bailey et al. (2012) also develop a similar analysis with an exponent for cross-sectional dependence under the presumed factor structure. Their analysis however, is based on the partial sums and canonical factor analysis. Furthermore, the factor loadings (although allowing to be random) are separated into two categories: ones that are bounded away from zero, and the other transitory ones that may decay to zero exponentially. Specifically, the cross-sectional exponent in their definition depends on the specific decaying rate of the cross-sectional factor loadings. The test statistics for these exponents are also established mainly to verify the cross-sectional dependence, not to identify the essential nondiversifiable factor(s) for empirical asset pricing models—although the CAPM is under discussion in their work. However, as stated earlier, these assumptions of cross-sectional exponents necessarily lead to the extent of cross-sectional dependence. The range of these exponents (similar to the above example) can be shown merely as sufficient conditions that lead to the cross-sectional long dependence.

In other words, there are many possible specifications that may cause the cross-sectional long dependence. These assumptions of cross-sectional exponents although explicit, do not cover all possible cases for efficient portfolios in identifying the nondiversifiable (systematic) factor(s). In particular, a supremum for the growth rates among all possible factors is applied for the verification of cross-sectional dependence in their work. Although useful, the method may mislead the analysis to an overfitted model with factor(s) or proxies that do not relate to the nondiversifiability since the arguments are based on the supremum of these cross-sectional exponents for factor loadings of the equally weighted portfolio.

The framework of the model search in the current context hence, does not include explicit assumptions for the decaying rates of factor loadings when equally-weighted portfolios are applied. Instead, following from Theorem 2.1.2, the cross-sectional long dependence is a *necessary* condition for the existence of nondiversifiable factor—whether there are specific assumptions for the cross-sectional exponents (for the growth rate of factor loadings) or not. The cross-sectional long dependence becomes sufficient condition for nondiversifiability only when certain specific memory conditions are given. In that case, assumptions for specific memory conditions will have to sacrifice the more general definition of nondiversifiability in factors.

Namely, the cross-sectional dependence (whether subject to specific assumptions for convergence rates) can not be treated as a sufficient condition in general, and applied for model specification tests. More specifically, the essence is to verify the persistent cross-sectional long dependence in which the functionals of fractional Brownian motion (with Hurst exponent) are only the approximations for the asymptotic representations in distribution. In other words, verification of cross-sectional long dependence using the replicates or extensions of long-memory tests in time series modeling is only a diagnostic tool without explicitly considering the decaying rates or exponents for the factor loadings (or else) across all asset returns. Therefore, verification of cross-sectional long dependence is the main interest of study for the existence of nondiversifiable (hidden) factor—whether it is based on the cross-sectional exponent, Hurst exponent, Allan variance, or others. Furthermore, more advanced methodology or definitions for long dependence can also be extended or applied here.

In addition, let $\lim\limits_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \sigma_{H,t}^2 \longrightarrow \overline{\sigma}_H^2 > 0$ for all N, and Assumption 2.A.7 holds (whether a cross-sectional exponent α is assumed or not)

$$\left[\frac{1}{T}\sum_{t=1}^{\mathrm{T}}\left(\sum_{m=1}^{\mathrm{N}}\left(\frac{1}{N}\sum_{i=1}^{m}(\epsilon_{it}-\overline{\epsilon_t})\right)^2\right)\right] \xrightarrow{d} \overline{\sigma}_H^2 \int_0^1 (\overline{B}_H(z))^2\, dz, \tag{2.2.12}$$

as $N \longrightarrow \infty$, $T \longrightarrow \infty$, where $\overline{B}_H(z) = B_H(z) - zB_H(1)$ is a fractional Brownian bridge, and $H = \frac{1}{2}$, if under the null hypothesis that the hidden factor is diversifiable. In other words, the test statistics in Equations (2.2.9), (2.2.10) can still be applied if the condition $\lim\limits_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \sigma_{H,t}^2 \longrightarrow \overline{\sigma}_H^2 > 0$

holds, and the consistent estimate for $\sigma^2_{H,t}$ is available. Alternatively, the analysis can be extended with a stronger condition. If the hidden factor is diversifiable, the subseries for all equally weighted idiosyncratic risk such as $\frac{1}{m}\sum_{i=1}^{m}\epsilon_{it}$will converge to zero in L^2-norm as $m \rightarrow \infty$.

Given suitable growth rates for both N and m, an alternative statistic can be devised. Assuming the weak law of large numbers such that $\frac{1}{T}\sum_{t=1}^{T}\left(\frac{1}{m}\sum_{i=1}^{m}\epsilon_{it}\right)^2 \xrightarrow{p} E[\frac{1}{m}\sum_{i=1}^{m}\epsilon_{it}]^2$ for all $1 \leq m \leq N$, as $T \longrightarrow \infty$, then for any N, as $T \longrightarrow \infty$,

$$\frac{1}{T}\sum_{t=1}^{T}\left(\sum_{j=1}^{[N/m]}\left(\frac{1}{m}\sum_{i=1}^{m}\epsilon_{it}\right)^2\right) \xrightarrow{p} \sum_{j=1}^{[N/m]} E\left(\frac{1}{m}\sum_{i=1}^{m}\epsilon_{it}\right)^2. \tag{2.2.13}$$

Notice that the above results are based on the assumption of the intensity of diversifiability α and the equally weighted portfolios. The purpose is to link the cross-sectional long dependence with the intensity of diversifiability under equally weighted schemes. The argument shows that the less diversifiable the hidden factor is (such that $\alpha > \frac{1}{2}$ and increasing), the stronger the cross-sectional long dependence for the (presumed) idiosyncratic risks. And hence, the cumulative statistics will converge to a functional of fractional Brownian bridge.

Theorem 2.2.2, instead, shows another form of panel CUSUM statistic in using residuals for empirical asset pricing models to verify (hidden) nondiversifiable factor (or component) in projection errors under the null hypothesis that the (hidden) factor is diversifiable. In addition, the tests apply the residuals from the presumed model of normal (or expected) returns and a sequential variable selection approach is suggested for the model search in empirical asset pricing models.

Theorem 2.2.2: Under the null hypothesis that the projection errors $\{\epsilon_{it}\}_{i=1,2,\ldots}$do not contain a hidden non-diversifiable factor so that $H = \frac{1}{2}$ in Assumption, 2.A.7, let $\left\{\hat{\epsilon}_{it}\right\}_{i=1,2,\ldots,N,}$ be the residuals of fitted regressions for asset returns at time t and $\{\epsilon_{it}\}_{i=1,2,\ldots,N,}$ follow the invariance principle of short-memory stochastic processes. Also let $\hat{s}^2_{N,t} = \frac{1}{N}\sum_{i=1}^{N}\hat{\epsilon}^2_{it} +$

$2\sum_{j=1}^{N-1}\theta(\frac{j}{q})\hat{\gamma}_t(j)$ be the heteroskedasticity and autocorrelation consistent (HAC) estimate for the cross-sectional asymptotic variance σ_t^2 at time t, $t = 1,2,\cdots,T$, $T \to \infty$,where $\hat{\gamma}_t(j) = \frac{1}{N}\sum_{i=1}^{N-j}\hat{\epsilon}_{it}\hat{\epsilon}_{i+j,t}$, and $\theta(\cdot)$ is the kernel function with bandwidth q, $q \to \infty$, $\frac{q}{N} \to 0$, and assume that $\hat{s}_{N,t} = \sigma_t + o_p(1) > 0$, as $N \to \infty$, $\overline{\hat{\epsilon}_t} = \frac{1}{N}\sum_{i=1}^{N}\hat{\epsilon}_{it}$, the panel CUSUM statistic $\hat{\Psi}^{\epsilon}_{n,T}$ such as

$$\hat{\Psi}^{\epsilon}_{N,T} = \frac{1}{T}\sum_{t=1}^{T}\left[N^{-1}\sum_{m=1}^{N}\left(\frac{1}{\sqrt{N}\hat{s}_{N,t}}|\sum_{i=1}^{m}(\hat{\epsilon}_{it}-\overline{\hat{\epsilon}_t})|\right)\right], \tag{2.2.14}$$

will show that

$$\hat{\Psi}^{\epsilon}_{N,T} \xrightarrow{d} \int_0^1 |\overline{\mathcal{B}}(z)|dz, \tag{2.2.15}$$

provided that $\left(\frac{1}{\sqrt{N}}\sum_{i=1}^{m}\epsilon_{it}\right) \xrightarrow{d} \sigma_t\mathcal{B}(z)$, $0 < z \leq 1$, $1 \leq m \leq N$, as $N \to \infty$, where $t = 1,2,..,T$, $\overline{\mathcal{B}}(z)$=$\mathcal{B}(z) - z\mathcal{B}(\infty)$ is a Brownian bridge and $\mathcal{B}(z)$, is a standard Brownian motion, where $z \in [0,1]$- provided that $\underline{\hat{\beta}_i} \xrightarrow{p} \underline{\hat{\beta}_i}$ where $\underline{\beta_i} = (\beta_{i1},\beta_{i2},\cdots,\beta_{ik})'$ and $\underline{\hat{\beta}_i} = \left(\hat{\beta}_{i1},\hat{\beta}_{i2},\cdots,\hat{\beta}_{ik}\right)'$ respectively, for all $i = 1,2,\cdots,N$.

The statistic $\hat{\Psi}^{\epsilon}_{N,T}$ is consistent in the sense that $\hat{\Psi}^{\epsilon}_{N,T} \to \infty$ as $N \to \infty$, $T \to \infty$, when there is a hidden nondiversifiable factor in $\{\epsilon_{it}\}_{i=1,2,...}$ such that $N^{-H}\sum_{i=1}^{[Nz]}\epsilon_{it} \xrightarrow{d} \gamma\,\mathcal{B}_H(z)$, $z \in [0,1]$, $H > \frac{1}{2}$, σ^2_{Ht} is the cross-sectional asymptotic variance of $\{\epsilon_{it}\}_{i=1,2,...}$where $\{\epsilon_{it}\}_{i=1,2,...}$ are of cross-sectional long memory and $\mathcal{B}_H(z)$ is a fractional Brownian motion.

The distribution for $\int_0^1 |\overline{\mathcal{B}}(z)|dz|$ is shown in Johnson and Killeen (1983) as

$$Pr(\int_0^1 |\overline{\mathcal{B}}(z)|dz \leq b) = (\frac{\pi}{2})^{\frac{1}{2}}\sum_{s=1}^{\infty}\left(\frac{1}{\delta_s^{\frac{3}{2}}}\right)\zeta\left(\frac{b}{\delta_s^{\frac{3}{2}}}\right), \tag{2.2.16}$$

$$\zeta(x) = \frac{3^{\frac{3}{2}} e^{\frac{-2}{27x^2}}}{x^{\frac{1}{3}}} A_i((3x)^{-\frac{4}{3}}, \tag{2.2.17}$$

$$\delta_s = \frac{-a_s'}{2^{\frac{1}{3}}}, \tag{2.2.18}$$

where A_i is the Airy function, a_s' is the s-th zero of A_i'.

Proof of Theorem 2.2.2: Applying the results of Proposition 2.2.1, it suffices to see that for $T \to \infty$,

$$\Psi_{N,T}^{\epsilon} = \frac{1}{T}\sum_{t=1}^{T}\left[N^{-1}\sum_{m=1}^{N}\left(\frac{1}{\sqrt{N}\sigma_t}|\sum_{i=1}^{m}(\epsilon_{it}-\overline{\epsilon_t})|\right)\right] \xrightarrow{d} \int_o^1 |\overline{\mathcal{B}}(z)|dz.$$

Hence, the proof requires to show that (1) $\hat{\Psi}_{N,T}^{\epsilon} - \Psi_{N,T}^{\epsilon} = o_p(1)$, and (2) $\widetilde{s}_{Nt}^2$ is also a consistent estimator for σ_t^2. The first part is easy to show since the factor structure is orthogonal to the error terms, the consistency of estimates for factor loadings is provided. Hence, the functionals of residuals will converge to those of error terms under law of large numbers. For the second part, the verification requires some works with matrix algebra.

Denote the fitted asset pricing model (*at any time t* where the time index is suppressed, for simplicity) as

$$R_n = B\widetilde{f} + \underline{\epsilon},$$

where $R_N = (r_1, r_2, \cdots r_N)'$is a N-by-one vector of excess returns in N assets, B is a N-by-p matrix of factor loadings such that $\{\beta_{ij}\}_{i=1,\cdots,N}^{j=1,\cdots,p}$ represents the elements on i-th row and j-th column of B matrix, $\widetilde{f} = (1, f_1, f_2, \cdots f_p)' = (1, f')'$is the $(p+1)$-by-one vector of known factors in the presumed model, $\underline{\epsilon} = (\epsilon_1, \epsilon_2, \cdots, \epsilon_N)'$is a N-by-one vector of idiosyncratic risks, $\hat{\underline{\epsilon}} = (\hat{\epsilon}_1, \hat{\epsilon}_2, \cdots, \hat{\epsilon}_N)'$is a N-by-one vector of forecast errors conditional on the factors included in the model. Let $\overline{\epsilon}_N = \frac{1}{N}\sum_{i=1}^{N}\epsilon_i$, and $\ddot{\epsilon}_N = \frac{1}{N}\sum_{i=1}^{N}\hat{\epsilon}_i$, respectively. Also let $e_m = (1, 1, \cdots, 0, 0, \cdots, 0)'$ be a N-by-one vector of m elements of one's and $N - m$ elements of zero's, $m = 1, 2, \cdots, N$. Then, let $\hat{B}$ be the matrix of consistently estimated factor loadings, it can be shown with simple linear algebra that

$$\underline{\hat{\epsilon}} - e_N \ddot{\epsilon}_N = \left(\underline{\epsilon} - e_N \overline{\epsilon}_N\right) - \left(I_N - \frac{1}{N} e_N e_N'\right)\left(\hat{B} - B\right)\widetilde{f},$$

and the cumulative sums as $m = 1, 2, \cdots, N$,

$$\begin{aligned}\sum_{i=1}^{m}\left(\hat{\epsilon}_i - \ddot{\epsilon}_N\right) &= e_m'\left[\left(\underline{\epsilon} - e_N \overline{\epsilon}_N\right) - M_N\left(\hat{B} - B\right)\widetilde{f}\right] \\ &= \sum_{i=1}^{m}\left(\epsilon_i - \overline{\epsilon}_N\right) - e_m' M_N\left(\hat{B} - B\right)\widetilde{f},\end{aligned}$$

where $M_N = I_N - \frac{1}{N} e_N e_N'$ and I_N is N-by-N identity matrix. Since $\hat{B} - B \to o_p(1)$ when estimated consistently, $M_N \to I_N$ as $N \to \infty$, it is easy to see that $\sum_{i=1}^{m}(\epsilon_i - \bar{\epsilon}_N) = \sum_{i=1}^{m}\left(\hat{\epsilon}_i - \ddot{\epsilon}_N\right) + o_p(1)$. This, in turns shows that $\hat{\Psi}^{\epsilon}_{N,T} - \Psi^{\epsilon}_{N,T} = o_p(1)$. Given that $\gamma(j) = \frac{1}{N}\sum_{i=1}^{N-j}(\epsilon_i - \overline{\epsilon}_N)\left(\epsilon_{i+j} - \overline{\epsilon}_N\right)$, let $A_j = \left[I_{N-j}, O_{(N-j),j}\right]$, $B_j = \left[O_{(N-j),j}, I_{N-j}\right]$, where $O_{(N-j),j}$ is a $(N-j)$-by-j matrix of zeros. We have

$$\begin{aligned}\ddot{\gamma}(j) &= \frac{1}{N}\left[\left(A_j\left(\underline{\hat{\epsilon}} - e_N \ddot{\epsilon}_N\right)\right)'\left(B_j\left(\underline{\hat{\epsilon}} - e_N \ddot{\epsilon}_N\right)\right)\right] \\ &= \frac{1}{N}\left[\left(\underline{\hat{\epsilon}} - e_N \ddot{\epsilon}_N\right)' A_j' B_j\left(\underline{\hat{\epsilon}} - e_N \ddot{\epsilon}_N\right)\right] \\ &= \frac{1}{N}\left[\left(\underline{\hat{\epsilon}} - e_N \ddot{\epsilon}_N\right)' \ddot{N}(j)\left(\underline{\hat{\epsilon}} - e_N \ddot{\epsilon}_N\right)\right],\end{aligned}$$

where $\ddot{N}(j)$ is a N-by-N matrix. Hence, as $N \to \infty$,

$$\begin{aligned}\ddot{\gamma}(j) &= \frac{1}{N}\left[\left(\left(\underline{\epsilon} - e_N \overline{\epsilon}_N\right) - M_N\left(\hat{B} - B\right)\widetilde{f}\right)' \ddot{N}(j)\left(\left(\underline{\epsilon} - e_N \overline{\epsilon}_N\right) - M_N\left(\hat{B} - B\right)\widetilde{f}\right)\right] \\ &= \hat{\gamma}(j) - \frac{2}{N}\left(\underline{\hat{\epsilon}} - e_N \ddot{\epsilon}_N\right)' \ddot{N}(j) M_N\left(\hat{B} - B\right)\widetilde{f} \\ &+ \frac{1}{N}\left[\widetilde{f}'\left(\hat{B} - B\right)' M_N N(j) M_N\left(\hat{B} - B\right)\widetilde{f}\right] \\ &= \hat{\gamma}(j) - o_p(1) + \frac{1}{N} tr\left[\left(\hat{B} - B\right)' M_N \ddot{N}(j) M_N\left(\hat{B} - B\right)\widetilde{f}\widetilde{f}'\right] \\ &= \hat{\gamma}(j) - o_p(1) + \frac{1}{N} Vec\left[\left(\hat{B} - B\right)' M_N \ddot{N}(j) M_N\left(\hat{B} - B\right)\right]' \mathrm{Vec}\left(\widetilde{f}\widetilde{f}'\right)\end{aligned}$$

$$= \hat{\gamma}(j) - o_p(1) + Vec\left[\left(\hat{B}-B\right)' M_N \ddot{N}(j) M_N \left(\hat{B}-B\right)\right]' \text{Vec}\left(\frac{1}{N}\widetilde{f}\widetilde{f}'\right)$$
$$= \hat{\gamma}(j) - o_p(1),$$

provided that $\hat{B} - B \rightarrow o_p(1)$ and $\frac{1}{N}\widetilde{f}\widetilde{f}' \xrightarrow{p} \Sigma_f$, where Σ_f is a positive-definite matrix. Thus, given the consistency of $\ddot{\gamma}(j)$ and $\hat{\gamma}(j)$ in estimating the cross-sectional covariances under the null hypothesis of cross-sectional short memory in $\{\epsilon_{it}\}_{i=1,2,\cdots}$, the estimate $\widetilde{s}^2_{Nt}$ for long-run cross-sectional variance σ_t^2 is also consistent since $\widetilde{s}^2_{Nt} - \widehat{s}^2_{Nt} = o_p(1)$. For consistency of the test, consider $\Psi^{\epsilon}_{N,T} = \frac{1}{T}\sum_{t=1}^{T}\left[N^{-1}\sum_{m=1}^{N}\left(\frac{1}{N^H}|\sum_{i=1}^{m}(\epsilon_{it} - \overline{\epsilon_t})|\right)\right] \xrightarrow{d} \ddot{\sigma}_H \int_0^1 |\mathcal{B}_{\mathrm{H}}(z)|dz$, where $H > \frac{1}{2}$ as $N \rightarrow \infty$, $T \rightarrow \infty$, under the alternative hypothesis. Notice that $\hat{\Psi}^{\epsilon}_{N,T} - \Psi^{\epsilon}_{N,T} = o_p(1)$, and assume that $\frac{1}{T}\sum_{t=1}^{T}\sigma_{Ht} + o_p(1) \xrightarrow{p} \ddot{\sigma}_H$, where $H > \frac{1}{2}$, then as $N \rightarrow \infty$, $T \rightarrow \infty$,

$$\hat{\Psi}^{\epsilon}_{N,T} \approx \Psi^{\epsilon}_{N,T} N^{H-\frac{1}{2}} + o_p(1) \longrightarrow \infty.$$

In particular, Johnson and Killeen (1983) also provides a table of probability distribution with respect to different critical value for the distribution of L^1–norm of Brownian bridge. Thus, we may check to see if these statistics are significant or not. Another interesting feature of the statistic is that it also the convergence toward the area underneath the reflected Brownian bridge.

Following the reasoning of Theorem 2.1.1 and Corollary 2.1.2, we have an additional condition to ensure that the selected proxies are of nondiversifiable factor loadings. An alternative method suggested in this paper is to perform a bottom-up sequential model search by expanding the dimension of model for nondiversifiability of selected proxies. The notion is not only to obtain the most parsimonious model of relevant variables. Instead, we need to identify the model that most likely represents the non-diversifiable components of excess returns. More specifically, it is necessary to pick up the variables that ensure the nondiversifiability of their factor loadings to specify the nondiversifiable or "systematic" components in excess returns. Thus, a sequential model search is devised here to make certain the selected variables are of such property. Moreover, given that model search with all different possible permutations of variables may be too computationally intensive, a sequential approach may reduce the number of search needed.

For instance, if we perform the model search with some model selection criteria, this will require model search for all 2^p possible models among p variables. Instead, the following bottom-up sequential model search approach only needs to check for $\frac{p(p+1)}{2}$ models.[7]

More specifically, even though the above $\hat{\Psi}_{N,T}$ is devised for projection errors, we may also introduce the $\hat{\Psi}_{N,T}$ statistics in Theorem 2.2.2 for the fitted values to establish the sequential model search as $\hat{\Psi}^{f(j)}_{N,T} = \frac{1}{T}\sum_{t=1}^{T}\left[N^{-1}\sum_{m=1}^{N}\left(\frac{1}{\sqrt{N}\hat{s}^{(j)}_{N,T}}|\sum_{i=1}^{m}(\hat{h}^{(j)}_{it} - \overline{\hat{h}^{(kj)}_{t}})|\right)\right]$, where $\hat{h}^{(j)}_{it} = \hat{\beta}^{(j)}_{i} f^{(j)}_{t} = \beta^{(j)}_{i} f^{(j)}_{t} + o_p(1)$, is the fitted value of the j-th proxy in the empirical model, $1 \leq j \leq k$. The intuition for sequential model selection tests is simple. Following from Corollary 2.1.2, if the k-th proxy is indeed needed and is with nondiversifiable factor loadings, the $\hat{\Psi}^{f(j)}_{N,T}$ for the fitted values of this proxy will be statistical significant and increasing with N and T due to cross-sectional long dependence.

In particular, the higher the $\hat{\Psi}^{f(j)}_{N,T}$ is, the more essential the variable is for describing the nondiversifiable component of excess returns. Many model selection tests or criteria are to obtain the optimal model(s) of relevant variables. Unfortunately, locating relevant variables for the model is not entirely identical to searching for a model with nondiversifiable factors (or proxies). Locating some models with statistically significant diversifiable variables or proxies may not necessarily improve the model specification since the findings may not necessarily identify the "systematic" "nondiversifiable" components for excess returns within finite samples. Thus, we may devise a sequential search for the model of excess returns using the above model selection test.

Corollary 2.2.3: Given Theorems 2.1.2 and 2.2.2, let the test statistics of a candidate variable for the j-th proxy (where $1 \leq$

$j \le k$) $\hat{\Psi}^{f(j)}_{N,T}$ be defined as

$$\hat{\Psi}^{f(j)}_{N,T} = \frac{1}{T}\sum_{t=1}^{T}\left[N^{-1}\sum_{m=1}^{N}\left(\frac{1}{\sqrt{N}\hat{s}^{(j)}_{N,T}}|\sum_{i=1}^{m}(\hat{h}^{(j)}_{it} - \overline{\hat{h}^{(j)}_{t}})|\right)\right], \tag{2.2.19}$$

where $h^{(j)}_{it} = \beta^{(j)}_i f^{(j)}_t$, and $\hat{h}^{(j)}_{it} = \hat{\beta}^{(j)}_i f^{(j)}_t = h^{(j)}_{it} + o_p(1)$ is the fitted value of the j-th proxy in the empirical asset pricing model. Then if this candidate variable is of diversifiable loadings, it follows that

$$\hat{\Psi}^{f(j)}_{N,T} \xrightarrow{d} \int_0^1 |\overline{\mathcal{B}}(z)|dz, \tag{2.2.20}$$

provided that $\left(\frac{1}{\sqrt{N}}\sum_{i=1}^{m} h^{(j)}_{it}\right) \xrightarrow{d} \sigma_t \mathcal{B}(z),\ 0 < z \le 1,\ 1 \le m \le N$, as $N \to \infty$, where $\overline{\mathcal{B}}(z)$=$\mathcal{B}(z) - z\mathcal{B}(\infty)$ is a Brownian bridge and $\mathcal{B}(z)$, is a standard Brownian motion defined on the interval $[0,1]$. Let $\hat{s}^{(j)2}_{N,T} = \frac{1}{N}\sum_{i=1}^{N}\hat{h}^{(j)2}_{it} + 2\sum_{\tau=1}^{N-1}\theta(\frac{\tau}{q})\hat{\gamma}^{(j)}_t(\tau)$ be the heteroskedasticity and autocorrelation consistent (HAC) estimate for the cross-sectional asymptotic variance $\sigma^{(j)2}_t$ of $\left\{h^{(j)}_{it}\right\}_{i=1,2,...}$ at time t, $t = 1,2,\cdots,T$, where $\hat{\gamma}^{(j)}_t(\tau) = \frac{1}{N}\sum_{i=1}^{N-j}\hat{h}^{(j)}_{it}\hat{h}^{(j)}_{i+\tau,t}$, and $\theta(\,\cdot\,)$ is the kernel function with bandwidth q, $q \to \infty$, $\frac{q}{N} \to 0$, and assume that $\frac{1}{T}\sum_{t=1}^{T}\hat{s}^{(j)}_{N,T} = \frac{1}{T}\sum_{t=1}^{T}\sigma^{(j)}_t + o_p(1) \xrightarrow{p} \ddot{\sigma} > 0$.

The statistic $\hat{\Psi}^{f(j)}_{N,T}$ is consistent that $\hat{\Psi}^{f(j)}_{N,T} \to \infty$ as $N \to \infty$, $T \to \infty$ if $\left\{h^{(j)}_{it}\right\}_{i=1,2,...}$ follow the alternative hypothesis such that $N^{-H}\sum_{i=1}^{[Nz]} h^{(j)}_{it} \xrightarrow{d} \gamma\,\mathcal{B}_H(z),\ z \in [0,1],\ H > \frac{1}{2}$, σ^2_{Ht} is the long-run variance of $\left\{h^{(j)}_{it}\right\}_{i=1,2,...}$ when $\left\{h^{(j)}_{it}\right\}_{i=1,2,...}$ are of cross-sectional long memory and $\mathcal{B}_H(z)$ is a fractional Brownian motion.

Proof of Corollary 2.2.3: Applying the results of Theorem 2.2.2, and set $\hat{e}_{it} = \hat{h}^{(j)}_{it}$ for each k-th proxy or identified variable, $1 \le j \le k$, the claims will follow.

For the arguments of (bottom-up) sequential model search, Assumption 2.A.8 implies that (after the k-th stage of model search, $k \geq 2$) the modeler has (at least) already included some $k-1$ proxies that may partially approximate the nondiversifiable components of excess returns. That is, at the k-th stage of model search, some $k-1$ relevant proxies with nondiversifiable factor loadings $\ddot{\beta}_j$ should have been identified. Applications of the new model selection criteria are to verify if further expansion of the model is needed. Otherwise, if $\ddot{\beta}_j$ is diversifiable, it will imply the modeler has simply included some irrelevant (diversifiable) proxies at first. For instance, suppose that CAPM is valid for the first stage of model search, the best candidate proxy will be some rates of returns of market indices. The factor loadings of all excess returns projected on this proxy should be non-diversifiable if these market indices are closely associated with market portfolio.

Assumption 2.A.8: Let the loadings of the already-selected $k-1$ proxies $\left\{f_t^{(j)}\right\}_{j=1,\ldots,k}$ be non-diversifiable at the k-th stage of model search that $\inf_{\underline{\omega}\in W^\star} G(\underline{\omega},\ddot{\beta}_j) \neq 0$, $\ddot{\beta}_j = \left(\beta_1^j,\beta_2^j,\cdots,\beta_N^j,\cdots\right)'$, $j = 1,2,\ldots,k-1$ for all well-diversified portfolios $\underline{\omega}$ in $W^\star$, where $k \geq 2$.

The detailed algorithm is stated as follows;

- (1). Select one proxy variable f_{jt}, $j = 1,2,\cdots,p$ initially from $\left\{f_{jt}\right\}_{j=1,2,\ldots,p} \in \underline{f_t}$, where $\underline{f_t}$ is the set of all presumed proxy variables for factors at time t. . Run the univariate regressions of $\{r_{it}\}_{i=1,2,\ldots,N,}$ on the f_{jt} for each $i = 1,2,\ldots,N$ to obtain the factor loadings $\hat{\beta}_{ij}$. Select those proxies where their $\hat{\Psi}_{N,T}^{f(k)}$'s are statistically significant. Choose the proxy variable among these significant proxies with maximum $\hat{\Psi}_{N,T}^{f(k)}$ and denote it as $f_t^{(1)}$ for the first stage of model search as $k = 1$.
- (2). Pursue further model search at k-th step, where $k = 2,\cdots,s$, $s \leq p$. Let $\hat{\epsilon}_{it}^{(k-1)} = r_{it} - \hat{\alpha}_i - \hat{\beta}_{i1} f_t^{(1)} - \cdots - \hat{\beta}_{i,(k)} f_t^{(k-1)}$, which represents the residual after $(k-1)$-th steps of search. Select and obtain orthogonalized kth-stage

regressor $\hat{\lambda}_t^{(k)}$, where $\hat{\lambda}_t^{(k)}$ is the residual after regressing each candidate variable f_t^* in $\hat{\Pi}_t = \underline{f_t} / \left\{ f_t^{(1)}, f_t^{(2)}, \ldots, f_t^{(k-1)} \right\}$ on the already selected proxies $\left\{ f_t^{(1)}, f_t^{(2)}, \cdots, f_t^{(k-1)} \right\}$, and $\hat{\Pi}_t = \underline{f_t} / \left\{ f_t^{(1)}, f_t^{(2)}, \ldots, f_t^{(k-1)} \right\}$ represents the subset of $\underline{f_t}$ that excludes the already chosen proxies $\left\{ f_t^{(1)}, f_t^{(2)}, \cdots, f_t^{(k-1)} \right\}$.[8] Perform univariate regressions of $\left\{ \hat{\epsilon}_{it}^{(k-1)} \right\}_{i=1,2,\ldots}$ on $\hat{\lambda}_t^{(k)}$ to get the factor loading $\left\{ \hat{\beta}_{i,(k)} \right\}_{i=1,2,\ldots,}$. Construct the new residual as $\hat{\epsilon}_{it}^{(k)} = r_{it} - \hat{\alpha} - \hat{\beta}_{i1} f_t^{(1)} - \cdots - \hat{\beta}_{i,(k-1)} f_t^{(k-1)} - \hat{\beta}_{i,(k)} \hat{\lambda}_t^{(k)}$ for the k-th step, and choose the proxy $\hat{\lambda}_t^{(k)}$ which maximizes $\hat{\Psi}_{n,T}^{f(k)}$ and denote it as $f_t^{(k)}$. Continue step 2.

- (3). If for some sufficiently small $\delta \in R$, $\delta > 0$, $|\hat{\Psi}_{n,T}^{f(k)} - \hat{\Psi}_{n,T}^{f(k-1)}| < \delta$, check to see if $\hat{\Psi}_{n,T}^{\epsilon}$ is significant in asymptotic distribution of $\int_0^1 |\overline{\mathcal{B}}(z)| dz$. If not, the search stops.

In other words, the procedures start with variables search on the proxy variable(s) by fitting the returns with some linear regressions that contain a constant term. Then, perform model search by expanding the model with more proxies using sequential model selection tests. The search will stop when there is no indication of any possible hidden nondiversifiable factor(s). Therefore, even if additional variable(s) that are diversifiable and may still contribute to the predictability of the model, the search will stop. Adding additional diversifiable explanatory variables may perhaps improve the temporal predictability of empirical asset pricing models tentatively. Yet, this inclusion will not necessarily improve the event studies in using the abnormal returns.[9] Perhaps, an easier way to say is "...let the abnormal returns remain as abnormal, where the firms specific information may still stay with them."

Chapter 3

Cumulative Abnormal Returns or Structural Change Tests?

Introduction

In this chapter, it is shown that if the impact of event(s) is considered as permanent, the cumulative abnormal return statistics in event studies coincides with the CUSUM statistics in the tests for parameter changes of regressions such as market models. Namely, the applications for the tests on abnormal returns are closely related with the model specification of normal returns, especially with the regression models assumed for the normal (expected) returns. If the statistical approach is reterospective (where the studies of interest are to identify the possible (permanent) change in parameters within the given history of stock returns), and if the presumed initial date for event window is the correct time period where parameter changes, the hypotheses testings of the conventional CARs and CUSUM statistics are almost identical except for the asymptotic distributions applied. The CARs tests apply the (asymptotic) normality, while CUSUM tests are based on Brownian motion or Brownian bridge.

For instance, the CARs in event studies usually assume the length of estimation period to increase asymptotically for validity of CARs statistics. The CUSUM reterospective tests assume the entire time period of interest (i.e., estimation period plus event window) to expand asymptotically, while the monitoring tests assume the training period can be of sufficient length. This compatibility is even vivid if the linear regression models are applied to construct the normal and abnormal returns. In the following, arguments on event window and

CARs are provided. Extension for these studies using the sequential approach for monitoring is also given. However, under circumstances where similarity of CARs and CUSUM tests established, it is questionable to verify that these two schools of statistical tests really depicts the essential answer for event studies. As well-known in corporate finance, the financial time series models and their parameters change rapidly over time—whether there is any significant event or not. These parameters can undergo abrupt changes or jumps over time even though there is no specific knowledge that leads to hypotheses of some corporate events.

Therefore, if one is interested in applying CARs to investigate the impacts of corporate events, caution should be applied since verification of these parameter changes is only to concur a possible impact or drift. A more thorough study on the abnormal returns should be added to consider the market adjustments and their durations. In particular, if the impacts from events are only temporary, the hypotheses of interest will then be identical to the epidemic changes in parameters. Specifically, the parameter changes occur only in a so-called event window. After the event window, the parameters will return to their original levels if the event window is correctly identified. For most event studies in corporate finance (since firm-specific information is not systematic), these settings of epidemic changes seem more closely fitted to the data when impacts from corporate finance events are mostly temporary. Yet, even with the multiple change points identified, it is still difficult to link these changes with the events of interest directly. Hence, given that the fluctuaions of security prices in capital market are so intensive, conventional CARs tests (and the CUSUM equivalents) are not entirely informative for event studies of corporate finance. Alternative methodology is needed.

3.1 Event Window and Sampling Period

One of the difficulties in event studies of corporate finance is the determination of event window and pre-event and post-event periods. For various event studies with daily returns, the event window usually covers a few days before the announcement (or event) date and some days after the announcement (or event) date if the date of interest is known. Although determining the pre-and post- mergers (or acquisitions) performance of either target or acquiring firms is essential to corporate finance, arbitrariness in choices of event window may dilue the soundness of the empirical findings.

For instance, various event studies when using daily returns select these days or windows subjectively or apply some pretest assumptions for the selection. As the conventional approach, the estimation period (or window) is set as the time period prior to this event window to estimate the coefficients of the presumed model for normal (or expected) returns. The prediction errors within the event window are collected for the tests on hypotheses of interest. The asymptotic arguments for the properties of statistical methods are based on the length of this estimation period. As the length expands, the asymptotic distributions for the estimates and the abnormal returns are then obtained.

The difficulty is that most of these methods assume either constant mean or constant coefficients on the presumed specification of normal (or expected) returns across event window under the null hypothesis. One reason for that is, under the null hypothesis where the firm-specific information such as mergers and acquisitions, the systematic components should not differ much even within the event window. The selection of event window that covers some days (or months) prior to the announcement (or event) date is to accommodate the possible influence from the information/announcement to the market. However, the choices on how extensive these days (or months) prior to and after the event date are either arbitrarily determined or dependent on selected samples.

For instance, the earlier study of Asquith et al. (1983) uses a period of –20 days to +20 days to cover the entire period of interest. Brown et al. (1985) show that the "beta's" are subject to structural change across the event periods where the number of changes depends on the determinations of various pre- and post- event periods. Brown and Warner (1989) apply monthly stock returns and cover –89 months to +10 months for the entire sample period (where month zero stands for the event month). In addition, Davidson et al. (1989) apply data for returns of failed mergers from –90 days before event to +250 days after the event day.

In particular, once these event windows are determined, the same length of event window (of various event dates) may usally be applied to selected firms for the similar events of interest. Given that capital market may adjust to new information of different firms in various ways or paces, it is hard to see why the length of event windows should be uniformly determined a priori even though the event dates may be different across all selected firms. Knowingly, even if these events are similar across the selected firms, they may not happen in the same time periods (or with the same length of event window). If, on the other hand, the event windows are not of the same dates (or length) across different firms, it is cumbersome to form some convenient statistics such as the average of cumulative abnormal returns using each firm's abnormal returns since the lengths of event windows or event dates all vary across selected firms.

Furthermore, across different time periods, the determination of pre-event period, event window, and post-event period should ideally be based on the adjustment speed of the capital market itself. Hence, an alternative for the selection of estimation period or else (or even, to completely drop the event window) is to let the data speak for themselves. That is, a recursive adjustment and its mechanism can be devised in tracking the normal (or expected) returns where windows for these adjustments are adaptive to the data (and time periods) so that the consequence of subjectivity in event-window selection can be minimized, especially when the newly available firm-specific information may leak out to influence the market expectations. More extensively, given the identified (market-wise) systematic determinants, the recursive estimation or adaptive tracking of the parameters in normal (or expected) returns may also allow other extraneous noises or information around the event (or announcement) date under some regularity conditions. Nevertheless, given the sliding windows (which may be different for each firm's returns) for tracking, the conventional methods in using cumulative abnormal returns that start from a specific date before the event date will no longer be applicable. An alternative method in verifying the impact of events must be devised. These methods then, are considered in Chapters 4 and 5 in Part 2 of this book.

While being arbitrary in setting the event windows, most conventional event studies ignore the back-testing for the post-event period. In fact, the quintessence of event studies is not only to identify the immediate impact, but the follow-ups from the events or announcements. Namely, how speedy the market or firms may adjust afterward may be as equally critical as the impacts of events. Unfortunately, conventional event studies usally stop at the findings on the impacts of events. As usual, the identification of impacts depends heavily on the known or presumed date when the impacts may start. More surprisingly enough, this setting is almost identical to the tests for structural change with some a priori information when the change date (or time) actually occurs within the presumed event window.

Alternatively, if the impacts from the events are considered temporary (i.e., the impacts from events will diminish to nil after the event windows per se), the hypothesis testing in this setting is similar to the structural change tests with epidemic change. That is (for instance), there are two possibly unknown dates (or points in time) for structural change; namely, one is for the drift when the parameter(s) differs from the pre-event period, and the other point where the parameter(s) returns to the level of pre-event period. These changes then may be considered as multiple change points in statistical inferences. The difficulty, however, is how to link these changes with the events of interest given that financial time series is notorious with the reputation of time-varying parametrization.

For instance, let the (known) event window (in discrete time) be denoted as $(T_1, T_2]$, where $T_1 < T_2$, and let the variable of interest (such as abnormal returns for firm i, $i = 1,2,...,n$) be denoted as ε_t^i, the usual event study can be formated as

$$r_{it} = E[r_{it}|\mathscr{F}_t] + \varepsilon_t^i, \tag{3.1.1}$$

$$\varepsilon_t^i = E[\varepsilon_t^i] + \xi_{it}, \tag{3.1.2}$$

where $\mathscr{F}_t$ is the conditional information set for the specification of normal (or expected) return, ξ_{it} is a random noise, $t = 1,2,...,T$, $T \geqslant T_2$. Notice that the conditional expectation $E[r_{it}|\mathscr{F}_t]$ is not restricted to the linear functional form for the conditioning variables. The means of idiosyncratic risks across all firms will have $E[\varepsilon_t^i] = 0$ uniformly for all i's and

t's under the null hypothesis such that the event has no significant impact on stock returns for all t's, or equivalently, $\varepsilon_t^i = \xi_{it}$ almost surely. In particular, if expressed in terms of cumulative abnormal returns, $\sum_{t=T_1+1}^{\tau} \varepsilon_t^i = \sum_{t=T_1+1}^{\tau} \xi_{it}$, $\tau > T_1$, if $E[\varepsilon_t^i] = 0$. However, alternatively, if the impact is essential, the mean $E[\varepsilon_t^i] \neq 0$ for almost all firm i's in studies within the event window. This is equivalent to say that $E[\varepsilon_1^i] = E[\varepsilon_2^i] = \cdots = E[\varepsilon_{T_1}^i] = E[\varepsilon_{T_2+1}^i] = \cdots = 0$, and $E[\varepsilon_{T_1+1}^i] = \cdots = E[\varepsilon_{T_2}^i] \neq 0$.

In other words, the alternative hypothesis simply states that the mean of the abnormal returns are nonzero within the event window across all firms.[1] This is, in fact, a typical example of epidemic change for the multivariate setting where identical length of event window is applied to each firm's returns. Notice that the setting does not assume these nonzero means are identical across all firms. Specifically, these means can be different in signs and magnitudes. On the other hand, if the impact is considered permanent, the alternative hypothesis will become $E[\varepsilon_1^i] = E[\varepsilon_2^i] = \cdots = E[\varepsilon_{T_1}^i] = 0$, and $E[\varepsilon_t^i] \neq 0$ for all $t \in (T_1, \infty)$. That is, the means of all abnormal returns will deviate from zero even after the event windows. In this case, the impacts of events may be considered as persistent and the usual statistical tests for structural changes can be applied also. To reduce the influence of cross-sectional dependence, one can obtain the cross-sectional average of these abnormal returns and denote it as $\bar{\varepsilon}_t$, where $\bar{\varepsilon}_t = \frac{1}{n} \sum_{i=1}^{n} \varepsilon_t^i$, n represents the total number of firms with similar events. In this case, the alternative hypothesis as epidemic change will become $E[\bar{\varepsilon}_1] = E[\bar{\varepsilon}_2] = \cdots = E[\bar{\varepsilon}_{T_1}] = E[\bar{\varepsilon}_{T_2+1}] = \cdots = 0$, and $E[\bar{\varepsilon}_{T_1+1}] = \cdots = E[\bar{\varepsilon}_{T_2}] \neq 0$.

In the statistical methodologies for structural changes however, the change point or date can be allowed as unknown a priori. Yet, in most event studies of corporate finance, the *presumed* event window is usually assumed and includes either a few days prior to the event (announcement) date or starts right on the event date itself according to the null hypotheses. Namely, the tests on abnormal returns are similar to the tests for structural changes with possibly presumed change point within event window. In addition, as will be discussed in Section 3.2, both the CARs tests and the monitoring tests assume some estimation periods where the starting time for CARs tests or recursive on-line detection is available. To generalize the idea, it seems more reasonable to assume that for some unknown dates T_1 and T_2, $E[\varepsilon_1^i] = E[\varepsilon_2^i] = \cdots = E[\varepsilon_{T_1}^i] = E[\varepsilon_{T_2+1}^i] = \cdots = 0$, and for some dates t's within $(T_1, T_2]$, $E[\epsilon_t^i] \neq 0$. In this case, not all abnormal returns for dates t's within $(T_1, T_2]$ will have their means equal to zero. Specifically, there will be no particular cut-off date applied to determine the estimation period, event window, or the starting time for monitoing tests. However, in this case, there could be sufficiently many abnormal returns in the interval $(T_1, T_2]$ that will have $E[\epsilon_t^i] \neq 0$ where T_1 and T_2 are unknown.

The question, however, the conventional CARs tests or the CUSUM tests (including monitoring tests) discussed in Section 3.2 all assume these changes are permanent afterward. These tests, however, can not deal with the general version where the changes are neither permanent afterward nor unanimous across the time period of interest, whether the event window is presumed or not.

In brief, regardless of the impacts of events are temporary or permanent, some thorough investigations for event studies should be extended to consider the entire sample period including post-event window. One difficulty, however, in most conventional event studies, is that subjective decision on the ending date of event window introduces additional havoc for robust conclusions. For instance, if the ending date is set prematurely before the actual expiration of impacts, the statistical results may ignore some residual effects from the events. If the ending date is set too long after the actual expiration of impacts, the statistical tests may include additional distinctives or the other irrelevant noises.

Another issue related to the estimation period and event window is the noisy information or unrelated events that may appear in the estimation period or before/on the event window. This is the so-called contamination problem in the estimation period and event window. One difficulty in analyzing these impacts is how knowledgeable the modeler can be to isolate these impacts from the events of interest. Most of the contemporaneous analyses assume that these noisy information in estimation period (which may be rumors in financial markets) will cause higher variance in the abnormal returns and hence most of the statistical tests (based on cumulative abnormal returns) are not robust or powerful.

Aktas et al. (2007) for instance, show that by specifically introducing the two-state switching market model for the returns, the controlling for unrelated events during the estimation period and on the event date (especially on the increase of event-induced volatility) will make the tests more powerful. For instance, in their study, the return-generating process is denoted as

$$R_i = X_i b_i + \underline{\varepsilon}_i = [\underline{1}, R_m, \underline{D}_i] \begin{bmatrix} \alpha_i \\ \beta_i \\ \gamma_i \end{bmatrix} + \underline{\varepsilon}_i, \qquad (3.1.3)$$

where R_i is a $T^\star \times 1$ vector of returns for firm i, X_i is a $T^\star \times 3$ matrix of explanatory variables for firm $i's$ returns including the $T^\star \times 1$ vecor of returns for a market portfolio proxy, denoted as R_m, $\underline{1}$ is a $T^\star \times 1$ vector of 1's, $\underline{D}_i$ is a $T^\star \times 1$ vector of dummy variable that is equal to 1 at the event date for firm i, and equal to 0 otherwise, $T^\star = T + TE$, where T represents the number of days in estimation period, and TE represents the number of days in event window. If under the homoskedasticity, the covariance matrix of OLS estimator can be shown as

$$\Sigma_{\text{OLS}} = \sigma_i^2 \left(X_i' X_i\right)^{-1}. \qquad (3.1.4)$$

However, if the residuals are assumed as state dependent, let the state variable $S_t = 1$ stand for the low-variance regime, and $S_t = 2$ stand for the high-variance regime, then the variance of residuals for each state can be expressed as

$$\begin{aligned} E[\underline{\varepsilon}_{i,1} \underline{\varepsilon}_{i,1}'] &= \sigma_{i,1}^2 I, \qquad S_t = 1 \\ E[\underline{\varepsilon}_{i,2} \underline{\varepsilon}_{i,2}'] &= \sigma_{i,2}^2 I, \qquad S_t = 2 \end{aligned} \qquad (3.1.5)$$

where $\sigma_{i,2}^2 > \sigma_{i,1}^2$. In this case, if the null hypothesis is true, the variance for residuals will all be equal to $\sigma_{i,1}^2$ regardless of the dates. And by assuming the transition between two regimes is following a Markov chain of order 1, the covariance matrix of OLS estimator can be shown as

$$\begin{aligned} \Sigma_{\text{OLS}} =& p(S_t = 1)\sigma_{i,1}^2 \left(X_i' X_i\right)^{-1} \\ &+ p(S_t = 2)\sigma_{i,2}^2 \left(X_i' X_i\right)^{-1}, \end{aligned} \qquad (3.1.6)$$

where $p(S_t = 1)$ and $p(S_t = 2)$ represent the probability of state 1 and state 2, respectively. Hence, given that $\sigma_{i,2}^2 > \sigma_{i,1}^2$, the lower standard error of OLS estimates under the null hypothesis of no event (that is, $\sigma_{i1}^2 \left(X_i'X_i\right)^{-1}$) will lead to higher rejection frequency for the null especially for the abnormal return γ_i.

Although their findings support the importance of reducing impact of unrelated events during the estimation period, the question is that these noises prior to the events of interest should already propagated through the media or spreaders in financial markets. That is, these impacts should be mediated over time when approaching (or passing) announcement dates or events, which in turn, may reduce the spreadings of rumours. In other words, if detailed specification is desired for the switching states of the event-induced volatility in estimation period and/or on event date, then the explanations for the winding-down of volatility as time approaching and passing the announcements (or events of interest) should also be provided. For instance, if the event-induced volatility increases on/after event date, a good question to ask then is how long this increase may last after the event dates. In other words, provision of the analyses for market adjustments (such as volatility) when approaching (or after) the announcements or events should be as essential as the analysis for the noise-induced and event-induced increases in volatility for event window and estimation period, respectively. Unfortunately, the switching model fails to explain why the switching happens abruptly on that particular event date and stays there for the entire event window without further evolution. Although the two-stage market model performs better than the other tests in their simulations, the result is based on the assumption of event-induced increase of volatility for the abnormal returns such that $\sigma_{i,2}^2 > \sigma_{i,1}^2$.

In particular, their simulations are done under the assumption that the underlying data generation process is indeed following the two-stage market model. To overcome the so-called contamination problem, Aktas et al. (2007) introduce additional variability by adding twice as the standard deviation of the stock returns into abnormal returns for the estimation period. Although this ad hoc setting for unrelated events during estimation period alleviates the possible consequences of contamination problem, it still lacks of theoretical explanations on the evolution of such symptom in the market. Under the hypothesis that no event-induced abnormal returns and no event-induced increase in variance, all the tests seem to perform compatibly well in their simulations. However, when there is a event-induced abnormal return and no event-induced increase in variance (with or without contamination problem), the Corrado's (1989) rank test has more power than the others. Yet, with the introduction of event-induced increase of variance, the rank test performs poorly. In their simulations, two-stage market model performs the best when there is event-induced abnormal return and event-induced increase in variance.

However, their simulations also confirm that the Corrado (1989) rank test statistics, which take both the estimation period and event window together as a single time series, peform better than the other conventional test statistics if there is no event-induced increase in return volatility. In other words, the nonparametric test that does not presume the ad hoc estimation period and event window has more power than the other conventional statistical tests for corporate finance events. This, in turn, suggests that the test statistics (especially reterospective) that apply the entire time series are better than the tests that apply the presumed event dates or windows. The impact of event-induced increase in return volatility can be reduced if one applies some heteroscedasticity-robust estimators for the variance of abnormal returns even though one does no assume any particular event or announcement date.

Karafiath and Spencer (1991) introduce the multiperiod event studies in multivariate setting where the length of event window is allowed to vary across different firms. In particular, the model specifications (for abnormal returns) in the estimation period and in the event window are considered together. The advantage is that the setting bring both the dynamics of the estimation period and the event window simultaneously into the discussions, which is more informative for analysis. However, the assumption for known initial dates for event window is also applied.

To demonstrate the impacts of increased variance due to events, Harrington and Shrider (2007) introduce the so-called true abnormal returns (within the event window) in addition to the idiosyncratic noises for the market model of returns. Their simulation results also indicate that increases in variance within the event window will cause many conventional tests on cumulative abnormal returns biased and prone to easier rejection of the null hypothesis. However, their setting is to intentionally introduce additional noises (the so-called true abnormal returns) in the event window. The difficulty is that such a setting fails to explain why the abnormal returns will have such a clear-cut difference in specification before and after the event dates. That is, there is no explanation that the abnormal returns are completely idiosyncratic noises before the event dates and become true abnormal returns plus noises after the event dates. In other words, the setting does not allow any possible evolving dynamics in the increases of variance. Furthermore, these increases of variance seem permanent after the event dates. They stay on the level and persist over time, especially over the event windows. The increases of variance due to events may possibly be observed empirically in a case-by-case basis.

However, the setting fails to consider that these increases of variance may dwindle down later on as the events become less white-hot or the rumors are gradually resolved. In fact, it is equally feasible to consider that the spreading of rumors before the event dates may cause extra noises for the market volatility. Announcements of these events (or event dates) may resolve some suspicions or uncertainty so that the market volatility may decline. As pointed out by Brown and Warner (1985), the increase of variance at the time of events was hypothesized as a possible cause for lower power of the conventional tests on abnormal returns. Yet, this hypothesized explanation is not the only possible cause for the deficiency of conventional tests. Nor is it a doctrine that all event studies (in corporate finance) should always introduce the event-induced variance increase into the methodology. To resolve this change of variance, and to devise a better methodology for event studies, a robust method that allows the heteroscedasticity of abnormal returns and the possible evolution of the market reaction should be considered. In other words, a methodolgy that allows various patterns of market resolutions (on the news or events) is an alternative for event-study methodology. Furthermore, although it is not necessarily considered as common sense in all cases, the time-varying parameters models are well known in financial time series for empirical finance. Hence, simply verifying or testing to see the parameter change (especially for the mean change of abnormal returns) may not necessarily certify the verification of impacts from the corporate events.

Alternatively, allowing the time-varying parameters in the models for normal (or expected) returns, and then verifying the abnormal returns afterward will provide perhaps some better results for empirical finance. The arguments may be considered as (1) if the time-varying-parameter models suffice to undertake the filterings such that all event-related information are absorbed into the models, the capital market will seem to have sufficient capacity to take advantage these event-related signals and then statistical verifications using the abnormal returns should show that no significant statistical results. In other words, the capital market is efficient in processing newly available information whether it is noisy or systematic. (2) In addition, if the time-varying parameter models are allowed in the expected returns, then conventional CARs tests or CUSUM tests can not be applied here since their purpose of study is only to identify the structural change in parameters. In the following section, some similarities can be found that the conventional CARs tests (using market model for stock returns) are related to CUSUM tests for structural changes even though the statistics are not entirely identical.

3.2 Applications with Cumulative Abnormal Returns (CARs)

For many event studies, applications with cumulative abnormal returns may be considered conventional. However, the usual hypothesis of interest is to consider that the mean of these cumulative abnormal returns tends to zero (at least) asymptotically as numbers of assets and time period expands. However, with the above discussions on event windows, it is likely to consider these tests are similar to the structural change tests that consider the difference in location parameter(s). For instance, if given the following structural equation for the asset returns $i = 1, 2, \ldots, n$, so that the normal (or expected) returns for each asset are expressed in some simple linear regressions such as the so-called market model

$$r_{it} = \alpha_i + \beta_i r_{mt} + \epsilon_{it}, \qquad (3.2.1)$$

the conventional method in using cumulative abnormal returns will be to obtain the estimates for the α_i and β_i and denoted as $\hat{\alpha}_i$ and $\hat{\beta}_i$, respectively, according to the dates from T_0 to T_1, where $(T_0, T_1]$ is the estimation window where the date T_1 may or may not be the actual event date in the time series but chosen by the empiricist of event studies. For instance, one may choose that T_1 is (say) k days prior to the event window of interest and $k > 0$. Notice that the model in Equation (3.1.2) can reduced to the simple mean-deviation model if the "beta" coefficients $\{\beta_i\}_{i=1,2,\cdots,n}$ are assumed to be zero. In this case, the conventional tests for event studies are to considered if there is a change in mean over the event window.

Then, calculate the abnormal returns in the event window (say, from $T_1 + 1$ to T_2) such as $\hat{\epsilon}_{it} = r_{it} - \hat{\alpha}_i - \hat{\beta}_i r_{mt}$. Then, for each asset with ending date τ, $T_1 + 1 \leq \tau \leq T_2$, the cumulative abnormal return is obtained as

$$\tilde{\varepsilon}_{i\tau} = \sum_{t=T_1+1}^{\tau} \hat{\epsilon}_{it}. \qquad (3.2.2)$$

The asymptotic arguments for the cumulative abnormal returns are based on (1) the length of estimation period $(T_0, T_1]$ is sufficiently large, (2) normality assumption for all ε_{it}'s. Specifically, the statistics in Equation (3.2.2) are calculated using the prediction errors for the model in Equation (3.2.1) across the event window $(T_1, T_2]$ by using the estimates for α_i and β_i obtained in estimation period. Hence, the statistics are also based on the assumption that the parameters α_i and β_i are constant over the event window under the null hypothesis that no significant impact is present in the event window.

Namely, in the null hypothesis such that there is no impact from the event(s), these cumulative abnormal returns should be centered around zero within the event window. Specifically, if the event is essential, the mean of these cumulative abnormal returns may differ from zero. If the impact from event(s) is permanent, the difference in means will persist after the event window. In other words, the mean of these abnormal returns will be different from zero from the event date and on. If the impact from these event(s) are only temporal within the event window, these impacts will disappear or decay after the event window (say after T_2+1). If so, and recalling from the statistical hypothesis testing of structural changes, these discussions are similar to the so-called epidemic changes in parameters.

That is, the parameters incur a rapid change within a certain period (with possibly unknown initial time of this period), and return to their original values after the period. In fact, if the event(s) or announcements are only for firm-specific information, it is conceivable to consider the impacts from event(s) are only temporal since the information will not have a permanent impact that all securities need to incorporate the information (or event(s)) as a systematic determinant of their pricings. Following from the setting in Equation (3.2.1), if the alternative hypothesis is to consider that $E[\varepsilon_1^i] = E[\varepsilon_2^i] = \cdots = E[\varepsilon_{T_1}^i] = 0$, and $E[\varepsilon_t^i] \neq 0$ for all $t \in (T_1, \infty)$, where a single point of permanent change in parameters is expected, this represents that the intercept α_i in Equation (3.2.1) has a jump when $t \in (T_1, \infty)$. For instance, the intercept in Equation (3.2.1) will become $\alpha_i + \theta_i$, where $\theta_i \neq 0$, when $t \in (T_1, \infty)$.

If the alternative hypothesis is for epidemic change, then the intercept in Equation (3.2.1) will become $\alpha_i + \theta_i$, where $\theta_i \neq 0$, only when $t \in (T_1, T_2]$. In either case, the cumulative abnormal return as in Equation (3.2.2) is similar to the cumulative sum (CUSUM) test for monitoring structural change in parameter(s) of the regressions such as Equation (3.2.1) in essence. That is to say, the CUSUM tests or similar recursive statistics for monitoring the time-dependent drift or the intercept α_i are equivalently applicable to the event studies, especially when the initial date for the event window is unknown in advance.

For instance, according to MacKinlay (1997), the abnormal return is formulated as

$$\hat{\varepsilon}_{it} = r_{it} - E[r_{it} | \mathscr{F}_{T_1}], \tag{3.2.3}$$

where $\mathscr{F}_{T_1}$ represents the conditioning information for the normal return model up to time T_1, and $t = T_1 + 1, \cdots, T_2$. If the market model is assumed, the asymptotic arguments for the distribution of these abnormal returns are based on the expansion of the estimation period $(T_0, T_1]$. In other words, the statistical properties of the estimated $\hat{\alpha}_i$, $\hat{\beta}_i$ and abnormal returns depend on the sufficient length of estimation period using the method such as ordinary least squares. For instance, let $L_1 = T_1 - T_0$,

$$\hat{\beta} = \frac{\sum_{t=T_o+1}^{T_1} (r_{it} - \hat{\mu}_i)(r_{mt} - \hat{\mu}_m)}{\sum_{t=T_0+1}^{T_1} (r_{mt} - \hat{\mu}_m)^2}, \tag{3.2.4}$$

$$\hat{\alpha}_i = \hat{\mu}_i - \hat{\beta}_i \hat{\mu}_m, \tag{3.2.5}$$

$$\hat{\sigma}_{\varepsilon_i} = \frac{1}{L_1 - 2} \sum_{t=T_0+1}^{T_1} (r_{it} - \hat{\alpha}_i - \hat{\beta}_i r_{mt})^2, \tag{3.2.6}$$

where

$$\hat{\mu}_i = \frac{1}{L_1} \sum_{t=T_0+1}^{T_1} r_{it}, \tag{3.2.7}$$

$$\hat{\mu}_m = \frac{1}{L_1} \sum_{t=T_0+1}^{T_1} r_{mt}. \tag{3.2.8}$$

The abnormal return will be denoted as $\hat{\varepsilon}_{it} = r_{it} - \hat{\alpha}_i - \hat{\beta}_i r_{mt}$, $t = T_1 + 1, \cdots, T_2$.

Hence, the asymptotic arguments then are given under the assumptions that the length of estimation period L_1 will expand and the normal joint distribution for all abnormal returns is given. In other words, the statistics in using the cumulative abnormal returns resemble the CUSUM-based sequential detection tests for structural changes, except that the sequential detection tests do not assume the change point is known (or assumed) in advance and no normality assumption for $\{\epsilon_{it}\}_{t=1,2,\cdots,}$ is used. More specifically, if based on asymptotic arguments, both methods apply sequential setting as in Equation (3.2.10) where the cumulative sums for abnormal returns are expanding through the event window while the length of estimation period (denoted as L_1) is sufficiently large. Notice that under the normality assumption, the abnormal return in the event window where $t = T_1 + 1, \cdots, T_2$, will have the mean zero and the variance as

$$\overline{\sigma}^2_{\varepsilon_i t} = \sigma^2_{\varepsilon_i} + \frac{1}{L_1}\left[1 + \frac{(r_{mt} - \hat{\mu}_m)^2}{\hat{\sigma}^2_m}\right]. \tag{3.2.9}$$

Therefore, as the length of estimation period L_1 expands, the cumulative abnormal return for asset i denoted as

$$\tilde{\varepsilon}_{i\tau} = \sum_{t=T_1+1}^{\tau} \hat{\epsilon}_{it}, \tag{3.2.10}$$

will also be distributed normally with mean zero and variance as $(\tau - T_1)\sigma^2_{\varepsilon_i}$, $\tau \geq T_1 + 1$.

The difficulty of this framework is that the normality assumption is usually violated in the stock returns, especially for the daily observations or data of higher frequencies. In addition, the analysis is based on the constancy of coefficients α_i and β_i over the estimation periods. Unfortunately, it is also well known that the coefficients in the market model are mostly time-varying in different time periods. Additional difficulty lies in the interpretation of the statistical results in using Equation (3.2.10) even though normality assumption is confirmed. The test based on Equation (3.2.10) can not tell if the mean changes are unanimous for the entire event window or, there are sufficiently many (but not all) abnormal returns are of mean changes within the event window. Namely, they will show the changes indeed. But, they can not show the intensity of the mean changes. That is, under normality and if there are sufficiently many $\hat{\epsilon}_{it}$ (and hence, ϵ_{it}) that have nonzero means, Equation (3.2.10) will have the statistically significant result. However, if Equation (3.2.10) is significant statistically, various scenarios or frequencies (including the cases when mean changes concentrated either in early section or later section of event window (or even simply as a spike) may happen for $E[\epsilon_{it}] \neq 0$ over selected event window $(T_1, T_2]$. Therefore, the CARs or, especially the following CUSUM monitoring tests, are only to be considered if there is a structural change alone. No attention is given to the intensity or duration of the impacts following from events. Although the conventional CARs tests are based on normality, and CUSUM monitoring tests are based on Brownian bridge respectively, they are so closely related with the hypotheses of interest since either one applies the residuals of presumed regression models.

Technically, the CUSUM-based sequential detection tests assume the functional central limit theorem where the partial sums of de-meaned abnormal returns (or de-meaned error terms in regressions) will converge in distribution to the Brownian bridge after normalization when the observations in pre-change-point period (or so-called training sample) expands. Recent theoretical works with robust estimation can be found in Chochola et al. (2013) where the emphasis is on the change point of the "beta's." Their analysis is basically sequential in detecting the change point of the "beta's" when additional observations are available. In other words, the analysis is to identify whether there is "going to be" a change point or not as sample size increases sequentially.

Alternatively, there are also retrospective statistical tests for the changes of interest. That is, the tests for the significance of events are to check if there is a change in mean for abnormal returns within the sampled time periods of the observations already collected. However, the conventional CARs tests (based on Equations (3.2.4)–(3.2.10)) are more similar to the monitoring tests for structural change since they both are based on (forward) forecast errors. In fact, the identification of change point(s) (single or multiple) is only to show the instability of the system of interest in either reterospective or sequential approach. More prominantly, whether applying on-line detection tests to identify a change point sequentially as number of observations increases gradually, or to check with parameter changes retrospectively, the essential issue is the setting of the hypotheses of interest. A more important issue to ask is that how long these changes may last, especially for financial data.

Unfortunately, the conventional approach in cumulative abnormal returns assumes the length of event window in advance, while the monitoring tests are only to identify the change that may happen if sequentially additional observations are available. None of them considers the duration of changes or impacts. Not surprisingly, if the length of event window (that is, $T_2 - T_1$) is so short that the impacts from events may disappear promptly, the identification of the change point is relatively less important because that change point (especially under the epidemic change) is possibly a resolution on the noisy new information introduced (which may not be related to the events of interest at all). Noticeably, if the actual (or true) event window is so short that these impacts disappear rapidly, it is hard to verify that the parameter change is actually caused by the event(s) of interest or noisy tradings, or more extremely, due to an outlier or a brief temporal drift even though the length of estimation period L_1 increases asymptotically. For compatibility with the CARs or CUSUM tests, the discussions for tests on parameter changes will start with the on-line detection or sequential approach first. Since there are so many discussions exploring this field in the econometrics literature, only a few related examples for this kind of structural change tests are discussed here. Nevertheless, the similarity of these tests and conventional CARs test is easy to discover.

For instance, the on-line monitoring of "beta" change of capital asset pricing model (CAPM) in Chochola et al. (2013) extends the methods to include robust M-estimates and residuals. However, their work is to monitor the "beta" change in the model. The framework is to assume that the daily log-returns follow the following framework such that

$$r_{it} = \alpha_i + \beta_i r_{Mt} + \epsilon_{it}, \tag{3.2.11}$$

where r_{it} is the daily log-return for the asset i, $i = 1, 2, \cdots, n$ and r_{Mt} is the log-return of the market portfolio at time t, ϵ_{it} is the error term for asset i (or so-called idiosyncratic risk) at time t, $i = 1, 2, \cdots, n$, n is the total number of assets included. Suppose there is a training sample of size m. Intuitively, this is similar to the estimation period (with length denoted as L_1 earlier) for the tests based on CARs where the estimates for the parameters in normal (expected) returns are obtained. As the setting in CARs, the monitoring test assumes that the coefficients such as intercepts and slope coefficients in Equation (3.2.11) are stable within this period such that

$$\begin{aligned} \alpha_{i1} &= \cdots = \alpha_{im}, \\ \beta_{i1} &= \cdots = \beta_{im}, \end{aligned} \tag{3.2.12}$$

for asset i across the time period up to time m, where m is sufficiently large . It is easy to see that these assumptions of stability of parameters resemble the conditions in corporate event studies where the error terms in Equation (3.2.11) are subject to the constant zero means over the entire estimation period. The null hypothesis is then the parameter “beta” will not change such that the beta’s will not change afterward where

$$H_o : \beta_{i1} = \cdots = \beta_{im} = \beta_{im+1} = \cdots , \tag{3.2.13}$$

where the alternative hypothesis is denotes as

$$H_a : \beta_{i1} = \cdots = \beta_{im+k^*} \neq \beta_{im+k^*+1} = \cdots . \tag{3.2.14}$$

More specifically, in Chochola et al. (2013), the stock returns are denoted in a d-dimensional vector of daily log-returns. The setting would imply that these d stock returns are subject to the same (unknown) change point over time if there is a structural change over subsequent time periods. However, according to the setting in Equation (3.2.14), it can be seen that the alternative is to assume a permanent change for the parameter after the (unknown) change point. Although the tests allow for robust estimates of beta’s, and the M-residuals will be applied, the setting only applies to the situations when the events such as beta’s changes for all selected firms may occur simultaneously along the time frame accordingly. In empirical finance, unless the event is noticeable and jointly shared by a number of firms simultaneously, it is hard to obtain such kind of samples in various cases of corporate finance. For instance, for some particular events such as mergers and acquisitions, the samples with simultaneous nature are difficult to obtain because these restructurings may occur with different timing across firms. What is interesting, however, is that this methodology allows to perform the test for some structural changes in “beta’s” (or in “alpha’s”) when certain similar events may happen to a group of firms that have significant cross-sectional dependence.

Nevertheless, the assumptions in Equations (3.2.12) and (3.2.13) seem somewhat far-fetched for financial time series with the well-known yet notorious nature of time-changing parameters. The occurence of these parameter changes may also increase as the frequency of observed financial time series becomes higher. One particular issue is that the asymptotic arguments for these monitoring tests and those in CARs test all demand sufficiently long estimation period for the training samples of size m ,where $m \to \infty$. However, if such an argument is applied, the setting actually assumes that Equation (3.2.12) holds for a sufficiently long period of constancy or stability of parameters. For the conventional CARs tests, this equation is considered to hold within the estimation period. For the monitoring tests, this condition is considered to hold prior to the change point identified later on.

Unfortunately, such an assumption may be considered too "luxurious" empirically in applying the CARs tests or CUSUM-based monitoring tests for parameter changes on financial time series. In fact, one can't even be so sure that the parameters in the so-called in-sample period or training period are actually stable over time for financial time series or data. In other words, the financial time series may have sufficiently many change points even before the event(s) of interest occurs. Even if the parameters for the model of interest are unchanged, the time period for such stability may only be for a very short while in financial time series. Namely, the asymptotic arguments for the monitoring tests may not hold since there is no assurance that such stability will last long enough for the asymptotic properties (or arguments). One possible reason for such a setting of in-sample stability is that the pre-change-point stability of parameters precludes the applications of recursive filtering that focuses on tracking instead of the asymptotic statistical properties such as consistency, weak convergence, or others. For the empirical applications in financial time series, a more essential question to ask perhaps is on how to detect the difference between time-varying nature and dramatic changes in parameters for financial time series and how persistent the difference may last. Unfortunately, these CUSUM-based monitoring tests (or so-called generalized version of CARs test) can not be directly applied to these issues of interest in event studies of corporate finance.

Instead of comparing the estimated "beta's" for the training sample and the after-training-sample period, Chochola et al. (2013) propose the use of functionals of robust M-residuals, where $\phi(\underline{\hat{\epsilon}_t}) = (\phi(\hat{\epsilon}_{1t}),\ldots,\phi(\hat{\epsilon}_{nt}))'$, $\hat{\epsilon}_{it} = r_{it} - \hat{\alpha}_{im} - \hat{\beta}_{im}r_{Mt}$, $\underline{\hat{\epsilon}_t}$ stands for a n-by-1 vector of $\{\hat{\epsilon}_{it}\}_{i=1,2,\cdots,n}$ and $\hat{\alpha}_{im}$, $\hat{\beta}_{im}$ are minimizers of the objective function $\vartheta(.)$ where its derivative

$\vartheta'(.) = \phi(.)$ as its score function

$$\sum_{t=1}^{m} \vartheta(r_{it} - \alpha_i - \beta_i r_{Mt}), \tag{3.2.15}$$

for all i = 1,2,...,n.
The test statistic based on the first $m+k$ observations is defined as

$$\hat{Q}(k,m) = \left(\frac{1}{\sqrt{m}} \sum_{t=m+1}^{m+k} \tilde{r}_{Mt}\phi(\underline{\hat{\epsilon}_t})\right)' \hat{\Sigma}_m^{-1} \left(\frac{1}{\sqrt{m}} \sum_{t=m+1}^{m+k} \tilde{r}_{Mt}\phi(\underline{\hat{\epsilon}_t})\right) \tag{3.2.16}$$

where the matrix $\hat{\Sigma}_m$ is an estimator (based on the first m observations) for the asymptotic variance matrix

$$\Sigma = \lim_{m\to\infty} var\left\{\frac{1}{\sqrt{m}} \sum_{t=1}^{m} (r_{Mt} - E(r_{Mt})\phi(\underline{\hat{\epsilon}_t})\right\}. \tag{3.2.17}$$

and $\tilde{r}_{Mt} = r_{Mt} - (\frac{1}{m}\sum_{t=1}^{m} r_{Mt})$. Then, the test statistic will reject the null hypothesis as

$$\frac{\hat{Q}(k,m)}{q_\gamma(k/m)} \geq c, \tag{3.2.18}$$

for the first time with an appropriately chosen $c = c_\gamma(\ddot{\alpha})$, and $q_\gamma(t)$, $t \in (0,\infty)$ is a boundary function. The stopping time is given as

$$\tau_m = \tau_m(\gamma) = \inf\left\{1 \leq k \leq [mT] : \frac{\hat{Q}(k,m)}{q_\gamma(k/m)} \geq c\right\}, \tag{3.2.19}$$

where $q_\gamma(t), = (1+t)^2\left(\frac{t}{t+1}\right)^{2\gamma}$, $t \in (0,\infty)$, γ is a tuning parameter lies in $[0,\frac{1}{2})$ and the critical value c is chosen such that under the null hypothesis with $\ddot{\alpha} \in (0,1)$

$$\lim_{m\to\infty} Pr(\tau_m < \infty) = \ddot{\alpha}, \tag{3.2.20}$$

and under the alternative hypothesis,

$$\lim_{m \to \infty} Pr(\tau_m < \infty) = 1. \tag{3.2.21}$$

In words, Equation (3.2.20) assumes that the false alarm rate is equal to "α" asymptotically, and the asymptotic power of the test is equal to one. Again, the test statistics are still based on the cumulative sums of residuals from robust estimation to perform test of a single change point. In fact, if examing CARs tests and the monitoring tests closely, they both assume that the estimation period m or $(T_o, T_1]$ grows asymptotically large and, the event window starts from $m+1$ (or T_1+1).

With the rigorous analytical framework given, one typical question to ask is what this identified change (by using on-line detection) may tell us about the impacts of corporate finance event(s). Given that the monitoring test of Chochola et al. (2013) identifies the change point of "beta's," this monitoring test (although based on the robust version of abnormal returns) may not be due to firm-specific event(s) in corporate finance —unless these events are systematic and prevail in the entire capital market. An important issue in corporate finance, however, is what may cause the "beta's" to change over time and how fast or often these changes happen. If indeed, the change happens when the first passage time of test statistic is shown significantly (in a short while or else), what next? Or, more specifically, while the monitoring test for changes in "beta's" provides some generalization of conventional CARs tests, will these changes of "beta's" tell us more about corporate finance event(s)?

For the other monitoring or sequential detection tests on parameter changes that include the change in constant term of regression, there are lots of articles following this fashion. For instance, following the notations and setting of Horváth et al. (2004, 2007), the setting for monitoring test for parameter changes in linear regression is shown as

$$r_{it} = x_{it}'\beta_{it} + \epsilon_{it}, \tag{3.2.22}$$

where x_{it} is a $k \times 1$ vector of explanatory variables (including a constant term), $\beta_{it}$is a $k \times 1$ vector of parameters, ϵ_{it} is a error term for the regression in Equation (3.2.22). Assuming there is no change in the historical sample of size *m*,

$$\beta_{it} = \beta_{io}, 1 \leq t \leq m, \tag{3.2.23}$$

Then, under null hypothesis, there is no change in the parameters,

$$\beta_{it} = \beta_{io}, t = m+1, m+2, \ldots,. \tag{3.2.24}$$

Under the alternative hypothesis, there exists a $k^{\star} \geq 1$ such that

$$\begin{cases} \beta_{it} = \beta_{io}, & m < t < m + k^{\star}, \\ \beta_{it} = \beta_{i\star} \neq \beta_{io}, & t \leq m + k^{\star} + 1. \end{cases} \tag{3.2.25}$$

Horváth et al. (2004) uses the stopping time as

$$\tau(m) = \begin{cases} \inf\{k \geq 1 : & \Gamma(m,k) \geq g(m,k) \\ \infty, & \Gamma(m,k) < g(m,k) \forall k = 1,2,\cdots, \end{cases} \tag{3.2.26}$$

where the detector $\Gamma(m,k)$, is defined as

$$\Gamma(m,k) = \frac{1}{\hat{\sigma}_m} | \sum_{m<t\leq m+k} \hat{\epsilon}_{it} |, \tag{3.2.27}$$

$\hat{\sigma}^2_m = (m-p)^{-1} \sum_{1\leq t\leq m} \hat{\epsilon}^2_{it}$, $\hat{\epsilon}_{it} = r_{it} - x'_{it}\hat{\beta}_{im}$, $\hat{\beta}_{im} = \left(\sum_{1\leq t\leq m} x_{it}x'_{it}\right)^{-1} \sum_{1\leq t\leq m} x_{it}r_{it}$, and the boundary function $g(m,k)$ are chosen so that under the null hypothesis

$$\lim_{m\to\infty} Pr\{\tau(m) < \infty\} = \alpha, \tag{3.2.28}$$

$0 < \alpha < 1$ is a prescribed number similar to the significance level, and under the alternative

$$\lim_{m\to\infty} Pr\{\tau(m) < \infty\} = 1. \tag{3.2.29}$$

The boundary function is shown as

$$g(m,k) = cm^{\frac{1}{2}} \left(1 + \frac{k}{m}\right) \left(\frac{k}{m+k}\right)^{\gamma}, 0 \leq \gamma < \frac{1}{2}. \tag{3.2.30}$$

Notice that the detector or on-line statistic chosen in equation (3.2.27) is similar to the CARs in event studies, except that the cumulative sums are performed without a priori information on the event dates. As shown by Horváth et al. (2004), under the null hypothesis,

$$\lim_{m\to\infty} P\left\{\Gamma(m,k) \le g(m,k), \forall 1 \le k < \infty \right\}$$
$$= P\left\{\sup_{0\le t\le 1} |B(t)|/t^{\gamma} \le c\right\}, \tag{3.2.31}$$

where $\left\{B(t), 0 \le t < \infty \right\}$ is a Wiener process (or so-called Brownian motion). In other words, the normalized cumulative sums of residuals from the presumed asset pricing model of normal returns will converge weakly to the supremum of Brownian motion.[2]

This, in turn, shows that the conventional CARs tests on event studies are the special cases of the on-line monitoring test for parameter change when the possible dates of parameter change for the events of interest are within the assumed event window. Or equivalently, if the unknown change point occurs right inside the assumed event window, the CARs tests are calculated similarly as the CUSUM-based monitoring tests. However, intuitively speaking, the test in Equation (3.2.31) denotes the possible parameter change when the cumulative sums statistic in Equation (3.2.27) goes beyond the boundary function in Equation (3.2.30).

In other words, it is considered there exists a parameter change only when the statistics are beyond the boundary, while allowing tentative fluctuations of parameters and accept the null if otherwise. The question then is if the statistic failed to reject the null, is it appropriate to accept that the parameters are constant over time? What if the parameters are indeed changing while the changes are insufficient to cross the boundary function? Are we willing to accept that and leave our portfolios un-rebalanced even if the parameters in financial time series are mildly changing and do not go beyond the boundary?

While examing financial time series, it is not surprising to see that the parameters can be moderately time-changing in the time period even when no particular events or issues occurred. Time-varying ARCH as shown in Dahlhaus and Rao (2007) can be seen as an example that even the conditional volatility could also be of time-varying coefficients as well. The reason is simple. Financial time series are famous of time-varying parameters since all models applied to them can only be considered as temporal approximations. However, the statistic in Equation (3.2.27) and its asymptotic distribution in Equation (3.2.31) is derived under the null when no structual change of parameters is considered. Yet, even with test confirming there is a structural change, the question remains as whether the change point(s) verifies the essentiality of event(s) or not. The issue is not just the change. How intensive and how long the change could be is perhaps a more important issue in event studies.

For the retrospective tests for structural change, the survey article of Aue and Horváth (2011) provides the essential concepts. For instance, in applying the signal-plus-noise model similar to Equation (3.1.1) in Section 3.1 for stock returns such as

$$r_{it} = \mu_{it} + \epsilon_{it}, \tag{3.2.32}$$

where the μ_{it} represents the expected (normal) return or signal, and ϵ_{it} represents the abnormal return or idiosyncratic risk for stock return i. Notice that the signal μ_{it} can be extended to the linear regression model such that $\mu_{it} = \beta'_{it} X_t$ where β_{it} and X_t are d-dimensional vector of coefficients and attributes, respectively, when applied for the systematic components of stock returns. Now that given the earlier discussions on Section 3.1, it is easy to see when the null hypothesis is given that the corporate event is not essential so that for all asset i's and over time period of interest, $E[\epsilon_{it}] = 0$ across time uniformly, this will imply that

$$\mu_{i1} = \mu_{i2} = \cdots = \mu_{iT} \equiv \mu_i, \tag{3.2.33}$$

where $[0, T]$ is the entire time period of interest. And the alternative hypothesis will become for $t^* \in [0, T]$

$$\mu_{i1} = \mu_{i2} = \cdots \neq \mu_{it^*} = \mu_{it^*+1} = \cdots . \tag{3.2.34}$$

If the linear regression model is assumed, this is equivalent to set the intercept for the regression as constant over time.

Starting from the simplest case in Equation (3.2.16), the retrospective approach of structural change tests using CUSUMs will apply the statistics such as

$$Z_T(k) = \frac{1}{\hat{\sigma}_\epsilon \sqrt{T}} \left(\sum_{t=1}^{[kT]} (\epsilon_{it} - \overline{\epsilon_i}) \right), \tag{3.2.35}$$

where $\overline{\epsilon_i} = \frac{1}{T}\sum_{t=1}^{T} \epsilon_{it}$, $k \in [0,1]$, $\hat{\sigma}^2_{\epsilon i}$ is a consistent estimator for the long-run variance $\sigma^2_{\epsilon i}$ of $\{\epsilon_{it}\}_{t=1,2,...}$. The statistic applied is similar to the CARs except that the normalization is different and the asymptotic distribution also differs from the normality assumption in CARs test. In particular, the change point time is unknown a priori. However, the objectives of these two approaches are identical - they all want to see if there's a structural change within the event window or over the time period of interest. Given that functional central limit theorem as $T \to \infty$, the weak convergence for cumulative sums in equation (3.2.7) will show that

$$Z_T(k) \xrightarrow{d} \overline{B}(k) \tag{3.2.36}$$

where $\overline{B}(k)$ is a standard Brownian bridge, $k \in [0,1]$. The CUSUM test statistic however, is emphasizing on finding the change point by setting $k = \frac{t}{T}$, where $1 \leq t \leq T$,

$$\psi_T = \max_{1 \leq t \leq T} Z_T(\frac{t}{T}), \tag{3.2.37}$$

and when $T \to \infty$

$$\psi_T \xrightarrow{d} \sup_{0 \leq k \leq 1} \overline{B}(k). \tag{3.2.38}$$

Yet, both monitoring test and the CUSUM test for structural change in Equations (3.2.16), (3.2.27), and (3.2.37) all assume that the change is permanent. That is, once the parameter changes, it will stay that way from then on (at least) hypothetically. However, for the event studies in corporate finance, it may show that these changes are only temporary even if one accepts that the parameters are stable before the change point. The reason is that the impacts from the events in corporate finance are usually considered idiosyncratic if the events are not systematic or influential toward the entire capital market. If so, these changes of corporate events should only be temporary. Given so, both the tests using CARs and CUSUM tests fail to consider the necessary setting to verify the impacts of corporate finance events regardless of how robust the statistical procedures are applied.

Notice that if under the market model in asset returns, the conventional tests with cumulative abnormal returns actually assume that the coefficients β_i's are unchanged over the entire sample period in this case. In particular, as shown in Equations (3.2.3)–(3.2.5) of conventional CARs tests, the abnormal returns are actually forecast errors based on the coefficients estimated in the pre-event estimation period. Namely, the hypothesis of interest is actually asking to see if the intercepts may have a drift or jump within the preselected event window, while assuming that the "beta's" remain constant whenever the time period is. To be more precise, this is only to ask if the intercepts may fluctuate within the event window while assuming the coefficients β_i hold constant over time. In fact, this appears as a very limited case for structural change in parameters. More explicitly, according to Equations (3.1.1) and (3.1.2), the hypothesis testing under the temporary parameter change in $(T_1, T_2]$ as epidemic change can be expressed as

$$H_0 : \begin{cases} r_{it} & = \alpha_i + \beta_i r_{mt} + \xi_{it}, \\ & t \in (0, \infty) \end{cases} \tag{3.2.39}$$

$$H_a : \begin{cases} r_{it} & = \alpha_i + \beta_i r_{mt} + \xi_{it}, \\ & t \in (0, T_1] \cup (T_2, \infty) \\ r_{it} & = (\alpha_i + \theta_i) + \beta_i r_{mt} + \xi_{it}, \\ & t \in (T_1, T_2] \end{cases} \tag{3.2.40}$$

where H_0 and H_a represent the null and alternative hypotheses, respectively. If only the constant-mean model is applied to obtain the abnormal returns, Equations (3.2.39) and (3.2.40) can be replaced with the setting as $\beta_i = 0$ for all i's. And if the change of parameter is permanent, and there is only one change point of parameter α_i over time, then the alternative hypothesis will become

$$\tilde{H}_a : \begin{cases} r_{it} & = (\alpha_i + \theta_i) + \beta_i r_{mt} + \xi_{it}. \\ & t \in (T_1, \infty) \end{cases} \tag{3.2.41}$$

The same reasoning applies also to the cross-sectional average of abnormal returns because the set of Equation (3.1.2) across all selected assets can be re-written as

$$\bar{r}_t = \bar{\alpha} + \bar{\beta} r_{mt} + \bar{\varepsilon}_t, \tag{3.2.42}$$

and denote that the cumulative average abnormal returns as $\bar{\varepsilon}^a_\tau = \sum_{t=T_1+1}^{\tau} \bar{\varepsilon}_t$, where $\bar{\varepsilon}_t = \frac{1}{n}\sum_{i=1}^{n} \varepsilon^i_t$, $\bar{\alpha} = \frac{1}{n}\sum_{i=1}^{n} \alpha_i$, $\bar{\beta} = \frac{1}{n}\sum_{i=1}^{n} \beta_i$. Again, the conventional tests with cumulative average abnormal returns show that if the cross-sectional average intercept $\bar{\alpha}$ (based on the estimates obtained in estimation period) will have a change point or not within event window given that $\frac{1}{n}\sum \theta_i \neq 0$ in the event windows. This framework can be generalized to accommodate the cross-sectional dependence possibly due to event dates clustering among the selected firms' abnormal returns by using the weighted average of asset returns. For instance, define the weighted average of asset returns as

$$r^\star_t = \sum_{i=1}^{n} w_i r_{it}, \tag{3.2.43}$$

where the scheme of the (say, randomly generated) weights satisfies the convergence condition that $\lim_{n \to \infty} w_n = 0$, and $\sum_{i=1}^{n} w_i = 1$ it is straightforward to re-write the set of Equation (3.1.2) across all assets into n

$$r^\star_t = \alpha^\star + \beta^\star r_{mt} + \varepsilon^\star_t, \tag{3.2.44}$$

setting $\alpha^\star = \sum_{i=1}^{n} w_i \alpha_i$, $\beta^\star = \sum_{i=1}^{n} w_i \beta_i$, and $\varepsilon^\star_t = \sum_{i=1}^{n} w_i \varepsilon_{it}$, the hypothesis setting in Equation (3.2.12) also applies. Given the assumption that the cross-correlations among these abnormal returns can be alleviated using weighted series, the issue of event-date clustering can be resolved. In other words, assuming that the "beta's" are constant in the event windows, the same setting for using the cumulative abnormal returns is still valid such that the hypothesis of interest is to see if $\alpha^{\star}$has a change point in the event windows, accordingly. In this setting, the test of whether $\alpha^{\star}$ has a change point or not is equivalent to the tests using cumulative abnormal returns if the assumption of known event dates is relaxed. The difficulty, however, is that although the weighting schemes alleviate the cross-sectional dependence among $\{\varepsilon_{it}\}_{i=1,2,\ldots,n}$when estimating the parameters α_i's and β_i's in normal (or expected) returns, some cross-sectional information may also be lost due to the weighting schemes.

Hence, from this perspective, the conventional tests based on CARs are equivalent to special cases of structural change tests for unconstrained parameters (such as α_i's) within the assumed event window provided that the "beta's" are assumed to hold constant over the entire estimation period. In essence, across all selected firms, and if the event windows $(T_1, T_2]$ are known where the change points may actually locate within them, the cumulative abnormal returns in Equation (3.2.10) are similar to the CUSUM statistics applied to the structural change tests in linear regressions of market model for each firm's asset returns-assuming that the change of parameters is permanent. From the initiative of hypothesis of interest, the purpose of using the cumulative abnormal returns is equivalent to that of the tests on change point in intercepts for the set of Equation (3.2.27) among all selected firms.

The arguments can be extended to more complicated asset pricing models for normal (or expected) returns other than the market model. Some may argue that the "beta's" are assumed constant since the systematic component for normal (or expected) returns would not change given that the firm-specific corporate events such as mergers and acquisitions are only unsystematic and diversifiable. However, many empirical studies show that the "beta's" (of market model) may still vary over time even though there is no specific corporate event identified in the sample periods. In particular, the setting assumes that the date where the parameters change is known a priori. While it is reasonable to argue that some firm-specific information may leak out to the market and cause some rumors in the market, it is hardly feasible to know when the impacts may occur to the market and returns in advance.

The above discussions for the nonstochastic change in the means of abnormal returns (and with respect to the market model, for example) show that there are two categories of studies: (a) a permanent change in the abnormal return's mean and (b) a temporal change in the abnormal return's mean. In either case, two methods can be devised: (1) retrospective methods and (2) sequential methods. The retrospective methods are in search of any change point within the sampled observations, and their asymptotic arguments depend on the growth rate of the sample size. This perspective of empirical studies is to prove or disprove (say) certain hypotheses of interest using the historical observations collected.The sequential methods, however, are to identify if there is any change point when additional sequential observations are available. Although the recursive identification on the change point(s) is useful for system control, the monitoring scheme usually assumes that the sample size of observations prior to the change point is sufficiently large to apply the asymptotic arguments. Besides, the parameters of interest are assumed stable within the sample period of observations prior to the unknown change point.

Unfortunately, for financial time series in particular, these conditions are difficult to hold. In empirical finance, one can easily find out that the coefficients in the models specified for normal returns (say, market model for instance) are not stable over various time periods and with different frequencies of data selected. More ironically, it is difficult to consider that the coefficients of interest are stable or constant within the in-sample period for financial time series (especially in high frequencies) before one can identify a change point when sequential out-of-sample observations are collected. Sequential identification of changes is essential indeed for many modelings on financial time series. However, with the dramatic time-varying parameters in financial time series, it may be considered luxurious if such a rarity for in-sample stability (of parameters) may appear.

Comparing the conventional CARs tests and robust CUSUM-based monitoing tests, it is straightforward to verify the similarity between traditional approach in using (cumulative) abnormal returns and the tests on the change point(s) over the event window(s). In particular, these methods of structural change although mostly equivalent to the tests using cumulative abnormal returns (when the event dates are known in advance) are all subject to the hypotheses of finite numbers of change points, especially under the market-model setting for normal returns. Given that there is no a priori knowledge of the evolution of parameters within the event windows (even when the event dates or length of event windows are allowed to vary across different firms), justifications on change points do not suffice to prove the significance of the events of interest.

The question remains, however, even if confirmation on the parameter changes is granted, it is difficult to tell that these parameter changes are indeed caused by the corporate-specific events or instead, these are market adjustments for newly available noisy information (such as spreading rumors or conjectures) that are possibly unrelated with the particular events of interest in financial time series. Even in the tests for epidemic change, the studies are actually finding if there are two change points: one is for the intercept to jump to $\alpha_i + \theta_i$, and another one when it returns to original level α_i. In other words, the tests for epidemic change are in fact, studies for multiple change points. As a result, extensions toward more than two change points seem to cope better with empirical findings since the parameters α_i's and β_i's for market model are likely to fluctuate over time even when there is no specific corporate event incurred.

Another difficulty with these tests and conventional CARs tests is that the models applied to specify the normal returns are presumed to be correctly specified. For instance, the functional forms and the inclusion of relevant explanatory variables are assumed. Hence, the tests applied to verify if there is a change in the mean of abnormal returns may result as the same tests for structural change in the presumed models. In other words, the tests are meaningful only if the models for normal returns are correctly specified. Then, in that case, the parameter changes will be interpreted properly as the impacts from the (announcements) of the events. However, all these models (for asset pricing models or else) applied to fit the normal returns should at best be considered as filters to screen out the so-called systematic components of asset returns. Consequently, the discussions for change point(s) here are only to demonstrate the similarity between the CUSUM-like methodology and the conventional event-study tests of cumulative abnormal returns and to provide critiques that these tests are relatively limited in applicability. In addition, hypotheses for epidemic change are also introduced here when the impacts from events may disappear as time span expands across the so-called event windows. What is more important, however, is that the similarity of the CUSUM-based monitoring tests and the conventional CARs tests does not necessarily imply that the verifications of change in parameters (say, the intercepts or means) are equivalent to verification for the significance of events of interest.

The hypothesis of significant events (where $E[\varepsilon_t^i] \neq 0$ within the event window) is only a sufficient condition for the parameter changes. Hence, these tests can only be considered as diagnostic, not specification tests. The statistical significance of parameter changes can not justify that the cause of structural changes are actually from the events of interest. The reason is that there are too many possible reasons or scenarios for the structural changes in the models for normal (or expected) returns. To the best extent, the so-called on-line detection or monitoring tests are only to confirm that the parameters in the financial time series are possibly subject to structural change of unknown change points. If tests of change in coefficients are of interest, the tests for epidemic change are possibly more suitable for the corporate event studies. Yet, even with epidemic changes, the tests will only result as confirmation for multiple change points for models of financial time series. Nothing more. In fact, a better way to consider impacts from corporate events is perhaps on the duration between change points (say, under the epidemic change scenarios) if changes of parameters are concerned. Namely, if changes of parameters are the emphases of study, then how long the changes may last across different change points is more important in assessing the impacts. Based on the above discussions, a few results can be stated:

Result 1: The CARs tests and CUSUM tests are similar in (a) hypotheses of interest since both will result in testing the change of coefficient(s) of the regression models assumed for the empirical asset pricing models, (b) if the change point indeed lies within the assumed event windows. In particular, the CARs tests are similar to the restricted form of CUSUM tests that normality assumption is applied to the abnormal returns, typically.

Result 2: Both tests are to identify if there is only one change point in the period of interest. When there is a possibility of multiple change points, these tests may not suffice to identify the underlying models.

Result 3: Since financial time series are almost always prone to time-varying changes in parameters or models, the results from both tests may still require verifications that the parameter change indeed, is the consequence of the event(s)

Result 4: Both tests are incapable to describe the duration of the impacts from events or structural changes.

Similarly, further detailed parametric specification that introduces time-varying intercepts and gradually-changing "beta's" for market model, Cyree and Degennaro (2002) generalize the event-study techniques to allow for parameter shifts (such as both α_i's and β_i's), variance shifts, and firm-specific event periods with different event dates and lengths. One of their major arguments is that any robust event study in corporate finance ought to consider the possible parameter changes in model specification before applications of CARs analyses. In particular, they allow the systematic risk such as β_i's to gradually change over the event periods and with regime-switching variance of the true abnormal returns. With various settings given to the time-varying-coefficient market model (which allows for multi-change points), part of their intent is to verify the power of test with the cumulative average abnormal returns (CAARs). In fact, their empirical findings show that the CAARs for Day(−1) (that is, the impact on one day before the event date) are only significant at 10 percent significance level when consider all these possible scenarios of parameters in specifying normal (expected) returns. More specifically, their setting can be specified as follows. For a firm-specific event date $T_{1,i}$, $i = 1, 2, \cdots, n$, and gradually-changing β_{it}, the null and alternative hypotheses (denoted as H_0 and H_a, respectively) are

$$H_0 : \begin{cases} r_{it} & = \alpha_{it} + \beta_{it} r_{mt} + \xi_{it}, \\ & t \in (0, \infty), \end{cases} \tag{3.2.45}$$

$$H_a : \begin{cases} r_{it} & = (\alpha_{it} + \delta_i) + \beta_{it} r_{mt} + \xi_{it}, \\ & t \in (T_{1,i}, \infty). \end{cases} \tag{3.2.46}$$

In particular, the firm-specific random variable ξ_{it} in abnormal returns may be subject to different schemes of variance such that

$$\text{Var}(\xi_{it}) = \begin{cases} \sigma_{\xi}^2(i), & t \in [0, T_{1,i}] \\ \overline{\sigma}_{\xi}(i), & t \in (T_{1,i}, \infty). \end{cases} \tag{3.2.47}$$

Namely, the setting simply assumes that (if the change is permanent) the time-varying intercept α_{it} should contain a jump δ_i (which can also be stochastic) around the unknown firm-specific event date $T_{1,i}$, $i = 1, 2, \cdots, n$, where the variance of abnormal returns are of different schemes for estimation period and event window, respectively.

However, with all these complexities, it would be more straightforward to specify the setting as the switching model with jump(s) across the firm-specific event dates. In other words, instead of applying CAAR's on the setting, model specification with Equations (3.2.45), (3.2.46) and (3.2.47) when using parametric methods to estimate and test the model will seem more powerful. Nevertheless, the setting requires estimations on the firm-specific event date $T_{1,i}$, $i = 1, 2, \cdots, n$ given that no a priori knowledge is applied. The question remains however, with such complexities allow, whether the tests using CAAR's will provide additional evidence on event studies or not. In particular, in Cyree and Degennaro (2002), the gradually changing systematic risks β_{it} are specified as

$$\begin{cases} \beta_{it} & = \left[\beta_{i,1} + \Delta^E \beta_{i,1} D_{1,i,t} + \Delta^P \beta_{i,1} D_{2,i,t}\right] r_{mt} + \xi_{it}, \\ D_{1,i,t} & = \begin{cases} 1, & t \in (T_{1,i}, T_{2,i}) \\ 0, & t \notin (T_{1,i}, T_{2,i}), \end{cases} D_{2,i,t} = \begin{cases} 1, & t \in (T_{2,i}, \infty) \\ 0, & t \notin (T_{2,i}, \infty), \end{cases} \end{cases} \tag{3.2.48}$$

where $\Delta^E\beta_{i,1}$ and $\Delta^P\beta_{i,1}$ represent the change of systematic risk during the event period, and post-event period, respectively. In addition, the firm-specific random variable ξ_{it} also follows equation (3.2.47) of different schemes of variance. To be more specific, they even consider some ad hoc specification for the gradual change in beta's for these periods such that

$$\Delta^E\beta_{i,1} = \beta_{i,2}(T_{1,i} - t)(t - T_{2,i}) + \beta_{i,3}(t - T_{1,i}) + \nu_{it}, \quad (3.2.49)$$

$$\Delta^P\beta_{i,1} = \beta_{i,3}(T_{2,i} - T_{1,i}) + \eta_{it}, \quad (3.2.50)$$

where ν_{it}, η_{it} are white noises. With all these parametrization on the time-varying coefficients and event-induced volatility in abnormal returns, it seems that the applications of CARs (or, CAARs) are not entirely essential now.

Nevertheless, such a conclusion is only primitive since all these time-varying coefficients are subject to presumed schemes of variation over time. The schemes of time-varying "beta's are not necessarily smooth or gradual, especially across the entire window of high-frequency observations. It is also suspicious that these time-varying smoothing actually obscure the impacts from the firm-specific information when Equations (3.2.49) and (3.2.50) are introduced. For instance, by substituting these two equations into Equation (3.2.48), the authors show that the abnormal return (say, denoted as $\varepsilon^{\star}_{i,t}$) will become $\xi_{it} + \nu_{it} r_{mt}$ when in the event period, and $\xi_{it} + \eta_{it} r_{mt}$ in the post-event period. Assuming that ν_{it} and r_{mt}, η_{it} and r_{mt}, are all mutually statistically independent, then it is easy to find that the volatility in abnormal return will become $\text{Var}(\xi_{it}) + \sigma^2_{\nu i}\sigma^2_{mt}$ or $\text{Var}(\xi_{it}) + \sigma^2_{\eta i}\sigma^2_{mt}$, respectively. Given that σ^2_{mt} is the volatility of market index return that represents a proxy for market volatility, the volatility of abnormal return may then include an argument of market volatility, or so-called systematic component of asset returns.

If this is the case, the separation between normal and abnormal returns to perform event studies in corporate finance seems weakened by these parametrization because the market volatility will influence the variance of abnormal returns where the latter ones are supposedly diversifiable. In other words, the reduction of power in the CAARs tests can possibly be due to overparametrization for smoothing in the models. What is more interesting however, is that their finding shows that (1) time-varying-coefficient models are influential in the specifications of normal (or expected) returns, (2) estimation of these time-varying coefficients should be adaptive in the sense that less subjective parametrization are applied, (3) some more robust statistical tests such as extended functionals of CARs (instead of CARs themselves) are needed for event studies. In other words, testing parameter changes in financial time series does not suffice to provide robust analysis for event studies in corporate finance unless they are under some exceptional occasions. Generalization for statistical tests beyond the CAR's approach for event studies in corporate finance is crucially needed. Parameter changes may possibly be considered as common knowledge in financial time series and hence tests (based on the abnormal returns) that have already taken these parameter changes into pre-filtering for the normal (or expected) returns would become more applicable for corporate finance issues.

Part II

Event Study Methodology II

Chapter 4

Recursive Estimation for Normal (or Expected) Returns

Introduction

Given that the updating information continue flowing into the capital market, modification on the model specifications of systematic components of normal returns are necessary for further discussions on firm-specific abnormal returns. In this chapter, since all models that approximate normal returns are prone to time-varying parameters, some recursive estimation methods are shown to cope with this nature. Given that the systematic components of asset returns can be approximated by proposed (time-varying coefficient) theoretical models of nondiversifiable variables or proxies, all the model specifications are similar to the adaptive filters for the data stream.

Although there are various methods for time-varying parameters in regressions, the contents of this chapter focus on the stochastic algorithm, recursive least squares, and system identification methods applied in tracking discussed in the literature of adaptive filtering. The intent is to introduce the filtering methods to approximate the systematic components of asset returns without overparametrization and to avoid ad-hoc specifications in empirical asset pricing models of normal (or expected) returns.

One advantage in the filtering with time-varying parameters is that it may help to correct some pricing noises that are subject to concurrent and diversifiable information in the market. In addition, the tracking with adaptive filtering provides more up-to-date adjustments for the normal (or expected) returns.

Thus, event studies for corporate finance based on (the functionals of) abnormal returns will be less contaminated with possible time-varying drifts, temporal impulses, and irrelevant noisy information. Another advantage is that no assumption of change point(s) (or event window) is provided to the equation of normal (or expected) returns. Hence, there is no need to verify if the change points actually lie within the assumed window or not.

Given that the purpose for the adaptive filters is to perform tracking on the underlying systematic components of stock returns, the emphasis is not on the (asymptotic) statistical properties such as stochastic convergence, consistency, or others for the specifications or parameters of interest. Instead, the purpose is to obtain filtering for the systematic components so that the tracking errors are as small as possible and the system identification is stable. These emphases have been discussed mainly as the recursive identification in engineering or automatic control. Given that most corporate finance event-study methodologies focus on the statistics of abnormal returns, the identification of the systematic components of stock returns is essential. In particular, since these methods can be applied in tracking the systems of time-varying parameters, they are useful in providing an alternative method for the estimation of normal (or expected) returns when (unknown) change time(s) for parameters prevail in most financial time series.

4.1 Why recursive?

Notice that the systematic components of stock returns are prone to change over time when new information is available, particularly upon receiving the market-wise information. These changes that are caused by new information may be well-approximated by the time-varying parameters in the asset-pricing models for normal (or expected) returns. To accomondate the changing expectations from the market, updating the information and modifying the parameters (or models) continuously over accessible tools or data is necessary.

Although updating normal (or expected) returns can be obtained in using dynamic forecasts (say, in time series regressions) when the parameters are constant, these forecasts are not robust enough to adjust for abrupt changes in expectations or in coefficients. Hence, in adjusting the normal (or expected) returns for the changing environment in capital market, recursive estimation for tracking these essential natures of systematic components in stock returns is useful. Even if the coefficients of underlying models are constant, such recursive estimation will provide on-line updatings for the models and forecasts where conventional statistical properties such as consistency can be obtained.

Developed in the literature of system identification of control theory for stochastic systems, these recursive estimation methods are feasible for applications in the corporate finance as well. Similar to the system identification for signal processing when immediate tracking (say, for moving targets or objects) must be done, financial market provides a rapidly changing environment where speedy adjustments either for expectations or for portfolio rebalancing are also in need. Hence, to provide certain adjustment schemes for model specifications in coping with new information, applications with recursive estimation when receiving the updated signals become some vehicles for adaptation in expectations.

As the information flow in the capital market, the tracking for systematic components such as normal (or expected) returns in financial securities is similar to the task in tracking a target or object for remote sensing or control. Since the investors in capital market are keen enough to update their information and modify their expectations, it is reasonable to assume that such analytical schemes can be applicable to the empirical modeling for financial data. Specifically, if the investors will endeavor to search for any tractable regularity in using all available information when individual optimality is concerned, it is reasonable to consider that the schemes of system identification for adaptive control can be applicable for the estimation of normal (or expected) returns.

Furthermore, the time-varying coefficient models for normal (or expected) returns help to reduce the possible model specification errors (such as drifts or jumps in parameters) over different time horizons. More specifically, these recursive algorithms do not impose subjective conditions such as estimation periods, event windows, or after-event periods. Temporal dependence or heteroskedasticity such as ϕ-mixing conditions of stochastic processes are also allowed in the data stream.

In the following, the recursive estimation such as Kalman filter, least mean squares and recursive least squares algorithms are discussed. In particular, the asymptotic properties of these algorithms are also covered in the general setting of identified empirical asset pricing models with nondiversifiable variables. There are vast amounts of research works for these properties. Thecurrent contents only cover the contribution such as in Guo (1990, 1994), Chen and Guo (1991), Guo et al. (1991) as an introduction to this field of studies for empirical finance. In particular, these recursive algorithms are available in various statistical software programs nowadays.

Although only asymptotic stability conditions for these tracking algorithms are presented here, their calculations can be peformed conveniently in many software packages such as Stata, R, or others. Given the technicalities, the extensions such as stochastic approximation with adaptive step-size or with purturbation method (Tang et al. (1999)), self-tuning recursive least squares, and many others are left for references.

4.2 The algorithms

Assumption 4.A.1: Let the identified (linear) model for normal (or expected) returns of security i at time t be denoted as

$$r_{it} = \psi_t' \theta_{it} + \epsilon_{it}, \tag{4.2.1}$$

where $i = 1, 2, \cdots$, θ_{it} is a *k-by-1* vector of time-varying coefficients (including the constant term) in the identified model, $k \geq 1$, ψ_t' is denoted as the transpose of ψ_t, ψ_t represents a *k-by-1* vector of nondiversifiable explanatory variables in the model of systematic components,[1] ϵ_{it} is the error term and consequently, the abnormal return of interest for security i.

As stated in Chapter 2, these selected explanatory variables are denoted as nondiversifiable to identify the normal returns. The parameters for these explanatory variables may also be time-varying, including the drifts. Hence, to obtain robust abnormal returns for corporate finance events, some trackings or recursive estimations for the normal returns must be applied first. Given the time-varying parameters in the model, two issues are essential. One is the stability of the adaptive algorithms and the other is the asymptotic convergence of tracking errors when various recursive algorithms are applied. A typical property for the models of time-varying parameters is that since the parameter variation does not vanish, usual recursive algorithms such as constant-step stochastic approximation for constant-coefficient model can not apply..

For instance, it is well-known that Kalman filter is usally applied in this case if the dynamics for the time-varying parameters is assumed. Given the linear regression model in equation (1) and with the following Markov model for the time-varying parameters θ_t such that for $t \geq 0$,

$$\theta_{i,t+1} = F\theta_{it} + \omega_{it+1}, \quad E|\theta_{io}|^2 < \infty \tag{4.2.2}$$

where F is a *k-by-k* matrix for transition of θ_t and the eigenvalues of F lie in or on the unit circle, these noises $\{\epsilon_{it}\}$ and $\{\omega_{it}\}$are assumed as mutually independent and serially indepent with zero mean and covariances $E[\omega_{it+1}\omega'_{i+1}] = Q_\omega \geq 0$, as a positive-semidefinite matrix and $E[\epsilon^2_{i,t+1}] = R_{\epsilon i} > 0$, as a positive-definite matrix.

The Kalman filter accordingly is stated as

$$\begin{aligned} \hat{\theta}_{it+1} &= F\hat{\theta}_{it} + \frac{FP_t\psi_t}{R + \psi'_t P_t \psi_t}(r_{it} - \psi'_t\hat{\theta}_{it}) \\ P_{t+1} &= FP_tF' - \frac{FP_t\psi_t\psi'_t P_t F'}{R + \psi'_t P_t \psi_t} + Q, \end{aligned} \tag{4.2.3}$$

where $P_o \geq 0$ as a positive semi-definite matrix and $Q > 0$, $R > 0$ are considered as a priori estimates for Q_ω and $R_{\epsilon i}$, respectively. In particular, when $\{\omega_{it}, \epsilon_{it}\}$ are Gaussian noises, the above Kalman filter is the optimal estimator for θ_t in the sense that for the sigma fileds $\{\mathscr{F}_{t-1}\}_{t=1,2,\cdots}$ generated by the past information of$\{r_{it}, \psi_t, \epsilon_{it}\}$,

$$\hat{\theta}_{it} = E[\theta_{it}|\mathscr{F}_{t-1}], \quad P_t = E[\tilde{\theta}_{it}\tilde{\theta}'_{it}|\mathscr{F}_{t-1}], \tag{4.2.4}$$

provided that $Q = Q_\omega$, $R = R_{\epsilon i}$, $\hat{\theta}_{io} = E[\theta_{io}]$, and $P_o = E[\tilde{\theta}_{io}\tilde{\theta}'_{io}]$, where $\tilde{\theta}_{it} = \theta_{it} - \hat{\theta}_{it}$ represents the tracking errors for time-varying parameters θ_{it}. In other words, if the underlying variance or covariance of noises are known, the Kalman filter will obtain the optimal tracking for the system of time-varying parameters. Unfortunately, these matrices are usually not known in advance.

For extending classicial results, Guo et al. (1991) generalize the asymptotic arguments for tracking error bound of Kalman filter where the time-varying parameters follow a Markov model and without the normality assumption. Yet, the results, although robust for nonnormal distributions, still require explicit specification for dynamics of time-varying parameters. Certainly there are many recursive algorithms for tracking the time-varying parameters for system identification. However, most of them require the specification of time series models on these time-varying parameters. For instance, Brockett et al.(1999) assumes the Markov-model dynamics for the time-varying beta's of market model. Their applications also include specifications of GARCH model for the error terms in the equations of stock returns.

In the following, the recursive algorithms are categorized into two sections: (1) with dynamic specification of the time-varying parameters and (2) without dynamic specification of the time-varying parameters. As stated in Equation (4.2.2), the extended Kalman filter algorithm can be applied here. Many research studies have done for various extentions. Following Guo (1994), the recursive algorithms, such as Kalman filter, least mean squares, and recursive least squares are discussed for tracking on time-varying parameters without the normality assumption. For simplicity, let the matrix $F = I$ as a *k-by-k* identity matrix and let $\tilde{\theta}_{it} = \theta_{it} - \hat{\theta}_{it}$ be the tracking error at time t, where the dynamics of the time-varying parameters assumed to follow a random-walk model such that

$$\theta_{it} = \theta_{it-1} + \Delta_{it}, \tag{4.2.5}$$

where Δ_{it} represents the variation of θ_{it} at time t. By assuming different models or properties of Δ_{it}, various extenstions obtain the bounds for tracking errors of recursive algorithms. For any matrix X (including vector) heretofore, the norm of matrix X

(denoted as $||X||$) is defined as $\left\{\lambda_{\max}\left(XX'\right)\right\}^{1/2}$, where $\lambda_{\max}$ represents the maximum eigenvalue of the matrix of interest. The other recursive algorithms such as least mean squares and recursive least squares are shown as

Least mean squares

$$\hat{\theta}_{it+1} = \hat{\theta}_{it} + L_t(r_{it} - \psi_t'\hat{\theta}_{it}), \qquad L_t = \mu\frac{\psi_t}{1+||\psi_t||^2} \tag{4.2.6}$$

where $\mu \in (0,1]$ is called the step size.

Recursive least squares

$$\begin{aligned} \hat{\theta}_{it+1} &= \hat{\theta}_{it} + L_t(r_{it} - \psi_t'\hat{\theta}_{it}) \\ L_t &= \frac{P_t\psi_t}{\alpha + \psi_t'P_t\psi_t} \\ P_{t+1} &= \frac{1}{\alpha}\left[P_t - \frac{P_t\psi_t\psi_t'P_t}{\alpha + \psi_t'P_t\psi_t}\right] \end{aligned} \tag{4.2.7}$$

where $P_o > 0$, and $\alpha \in (0,1)$is a forgetting factor.

In general, the recursive equations in (3), (6) and (7) can be expressed as

$$\tilde{\theta}_{i,t+1} = (I - L_t\psi_t')\tilde{\theta}_{it} - L_t\epsilon_{it} + \Delta_{i,t+1}. \tag{4.2.8}$$

This, in turn, can be expressed as in the following stochastic difference equation

$$x_{t+1} = (I - A_t)x_t + \xi_{t+1}, \tag{4.2.9}$$

where $\{A_t\}_{t=0,1,\cdots}$ is a sequence of *k-by-k* random matrices, x_t is the *k-by-1* vector of tracking errors, and $\{\xi_{t+1}\}_{t=0,1,\cdots}$ is the disturbance. Given so, the emphasis of these recursive algorithms is on the asymptotic stability of the stochastic difference equation (9) of tracking errors. Apparently, due to the time-varying parameters of the system, the adjustment matrix L_t (and hence, A_t) needs to update with new information. And hence, the stability of recursive algorithms requires analysis of the sequence of random matrices $\{A_t\}_{t\geq 0}$. According to

Guo (1994), the norm and stability of random matrix (or vector) sequence is defined accordingly. Let $||A_t||_{L_p}$be defined as the L_p−norm of matrix A_t, where $||A_t||_{L_p}\equiv [E||A_t||^p]^{1/p}$. The following definitions are to provide stability condition for the sequence of random matrices and henceforth, applied to the definition of stability for equation (9) for the recursive algorithms.

Definition 4.D.1: A random matrix (or vector) sequence $\{A_t\}_{t\geq 0}$ is called L_p−stable ($p > 0$) if $\sup_{t\geq 0} E||A_t||^p < \infty$, where $||A_t||$ represents the matrix norm of A_t.

Definition 4.D.2: A sequence of random matrices $\{A_t\}_{t\geq 0}$ (in equation (9)) is denoted as stably exciting of order p, $p \geq 1$, with parameter $\lambda \in [0,1)$, if it belongs to the following set that

$$S_p(\lambda) = \left\{ \{A_t\} : \| \prod_{j=i+1}^{t} (I - A_j) \|_{L_p} \leq M\lambda^{t-i}, \forall t \geq i, \forall i \geq 0, M > 0 \right\}. \qquad (4.2.10)$$

Definition 4.D.3: A scalar sequence $\{a_t\}_{t\geq 0}$is denoted as $S^o(\lambda)$-class if it belongs to the following set that

$$S^o(\lambda) = \left\{ \{a_t\}_{t\geq 0}, a_t \in [0,1], \quad E \prod_{j=i+1}^{t} (1 - a_j) \leq M\lambda^{t-i}, \forall t \geq i, \forall i \geq 0, \quad M > 0 \right\}. \qquad (4.2.11)$$

It follows from Guo (1994) that these definitions help to establish the stability of recursive algorithms applied to equation (1). The following condition is denoted as excitation condition for the model in equation (1). With this condition given, the above definitions for stability of recursive algorithms will be satisfied. Basically, since the system is of time-varying parameters, the explanatory variables must be of certain sufficient time-dependent variability to provide sufficient information for tracking. In other words, the included explanatory variables must provide sufficient fluctuations so that the time-varying coefficients can be updated sequentially. Intuitively, it appears that the applications of explanatory variables (for the empirical asset pricing models) for normal returns should provide some time-dependent variability over time. Otherwise, if the underlying parameters are time-varying, the recursive algorithms will not necessarily be stable, and the tracking errors can be unbounded.

Excitation condition (Guo (1994)): Let the regressors $\{\psi_t, \mathscr{F}\}$be an adapted sequence of random vector (that is, ψ_t is $\mathscr{F}_t$−measurable for all t's, where $\{\mathscr{F}\}$is a sequence of non-decreasing σ- algebras), and there exists an integer $h > 0$, such that $\{\lambda_t\} \in S^o(\lambda)$ for some $\lambda \in (0,1)$, where

$$\lambda_t \equiv \lambda_{\min}\left\{E\left[\frac{1}{1+h}\sum_{i=th+1}^{(t+1)h}\frac{\psi_i\psi_i'}{1+||\psi_i||^2}|\mathscr{F}\right]\right\}, \qquad (4.2.12)$$

and $\lambda_{\min}$ stands for the minimal eigenvalue of the random matrix, $E[x_{t+k}|\mathscr{F}]$represents the conditional expectation of random variable x_{t+k} based on information set $\mathscr{F}$.

In addition, Guo (1994) shows that if the vector of explanatory variables follows a ϕ−mixing process, then the necessary and sufficient conditions for the excitation condition in equation (12) to hold is that there exist an integer $h > 0$, such that

$$\inf_{t\geq 0}\lambda_{\min}\left\{\sum_{i=th+1}^{(t+1)h}E\left[\frac{\psi_t\psi_t'}{1+||\psi_t||^2}\right]\right\} > 0. \qquad (4.2.13)$$

Specifically, if the sequences of explanatory variables follow some mixing process, the excitation condition will require the covariance matrices of these explanatory variables do not become degenerated over time uniformly for any time interval with length h, $h > 0$.

For a similar yet more restrictive condition, Zhang et al. (1991) show it as $\forall m \geq 0$,

$$E\left(\sum_{t=m+1}^{m+h}\frac{\psi_t\psi_t'}{1+||\psi_t||^2}\right) \geq \frac{1}{\alpha_m}I,\ a.s., \qquad (4.2.14)$$

and $\{\alpha_m, \mathscr{F}_m\}$is an adapted non-negative sequence satisfying $\forall m \geq 0$,

$$\alpha_{m+1} \leq a\alpha_m + \eta_{m+1}, M_o \equiv E\alpha_o^{1+\delta} < \infty, \tag{4.2.15}$$

where $\{\eta_m, \mathscr{F}\}$is an adapted non-negative sequence such that

$$\sup_{m \geq 0} E\left[\eta_{m+1}^{1+\delta} | \mathscr{F}\right] \leq M, a.s., \tag{4.2.16}$$

where $a \in [0,1), 0 < \delta < \infty, 0 \leq M < \infty$, $a.s.$ stands for almost surely. A similar condition is applied in Chen and Guo (1991) and denote it as conditional richness condition. Given the settings, Guo (1994) shows that the tracking error bounds for the recursive algorithms. Intuitively, the "richness" condition such as Equation (4.2.12) is to ensure that the regressors are informative enough so that the stability of stochastic difference equation in equation (8) can be of asymptotic (exponential) stability. And hence, the recursive algorithms can be feasible to track down the time-varying coefficients. The following theorem is from Guo (1994) by re-writing three theorems into one and with additional sub-index i for a particular security of interest in empirical studies.

Theorem 4.1.1: (Guo (1994)): Given the time-varying coefficient model in Equation (4.2.1), and suppose the condition in Equation (4.2.12) is satisfied for all the following recursive algorithms. In addition, for each security i, $i = 1, 2, \cdots, n$, and for some $p \geq 1$, $\beta > 2$, and let

$$\sigma_p^{(i)} \equiv \sup_t ||\xi_{it} \log^\beta (e + \xi_{it})||_{L_p} < \infty \tag{4.2.17}$$

and

$$||\tilde{\theta}_{io}||_{L_{2p}} < \infty, \tag{4.2.18}$$

where $\xi_{it} = |\epsilon_{it}| + ||\triangle_{i,t+1}||$, $\theta_{it} = \theta_{it-1} + \Delta_{it}$, $\tilde{\theta}_{io} = \theta_{io} - \hat{\theta}_{io}$. Then for the Kalman filter in equation (3), the tracking error $\left\{\theta_{it} - \hat{\theta}_{it}\right\}_{t \geq 0}$ is L_p-stable and

$$\limsup_{t \to \infty} ||\theta_{it} - \hat{\theta}_{it}||_{L_p} \leq c\left[\sigma_p^{(i)} \log^{1+\beta/2}(e + \sigma_p^{(i)^{-1}})\right], \tag{4.2.19}$$

where c is a constant depending on $\{\psi_t\}_{t\geq 0}$, R, Q and p only.[2]

In addition, if least mean squares algorithm is applied, and if for some $p \geq 1$, $\beta > 1$, equations (17) and (18) hold, then $\left\{\theta_{it} - \hat{\theta}_{it}\right\}_{t\geq 0}$ is L_p-stable and

$$\limsup_{t\to\infty} ||\theta_{it} - \hat{\theta}_{it}||_{L_p} \leq c\left[\sigma_p^{(i)} \log(e + \sigma_p^{(i)^{-1}})\right], \qquad (4.2.20)$$

Furthermore, if recursive least squares algorithm is applied for tracking, and if for some $p \geq 1$

$$\sup_t \left(||\epsilon_{it}||_{L_{3p}} + ||\triangle_t||_{L_{3p}}\right) \leq \sigma_{3p}^{(i)}, \qquad (4.2.21)$$

$$\sup_t ||\psi_t||_{L_{6P}} < \infty, \qquad (4.2.22)$$

where the forgetting factor α satisfies that $\lambda^{[48hk(2h-1)p]^{-1}} < \alpha < 1$, k is the dimension of ψ_t, then there exists a constant c such that

$$\limsup_{t\to\infty} ||\theta_{it} - \hat{\theta}_{it}||_{L_p} \leq c\sigma_{3p}^{(i)}. \qquad (4.2.23)$$

In other words, Theorem 4.1.1 states that if given the boundedness of the moments in the fluctuations of time-varying parameters and error terms of regression, the tracking errors of these recursive algorithms will also be bounded asymptotically.[3] This implies that if the smoother the time-varying fluctuations of parameters, the smaller the tracking errors of these recursive algorithms may be. For further extended conditions, Guo and Ljung (1995a, 1995b) analyzes the exponential stability of some well-known algorithms such as extended Kalman filter, least mean squares, and recursive least squares with the forgetting factor in more general settings. The results simply identify that the tracking errors will be bounded if the conditions such as equation (4.2.12) is satisfied and specific conditions for each recursive algorithm are provided.

However, according to Grillenzoni (2000), when the dynamics for the time-varying parameters θ_t in equation (4.2.1) is unknown, the classical Kalman filter can not be applied for the estimation. Hence, if applying some functionals in approximating the time-varying coefficients, some more general algorithms such as extended recursive least squares (with forgetting factor) for the normal returns are needed. The studies of Hastie and Tibshirani (1993), Mei et al. (2001), Nielsen et al. (2000), Joensen et al. (2000), and Grillenzoni (2008) are some examples for the studies in this field. Following from Hastie and Tibshirani (1993), the time-varying coefficient regression for empirical asset pricing model can be expressed as

$$r_{it} = \beta_{io} + \sum_{j=1}^{k} \beta_{ij}(x_{jt})\psi_{jt} + \epsilon_{it}, \tag{4.2.24}$$

where $\{x_{jt}\}_{j=1,\cdots,k}$ are explanatory variables for the time-varying "beta's",$\{\psi_{jt}\}_{j=1,\cdots,k}$ are the nondiversifiable variables identified for the normal (expected) returns. For the specific identification, one may assume that these variables $\{x_{jt}\}_{j=1,\cdots,k}$ may be reduced to time variable t alone. To make the model in (4.2.24) more applicable, some apprroximations for the time-varying functions $\{\beta_{ij}(x_{jt})\}_{j=1,\cdots,k}$ such as polynomials, piece-wise constant, some smooth parametric functions or some spline functions can be applied. Specifically, the framework can be extended to include nonparametric setting. However, for simplicity, the "beta" functions are assumed to be of time variable t alone. The system in equation (4.2.24) is certainly more general than the random walk assumption for the transition of paramters. Neverthelss, to allow for applications of recursive estimation algorithms, more conditions are needed.

In particular, Grillenzoni (2008) introduces the time-varying coefficient model with slowly-varying parameters as

$$r_{it} = \psi_t' \beta_i(t) + \epsilon_{it}, \tag{4.2.25}$$

where the time-varying parameters $\beta_i(t)$ are considered as some smooth deterministic functions of time. [4]In addition to equation (4.2.25), Grillenzoni (2008) introduces additional assumptions for the system to allow for recursive estimation such as recursive least squares with forgetting factor:

Assumption 4.A.2. The noise process $\{\epsilon_{it}\}_{t=1,\cdots,}$ is idependent and identically distributed with mean zero and finite variance σ_i^2, where $\{\epsilon_{it}\}_{t=1,\cdots,}$ is independent of $\{\psi_t\}_{t=1,\cdots,}$, where $\psi_t = \begin{bmatrix} \psi_{1t} \\ \psi_{2t} \\ \vdots \\ \psi_{kt} \end{bmatrix}$.

Assumption 4.A.3. The explanatory variables $\left\{\psi_{jt}\right\}_{j=1,\cdots,k}$ are zero-mean, second-order stationary and ergodic with $E\left(\psi_{jt}^4\right) < \infty$, $j = 1,\cdots,k$ and $E[\psi_t\psi_t']$ is positive definite for all t's.

Assumption 4.A.4. The system is stable if $\left\{\psi_{jt}\right\}_{j=1,\cdots,k}$ incurs lagged dependent variables in the sense that the roots of polynomial for autoregression fall outside of unit circle.

Assumption 4.A.5. The $\beta_i(t)$ are deterministic functions and are differentiable up to the second order in continuous time, where the derivatives are rescaled as $\dot{\beta}_i(t) = \frac{\partial \beta_i(t)}{\partial t}/T$, $\ddot{\beta}_i(t) = \frac{\partial^2 \beta_i(t)}{\partial t^2}/T^2$.

Given the setting, a weighted least squares method can be applied for the estimation of equation (4.2.25)

$$\hat{\beta}_i(t) = \left(\sum_{h=1}^{t} \omega_{h,t}\psi_t\psi_t'\right)^{-1} \sum_{h=1}^{t} \omega_{h,t}\psi_t r_{it}, \tag{4.2.26}$$

where $0 < \omega_{h-1,t} \le \omega_{h,t} \le 1$, $h \ge 1$, $\omega_{h,t} = \lambda^{t-h}$, $\lambda \in (0,1)$. Equation (4.2.26) can be re-stated in a recursive implementation such as[5]

$$\begin{aligned} R_t &= \lambda R_{t-1} + \psi_t\psi_t' + (1-\lambda)I \\ \hat{\beta}_i(t) &= \hat{\beta}_i(t-1) + R_t^{-1}(r_{it} - \psi_t'\beta_i(t-1)), \end{aligned} \tag{4.2.27}$$

where R_t is a matrix of weighted squared regressors.

According to Grillenzoni (2008), the variance–covariance matrix of tracking errors can be shown as

$$
\begin{aligned}
&E\left[\left(\hat{\beta}_i(t)-\beta_i(t)\right)\left(\hat{\beta}_i(t)-\beta_i(t)\right)'\right] \\
&\quad=\left(\frac{1-\lambda}{1+\lambda}\right)E\left(\psi_t\psi_t'\right)^{-1}\sigma_i^2+O\left((1-\lambda)^{1.5}\right) \\
&\quad+\left(\frac{\lambda}{1-\lambda}\right)^2\left(\frac{\dot{\beta}_i(t)}{T}\frac{\dot{\beta}_i(t)'}{T}\right)+O\left(\frac{1}{T^2(1-\lambda)^{1.5}}\right) \\
&\quad+O\left(\frac{1}{T^3(1-\lambda)^3}\right).
\end{aligned}
\tag{4.2.28}
$$

Namely, the asymptotic variance–covariance matrix of tracking errors depends on the matrices $E\left(\psi_t\psi_t'\right)$ and $\dot{\beta}_i(t)\dot{\beta}_i(t)'$. That is to say, the higher the "richness" of the included explanatory variables, the smaller the norm of variance–covariance matrix tracking errors. In other words, the more informative the explanatory variables are, the less the tracking errors will be. In addition, from equation (4.2.28), it also shows that the less steeper the change of parameters is, the smaller the tracking errors naturally. Hence, this allows the situation when the parameters are time-varying mildly. What is important though, the determination of the exponential forgetting factor λ is also critical to this recursive algorithm for time-varying coefficients.

In summary, these algorithms are devised to track the possibly time-varying parameters associated with the empirical asset pricing models of normal (or expected) returns. Basically, the intent is to follow closely with the concurrent information for system-wise nondiversifiable explanatory variables. As such, the algorithms applied in the time-varying parameter models will reduce the pricing errors in specifying the normal (or expected) returns. Likewise, the abnormal returns $\{\epsilon_{it}\}_{t=1,2,\cdots}$ in equation (4.2.1) can be more closely observed and hence, event studies pursued thereby will be more robust in analyses.

Chapter 5

Time Will Tell! A Method with Occupation Time Statistics

Introduction

In this chapter, an alternative method is introduced to assess the impact of corporate events such as mergers and acquisitions on the firms. The method differs from the conventional event study tests in that, instead of testing the parameter changes over time, the durability of the parameter changes and persistence of the impacts is idscussed. In other words, the method considers the intensitity of the impacts from announcements or events may last over time. In terms of properties of stochastic processes, this persistence over time can be represented by the so-called occupation time (or sojourn time) of the underlying stochastic processes constructed by the statistics of interest.

The advantage of this approach is that the method is applicable particularly when the specifications for normal (or expected) returns with time-varying parameters are not sufficient to assess the impacts from events. Not surprisingly, it is well known as an example of common knowledge that the parameters fitted to the in-sample models for financial time series are subject to changes over time. However, overfitting the in-sample observations with ad hoc time-changing mechanisms is not necessarily ideal since the system of parameterization itself may also be subject to changes over time. Therefore, an alternative is to allow the changes of underlying parameters, and assess how persistent the impacts may last over time to consider the intensity of impacts. Two statistical methods are provided; one is with the Banach-space central limit theorem to perform the test for cross-sectional average of the firms' occupation time statistics, the other is for the firm-by-firm statistics of occupation time and application of false discovery rate for the multi-hypothesis testing across all firms.

5.1 Intuition

As stated in earlier chapters, the conventional CARs tests and the structural change tests are not entirely suitable for corproate event studies. The time-varying coefficients models are usually considered as common knowledge for most of financial time series. Either using the in-sample or out-of-sample statistics, it is difficult to accept that the parameters for the fitted models for financial time series are time-invariant. If certain structural changes are allowed, even if verifications of structural change can be monitored or tracked with sequential on-line approach, the results are not sufficient to justify that these changes are indeed related to events of interest when the test statistics are statistically significant. In fact, it is hardly conceivable that the parameters (fitted or assumed) are stable in the a-priori (predetermined) time frame before (or after) the possible (unknown) change point(s) in financial time series.

However, if time-varying parameters are considered as a "norm" for model specifications of financial time series, an alternative method that incorproates this "norm" for verifications on corporate-event issues is in need. In other words, if allowing the time-varying parameters (over various time frames) for financial time series, the conventional tests applied to corporate events such as mergers or acquisitions will not be appropriate. Namely, the conventional test based on CARs or CUSUMs are not capable of providing convincing statistics if time-varying coefficients are considered de facto for financial time series modelings. In addition, the alternative must also consider the stochastic properties of financial data particularly without arbitrary choices on the length of event windows.

Consequently, an alternative method is proposed to address these problems by focusing instead on the time spans of the dynamics of stock returns. In other words, the sample period can be chosen as the data set allows and there is no need to determine the arbitrarily chosen event windows for the tests. For instance, in Jeng et al. (2014), the sample period (of daily returns) spans from two years before the merger announcement to two years after the merger completion. However, no arbitrary event windows for the announcements are selected. A new test based on Takacs (1998) for occupation times of reflected Brownian motion is provided. The method introduces the statistics based on occupation time of reflected Brownian motion and applies the theoretical results of Takacs (1998) to analyze the durations of the impact from the new information.

A significant corporte finance event is redefined that an event is significant when the duration of the impact from the new information is significantly larger or lower than the occupation time of a reflected Brownian motion. The results in Jeng et al. (2014) shows that the durations of impact from the new information based on the abnormal returns to acquirers are significantly lower than the occupation times of a reflected Brownian motion. In other words, the market adjusts to the information faster than the speed of reflected Brownian motion assumed in the null hypothesis and resolves the uncertainty surrounding the mergers and acquisitions. In that article, comparative studies versus conventional CARs tests are also provided.

The alternative test strategy heretofore considers the occupation time (or sojourn time) of the diffusion processes. In particular, the counting measure for level crossings of absolute values of cumulative abnormal returns is taken into account. Intuitively speaking, if an event is essential, its impact will not simply be a split-second burst or spike on the sample path of asset returns. Instead, the fluctuations (including the spike) will last a while when the market works to digest the new information (or shock) and adjust itself before it settles. In other words, the market needs a little more time to resolve the new information from the event if the event or announcement is really to everyone's surprise. This time period to react will be longer if the information is more essential.

Alternatively, situations occur when the cumulative abnormal returns may be more volatile before and after the event than close around the event. More specifically, it may not necessarily lead to an increase of variance within the neighborhood of the event time in contrast with the other time periods. In such case, a conventional thought might consider that there is no essential event occurred. Yet, the conventional event study could be wrong. Instead of the no-event scenario, it is possible that the announcement (within the event window, if given) simply resolves the uncertainty for the market. In other words, an essential corporate event does not necessarily increase the fluctuations of the markets at announcement. Instead, it may dampen the fluctuations.

Hence, any method that aims to identify essential corporate events should account for both cases of "volatile jitterings" and " tamed-downs" when the new information hits. The new statistic therefore, considers the relative frequency of the statistics for absolute cumulative abnormal returns that exceed some thresholds over the entire sampled period. If the impact causes "volatile jitterings", the relative frequency will increase significantly.

Conversely, if the event impact causes " tamed-downs" of the market, the relative frequency will then decrease significantly. This statistic converges in distribution to the occupation time of a reflected Brownian motion provided that there is no essential event under the null hypothesis and the invariance principle for the cumulative abnormal returns holds when the sample period for each firm's abnormal returns is sufficiently large. Instead of testing for structural changes (multiple change points or epidemic change), the test statistic is to study the length or duration of the possible impacts caused by the shocks or corporate finance events. More specifically, the intuition is to take the time-varying coefficients (in market model or else) as a underlying phenomenon for stock returns and consider the duration of these impacts as level crossings of absolute cumulative abnormal returns as the statistic to verify the intensity of the events.

Given the formula provided by Takacs (1998) for distribution and moments of occupation time of Brownian motions, two methods can be applied to perform tests on the hypotheses of interest. First, the Banach-valued central limit theorem is applied for the cross-sectional average of occupation time functional across all firms. Under the null hypothesis, this test statistic will converge to a normally distributed random variable where the mean and variance of the occupation time of reflected Brownian motion are given with the formula of Takacs (1998). Second, the distribution of occupation time of reflected Brownian motion can be applied to perform test on each firm's abnormal returns individually across all firms. However, since there are many firms in the studies, the entire study may incur the mutiple hypotheses testings. This multiple hypothesis testing if with identical significance levels assumed in frequenist approach may incur overrejections across all hypotheses of interest. In other words, for performing the simultaneous inferences (across all selected firms over the null hypotheses of interest), the control over false discovery rate (FDR) will be applied to reduce the overrejection problems.

Notice that the test statistics allow for time-varying parameters in the fitted models on asset pricings or normal (or expected) returns by filtering the return series with recursive estimators. Hence, the time-varying parameters that are possibly subject to noise tradings and/or systematic feedbacks from the new information in the market are already filtered through the specifications of normal returns. In other words, for the well-known fact such as structural change (that is possibly of multiple change points) in financial time series modeling on stock returns, the filters applied will absorb these impacts first. The abnormal returns thereafter, will be "cleaner" in the sense that these time-varying coefficients from possibly systematic factorizations on stock returns are prefiltered before applying to the statistical tests on corporate events.

5.2 The Alternative Test Methodology

In order to assess the return process, a time-varying coefficient model is assumed for the empirical asset pricing models. Hence, the recursive estimation algorithms discussed in Chapter 4 can be applied here. Specifically, the following Assumption 5.A.1 is for the model specification of conditional expectations of excess returns in terms of presumed explanatory variables. Assumption 5.A.2 is the invariance principle for the cumulative sums of abnormal returns.

Assumption 5.A.1: Let the excess return r_{it} for each firm i and time t be a stochastic process in a complete probability space $(\Omega, \mathscr{F}, P)$ such that a k-factor model

$$r_{it} = E[r_{it}|\mathscr{F}_{it}] + \epsilon_{it} = \alpha_{it} + \sum_{j=1}^{k} \beta^{i}_{jt}\phi_{jt} + \epsilon_{it}, \tag{5.2.1}$$

holds for the excess returns where $\mathscr{F}_{it}$ represents the information filtration for firm i up to time t, α_{it} is the drift for the excess return, β^{i}_{jt} represents the possibly nonstochastic j-th time-varying coefficient with respect to nondiversifiable factor ϕ_{jt}, $j = 1, 2, \ldots, k$, $\epsilon_{it}$is the so-called abnormal return for firm i at time t.

Notice that Assumption 5.A.1 does not assume that the model fitted in equation (1) has already included all dynamic specifications. Hence, the abnormal returns $\{\epsilon_{it}\}_{i=1,2,..,t=1,2,...,}$may be subject to some short memory conditions such as mixing and heteroskedascity. However, the pricing variables $\left\{\phi_{jt}\right\}_{j=1,\ldots,k}$ should already be identified as proxies for nondiversifiable factors as discussed in Chapter 2.

Based on Assumption 5.A.1, a recursive estimation (such as recursive least squares method or stochastic approximation) in tracking can be applied to estimate the asset pricing models. In other words, the recursive estimation itself is a (dynamic) filter to retrieve possibly relevant impacts on the coefficients of asset-pricing model in reponse to the corporate information or event(s). Hence, the possible drifts or jumps (in the empirical asset pricing models) that occur during the sampled period can be tracked down so that the impacts from the time-varying coefficients are considered.

Specifically, the first step of methodology is to trace the possible market perception and impact when possible leakage of corporate information prevails in the market as the event day draws nearer. Thus, we filter the excess returns with adaptive (recursive) filter to accommodate possible impacts on the asset-pricing models before testing for corporate-finance events in using the abnormal returns. If a simpler model is desired, the conditional expectation $E[r_{it}|\mathscr{F}_{it}]$ can be assumed as a time-varying drift over time if $\beta_{jt} = 0$, $j = 1, 2, \ldots, k$, uniformly for all selected firms. More extensively, the conditional expectation $E[r_{it}|\mathscr{F}_{it}]$ can also be approximated by some functionals of explanatory variables in certain function spaces (with certain regularity conditions). For instance, the pricing kernel can also be extended to the generalized additive model where the pricing variables $\left\{\phi_{jt}\right\}_{j=1,\ldots,k}$ are functionals of some known explanatory variables. However, the non-diversifiability of these pricing kernels must be verified as claimed in Chapter 2 earlier.

The purpose to apply the time-varying coefficient models is that the models may approximate the adjustments of market's expectations where optimal filtering or recursive estimation can be applied for tracking of the normal (expected) returns. In addition, the structural changes of parameters in systematic components of asset pricing models where hypotheses of corporate finance events in CARs tests (and/or CUSUM tests) are included. In this case, the optimal filtering performs tracking for the normal returns with possible sturctural changes (due to various reasons such as noise trading, rumors, market overreaction, or even pricing errors, for instance) are filtered out and the error terms in equation (1) become "cleaner."

As stated in Chapter 2, the conditional expectation $E[r_{it}|\mathscr{F}_{it}]$ should be specified as comprensive as possible so that the systematic nondiversifiable components for stock returns are taken into account. However, given that the information set $\mathscr{F}_{it}$ applied may not be entirely exhaustive, there is a possibility that some pricing errors of asset returns may occur in $\{\epsilon_{it}\}_{t=1,2,\cdots}$ when using the empirical asset pricing models with limited information for some time periods. For instance, if there is a missing (unobservable) variable (say, η_t that should be included in the model), and $\{\phi_{jt}\}_{j=1,2,\cdots k}$are orthogonal to η_t, the correct specification should be stated as

$$\begin{aligned} r_{it} &= \alpha_{it} + \sum_{j=1}^{k} \beta^i_{jt}\phi_{jt} + \eta_{it} + \epsilon_{it}, \\ &= \ddot{\alpha}_{it} + \sum_{j=1}^{k} \beta^i_{jt}\phi_{jt} + \epsilon_{it}, \end{aligned} \tag{5.2.2}$$

where $\ddot{\alpha}_{it} = \alpha_{it} + \eta_{it}$, and η_{it} represents the pricing error due to omitted explanatory variables. This, in turn, shows that the pricing errors are simply absorbed in the time-varying drifts of assumed pricing models.[1] The recursive filterings based on equation (5.2.1) can still apply to approximation of normal (or expected) returns (as discussed in Chapter 4) and to obtain the abnormal returns accordingly. The tracking of recursive estimation will pick up the bias or pricing errors of asset pricing models through the intercept or drift.

Therefore, as long as the procedures for model search is thoroughly applied, the pricing errors from asset pricing models for normal (expected) returns should only be diversifiable. In other words, there is only a possibility of some negligible temporal drifts for the error terms of asset pricing models. Yet, even so, through the time-varying coefficient models in equation (1), these temporal drifts can be picked up by the recursive schemes of estimation. In other words, the abnormal returns are subject to much less pricing errors for the corporate event studies if recursive estimation is applied onto the time-varying coefficient models.

Brockett et al. (1999) apply a similar framework of time-varying beta market model and CUSUM test of standardized one-step-ahead forecast errors for the event study on California Proposition 103. As they claimed and discussed in Chapter 3 of this book, the applications of CUSUM test can avoid the arbitrary or subjective determination of width of event windows and estimation period. In particular, Brockett et al. (1999) apply the Kalman filter technique to estimate the fixed parameters of the model and update the prediction using new information available where no a priori assumption for event dates is applied. In addition, the GARCH(1,1) is also introduced to specify the possible fat-tailed distribution in error terms of regressions.

However, the CUSUM test (based on the residuals) is usually applied in the test for structural changes for parameters. Although their empirical result shows that the test statistics of the cumulative abnormal returns (based on the prediction errors) are different from zero, it implies that either (a) the constant term of the time-varying beta market model in their applications is also a time-changing drift (according to the discussions in Chapter 3), or (b) the assumption of AR(1) (autoregression of beta's with one lag) for the time-varying beta is possibly incorrect. In this case, the CUSUM test using the cumulative abnormal returns of prediction errors may not necessarily be more robust than conventional methods.

A good question to ask then is whether the result shows that the corporate event is indeed significant in influence or simply the time-varying beta market model is not adequate to pick up the time-chaing patterns of normal returns. On the other hand, if there are event-induced increases (or decreases) of excess volatility when the event causes extreme movements of the abnormal returns, GARCH(1,1) may not be sufficient to catch up. This also shows that allowing the time-varying parameters in model of normal returns must be accompained with an alternative methodology that is not based on the parametric or nonparametric tests on the location (or scale) of cumulative abnormal returns because these changes of location or scale are typical in financial time series even though there is no significant events. These changes of parameters, however, may or may not be event-oriented.

The following assumptions are for the test based on cumulative abnormal returns and provide the asymptotic functional where explicit formulas for the asymptotic distribution and moments are available. Assumption 5.A.2 is to establish the statistic based on cumulative abnormal returns under the null that there is no "essential" event in the data and consider the relative frequency of occurrence when cumulative abnormal returns exceeding some thresholds. Assumption 5.A.3 assumes that the occupation time functionals lie in the proper Banach space such as L^2-space equiped with the L^2-norm. Assumption 5.A.4. assumes strong law of large number for occupation time functionals in the Banach space. Assumption 5.A.5 provides the conditions that the de-meaned occupation time functionals across all selected firms will not explode asymptotically in their tails.

Assumption 5.A.2: Suppose under the null hypothesis that no impacts from event(s) are significant, and let the cumulative sums of abnormal returns $\{\epsilon_{it,}\}_{t=1,2,\ldots,}$ for each firm i follow the invariance principle such that as $T \to \infty$

$$\frac{1}{\sqrt{T}\sigma_{\epsilon_i}} \sum_{t=1}^{[\lambda T]} \epsilon_{it} \xrightarrow{d} B(\lambda), \tag{5.2.3}$$

where $0 \leq \lambda \leq 1$, and $\sigma^2_{\epsilon_i}$ represents the long run variance of $\{\epsilon_{it,}\}_{t=1,2,\ldots,}$, $B(\lambda)$ is a standard Brownian motion defined on interval $[0,1]$, the notation $\xrightarrow{d}$ stands for the convergence in distribution.[2]

Notice that Assumption 5.A.2 states that the cumulative sums of abnormal returns follow some invariance principles such that when properly weighted, they will converge to the Brownian motion in distribution. The assumption actually allow the abnormal returns to contain different serial dependence such as mixing condition or heteroscedasticity. Various invariance principles have been devised in the econometrics and statistics literature. Similar results that extend to almost sure invariance principles can be found in Eberlein (1986), and Wu (2007).

Many examples for the invariance principle of the error terms in regression have been applied in the related literature. For instance, under the null hypothesis that the drift and "beta" are not time-varying in the linear regression of Equation (5.2.1), Sen (1982) provides the invariance principle for recursive residuals from the linear regression even though the error terms $\{\epsilon_{it}\}_{i=1,2,\ldots,t=1,2,\ldots,}$ do not follow normal distribution. In the current context, Assumption 5.A.2 allows both serial dependence and heteroskedasticity given possible (dynamic) model misspecification.

Although Equation (5.2.3) incurs asymptotic normality where Brownian motion is assumed for the weak convergence of cumulative abnormal returns, the occupation time defined in the following for diffusion processes (including Brownian motion) is not normally distributed. In particular, the asymptotic convergence is for the entire sample period, and is not limited to the event windows or else. Specifically, with Assumption 5.A.2, the intent in the following is to consider whether the capital market under certain event(s) of interest has responded significantly more (or less) than the frequencies under the reflected Brownian motion. There is no need to consider if the statistics are obtained through the estimation period, event window, or post-event period since the invariance principle holds for the entire period of interest. The study is not based on the (asymptotic) normality assumption of stock returns or abnormal returns. Nor does the test verify the normality in asymptotic distribution of abnormal returns at all.

On the other hand, Assumption 5.A.2 is for the error terms from the time-varying-coefficient model of Equation (5.2.1) when the on-line recursive estimation will be applied to track and update the normal (or expected) returns sequentially. These error terms are different from the abnormal returns obtained from the conventional (say) market model (or else) estimated based on the presumed estimation periods prior to the assumed event window. More technically, the abnormal returns applied in conventional CARs tests are out-of-sample forecast errors of fitted regressions when using data from estimation period. No recursive updating of information is considered in the conventional approach. In addition, the asymptotic distribution (including normality) in the conventional CARs tests are obtained when number of observations in estimation period prior to event window grows sufficiently large. In Equation (5.2.3), the assumption does not assume that the estimation period must grow sufficiently large prior to the event date for the weak convergence to hold asymptotically. Instead, it assumes that the entire sample period T grows sufficiently large.

Although the conventional CARs tests may apply asymptotic normality in the cumulative sums of abnormal returns across the event window, the event window is assumed to have only finite number of dates or observations. The asymptotic arguments for normality in CARs tests is based on the large number of observations within the estimation period. Hence, Assumption 5.A.2 applies even to the occasions where no knowledge is feasible for the separation of estimation period and event window. Furthermore, since the on-line recursive estimation is applied, the estimation on normal (or expected) returns is updated with new observations. Therefore, there is no need to assume that the length of estimation period (prior to event window) must be sufficiently long. Nor is there the need to consider subjectively the number of days prior to and after the event dates to determine the event window even when the precise event date is known a priori.

However, an issue with applying structural change tests on event studies is that these parametric or distributional changes resolve gradually after certain time periods even if the event(s) are essential or significant. Thus, even with possible changes in parameters or distributions, the discussions on the essentiality of the event(s) should focus on "how long the impact may last." An event that is significant must have some occurrences of statistics (such as the frequencies of cumulative sums of abnormal returns that cross certain thresholds) persist over some time horizons.

In other words, similar to the assessments of earthquakes, the magnitudes of changes, or fluctuations of parameters provide only tentative estimates or partial information to the severity of the event. The measurement such as frequencies or durations of the impact (in time horizon) is a more valid method to describe the severity or essentiality of the event(s). This study provides an alternative method that allows for such a measurement. In addition, the method also encompasses the occurrence of permanent changes as special cases.

In the following, a general definition for occupation time of a nondegenerated real-valued diffusion process is provided. Furthermore, Assumption 5.A.2 serves only as an approximation for the weak convergence for the cumulative sums of abnormal returns. Given that, the following occupation time statistics are not to test whether these cumulative sums will converge to Brownian motion or not. Nor does it rely on the normality assumption for the increments of Browinan motion to perform the tests. Instead, the test statistics is to consider whether the hitting frequency or duration of impacts (for a certain level of threshold) is distinct from that of the reflected Brownian motion.

In other words, the purpose is to see if the duration of impacts from events of interest is significantly different from that of reflected Brownian motion. Although the cumulative sums of abnormal returns are similar to the statistics applied in the conventional CARs tests, the test statistics are not based on the asymptotic distribution of the CARs (or so-called cumulative abnormal returns). Instead, the statistical inference is based on the asymptotic distribution of occupation time of reflected Brownian motion.

In addition, the definition of occupation time for the diffusion process in the following may include various stochastic processes such as Lévy process where the jumps or discontinuities on the sample path may happen. In other words, the occupation time statistics can be extended to some jump processes in financial time series. For instance, in Fatalov (2009), the occupation time and its asymptotics are extended to the L^p–functionals of the Ornstein-Uhlenbeck process. Fitzsimmons and Getoor (1995) show the distribution for occupation times of Lévy bridges and excursions. These extensions can be applied to the following occupation time statistics when Assumption 5.A.2 is modified to different invariance principles for weak convergence to various diffusion processes of interest.

Definition 5.D.1: For a nondegenerated real-valued diffusion process with stationary increments, X_t, $t \in [0,1]$ and for a threshold h, the occupation time for the process X_t is defined as

$$\int_0^1 \delta(X_t > h)\mathrm{d}z, \tag{5.2.4}$$

where $\delta(x)$ is an indicator function for x to lie in a set A such that $\delta(x) = 1$ if $x \in A$, $A \subset R$, $A = \{x|x > h, x \in R^+\}$, and $\delta(x) = 0$, otherwise.

Notice that Definition 5.D.1 considers various nondegenerated real-valued diffusion processes that have the stationary increments. In other words, the processes such as Brownian meander, Brownian excursion, reflected Brownian motion, Brownian bridge, Lévy bridges, and excursions are included. Based on the above assumptions and Definition 5.D.1, we develop an alternative test statistic using the absolute values of the cumulative abnormal returns from the event study. The advantage of our test is that it does not require testing the differences in parameters or distributions across pre-event and post-event periods. Hence, it avoids the arbitrary choices of event window or pre-and post-event periods for statistical verification. Furthermore, based on the formula for asymptotic distribution and moments for the occupation time of reflected Brownian motion provided by Takacs (1998), a statistical test using the occupation times statistics across the entire sample period can be devised.

In other words, following Assumptions 5.A.1, and 5.A.2, and under the null hypothesis that the new (corporate finance) information has no essential or significant impact on the capital market, the underlying occupation time statistics defined in the following will converge to the occupation time of reflected Brownian motion in distribution asymptotically. Hence, the test that examines the essentiality of the corporate events can be formed as the test for significant difference between the occupation time statistics defined in the following Theorem 5.1.1 and the occupation time of reflected Brownian motion.

Theorem 5.1.1 is applicable to each individual firm selected. Thus, to perform the hypothesis testings for the impacts from event(s) of interest, two approaches are introduced hereon.

The first approach is to define the occupation time statistics as an element of a suitable function space and apply the central limit theorem for the function space to test the hypothesis on the average of occupation time statistics across all selected firms. The second one is based on the occupation time statistics of individual firms selected and apply corrections for over-rejection errors when multiple hypotheses are introduced. The first approach is more convenient since weak convergence of the average occupation time statistics (after normalization) will provide the well-known standardized normal distribution asymptotically under additional regularity conditions. In particular, if there is a pricing error that is hidden in the presumed abnormal returns where the recursive estimation did not catch up, the average statistics will reduce the errors in occupation time statistics under certain regularity conditions. The second approach is more informative since it takes all individual firm's responses into account and perform the tests over all firms by treating the null hypothesis for each firm as a separate hypothesis. And hence, the tests will require some adjustments based on multiple-hypothesis testings when sequential statistical inferences are applied.

Given the finite first and second moments of the test statistics $\psi_{i,T}(\omega, h)$ under the null hypothesis for occupation time of reflected Brownian motion , and using the Banach-valued central limit theorem on L^2 space, the test of the first approach is to verify the relative frequencies for exceedance of absolute values of CARs (for any given level of threshold) after normalization is significant or not with a standard normal distribution. More specifically, if under the null hypothesis that there is no significant impact from the corporate event(s), the cross-sectional average of occupation time from these abnormal returns when standardized using the mean and variance provided by Takacs (1998), should converge to a standardized normal random variable as the number of firms grows sufficiently large. Notice that the test statistic does not depend on the (asymptotic) normality in statistical distribution of abnormal returns. Nor is there any need to assume that there is no parameter change before the assumed event window or as the monitoring tests for structural changes. The following theorem shows the weak convergence of occupation time statistics for each firm's cumulative abnormal returns.

Theorem 5.2.1: Given Assumptions 5.A.1 and 5.A.2, and under the null hypothesis such that there is no essential impact from the new information disclosed in the event(s), let $\left\{\hat{\epsilon}_{it}\right\}_{i=1,2,\cdots}$ be the residuals of fitted regressions for asset returns at time t, and $\{\epsilon_{it}\}_{i=1,2,\ldots}$ follow the invariance principle of short-memory stochastic processes as stated in Assumption 5.A.2, and let $\tilde{\sigma}^2_{\epsilon_i} = \frac{1}{T}\sum_{t=1}^{T}\hat{\epsilon}^2_{it} + 2\sum_{j=1}^{T-1}k(\frac{j}{q})\hat{\gamma}_i(j)$ be the heteroskedasticity and autocorrelation consistent (HAC) estimate for the asymptotic variance $\sigma^2_{\epsilon_i}$ for asset i, $i = 1,2,\cdots,n$, where $\hat{\gamma}_i(j) = \frac{1}{T}\sum_{t=1}^{T-j}\hat{\epsilon}_{it}\hat{\epsilon}_{i,t+j}$, and $k(.)$ is the Bartlett-kernel function with bandwidth q, $q \to \infty$, $\frac{q}{T} \to 0$, let the occupation time statistic $\psi_{i,T}$ for all $\omega \in \Omega$, and for some $0 \leqq h < \infty$, $h \in R^+$ be defined as

$$\psi_{i,T}(\omega,h) = \frac{1}{T}\sum_{m=1}^{T}\delta\left[\frac{1}{\sqrt{T}\tilde{\sigma}_{\epsilon_i}}\left|\sum_{t=1}^{m}\hat{\epsilon}_{it}\right| > h\right], \tag{5.2.5}$$

where $\delta(x)$ is an indicator function for x to lie in a set A such that $\delta(x) = 1$ if $x \in A$, $A \subset R$, $A = \{x|x > h,\, x \in R^+\}$, and $\delta(x) = 0$, otherwise. Then, for any fixed n,

$$\psi_{i,T}(\omega,h) \xrightarrow{d} \int_0^1 \delta(\left|B(z)\right| > h)\mathrm{d}z \tag{5.2.6}$$

as $T \to \infty$, under the null hypothesis when no significantly essential events occur. In other words, when under the null hypothesis where there is no significantly essential event, the statistic $\psi_{i,T}$ will converge to an occupation time for the reflected Brownian motion (reflected at zero) $\left|B(z)\right|$ with some thresholds h, $h \in R^+$, where $B(z)$ is a standard Brownian motion defined on interval $[0,1]$, $\tilde{\sigma}^2_{\epsilon_i}$ is the estimate for the long run variance of$\{\epsilon_{it}\}_{t=1,2,\cdots,}$for firm i's abnormal returns.

Proof of Theorem 5.2.1:

Given Assumptions 5.A.1, 5.A.2, and according to Pötscher (2004), we only need to verify that the function $\delta\left[\frac{1}{\sqrt{T}\widetilde{\sigma}_{\epsilon i}}\left|\sum_{t=1}^{m}\epsilon_{it}\right| > h\right]$ is locally integrable. Pötscher (2004) defines that a real-valued function $f(x)$ is locally integrable if and only if for any $0 < K < \infty$,

$$\int_{-K}^{K} |f(x)|\mathrm{dx} < \infty.$$

Since the indicator function $\delta(x)$ is bounded with values as zero or one, we have $\sup_{x \in R^+} |\delta(x > h)| \leq 1$. Hence, for any level of threshold h, $0 < K < \infty$,

$$\begin{aligned}
&\int_{-K}^{K} \sup_{x \in R^+} |\delta(x > h)|\mathrm{dx} \\
&\leq \int_{-K}^{K} \mathrm{dx} \\
&\leq 2K < \infty.
\end{aligned}$$

Hence, the function as $\delta\left[\frac{1}{\sqrt{T}\widetilde{\sigma}_{\epsilon i}}\left|\sum_{t=1}^{m}\epsilon_{it}\right| > h\right]$ is also locally integrable. Following Theorem 2.1 in Pötscher (2004), for any fixed n, as $T \to \infty$, it is straightforward to verify that the weak convergence of $\psi_{i,T}(\omega, h)$ to an identically distributed occupation time of reflected Brownian motion such that $\psi_{i,T}(\omega, h) \xrightarrow{d} \int_0^1 \delta(\left|B(z)\right| > h)\mathrm{dz}$ for all $i = 1, 2, ..., n$.

Notice that, even though the above statistic is denoted as *occupation time statistic*, it is actually a statistic of normalized counting measure that considers the hitting frequency of the impacts from events when represented by the absolute values of cumulative abnormal returns. The cumulative abnormal returns are similar to the conventional approach in event studies of corporate finance. Yet, this cumulative sum of abnormal returns is not solely across event windows. Instead, the cumulative sums are obtained for the entire sample period without any a prior determination of estimation period, event windows and post-event period. In other words, the test statistic for the event(s) focuses on the frequency of the occurrence when the absolute values of cumulative abnormal returns may cross some thresholds. Intuitively, the higher the frequency, the more intensive the impacts are. However, under the Assumption 5.A.2 and assuming the events have no essential impact, this statistic will converge weakly to the occupation time of reflected Brownian motion asymptotically where explicit formulas for its moments and distribution are available. That is the reason why it is denoted as the occupation time statistic, so to speak.

Hence, if an event is essential, then its impacts are noticeable and therefore the frequency for such occurrence will either be significantly higher or significantly lower than the occupation time of reflected Brownian motion over the time horizon for any given threshold. On the other hand, if the event is not essential, this statistic will converge to the occupation time of a reflected Brownian motion. That is, for any given level of threshold, the occupation time statistic for the entire sample period will converge in distribution to the occupation times of reflected Brownian motion under the null hypothesis of no essential event. The formula of distribution and moments for the occupation times of a reflected Brownian motion versus different thresholds is provided by Takacs (1998).

As a result, we consider that the occupation time statistics $\psi_{i,T}(\omega, h)$ will converge weakly to $\int_0^1 \delta(\left|B(z)\right| > h)dz$ and are identically distributed for all firm i's under the null. Hence, to perform hypothesis testing across all selected firms, the average of these occupation time statistics is considered. Given certain regularity conditions and assume that the occupation time statistics $\left\{\psi_{i,T}(\omega, h)\right\}_{i=1,2,\ldots}$ belong to some proper Banach spaces, we may apply the Banach-valued central limit theorem and use their moments under the null hypothesis. Detailed mathematical arguments can be found in Ledoux and Talagrand (1991). In the following, two sections are provided for the applications of occupation time statistics. Section 5.2.(i) discusses the sample-average type test for occupation time statistics where asymptotic normality in Banach function space is applied. Section 5.2.(ii) applies the multiple hypothesis tests with occupation time statistics of each individual firm. Both methods can be applied to event studies in corporate finance.

5.2.1 (i) Hypothesis testing based on average occupation time statistics

Assumption 5.A.3: Let the occupation time statistics of abnormal returns for each firm i, $i = 1,2,...$, such as $\psi_{i,T}(\omega,h)$: $(\Omega,\mathscr{F},P) \to [0,1]$ for any threshold h, $0 \leqq h < \infty$, belong to a type-2 separable Banach L^2–space of all squared integrable real-valued Borel-measurable functions g, where $g : \Omega \to R$ with respect to the probability measure P and be equipped with the L^2–norm (denoted as $\| . \|$ heretofore).[3] A separable Banach L^2–space is of type 2 if and only if for a sequence of Rademacher random variables $\{\theta_i\}_{i=1,2,...,}$defined on $(\Omega,\mathscr{F},P)$ and for all finite sequences of $\{x_i\}_{i=1,2,...,}$, $x_i \in L^2$–separable Banach space, there exists a constant $C > 0$ such that

$$\| \sum_i^n \theta_i x_i \| \leq C \left(\sum_i^n \| x_i \|^2 \right)^{\frac{1}{2}}. \tag{5.2.7}$$

Assumption 5.A.4: Let the the occupation time functional $\{\psi_{i,T}(\omega,h)\}_{i=1,2,...,}$ follow the strong law of large numbers in the L^2–separable Banach space such that as $n \to \infty$,

$$\frac{1}{n}\sum_{i=1}^{n}(\psi_{i,T}(\omega,h) - E[\psi_{i,T}(\omega,h)]) \xrightarrow{a.s.} 0, \tag{5.2.8}$$

where $E[\psi_{i,T}(\omega,h)]$ denotes the Bochner integral of $\psi_{i,T}(\omega,h)$ with respect to P, the notation $\xrightarrow{a.s.}$ stands for almost surely convergence.

Assumption 5.A.5: The de-meaned occupation time functional $\{\xi_{i,T}(\omega,h)\}_{i=1,2,...,}$ satisfies the small ball criterion such that for each $\epsilon > 0$,

$$\lim_{\tau\to\infty} \tau^2 P\{||\xi_{i,T}(\omega,h)|| > \tau\} = 0, \tag{5.2.9}$$

$$\alpha(\epsilon) = \liminf_{n\to\infty} P\left\{||\sum_{i=1}^{n}\xi_{i,T}(\omega,h)/\sqrt{n}|| < \epsilon\right\} > 0, \tag{5.2.10}$$

where $\xi_{i,T}(\omega,h) = \psi_{i,T}(\omega,h) - E[\psi_{i,T}(\omega,h)]$.

These equations in Assumption 5.A.5 follow the small ball criterion in Ledoux and Talagrand (1991). Equation (5.2.7) specifies the tail condition such that the second-order moments of functionals $\{\xi_{i,T}(\omega,h)\}_{i=1,2,...,}$ will not be expanding too rapidly, while Equations (5.2.9) and (5.2.10) ensure$||\sum_{i=1}^{n}\xi_{i,T}(\omega,h)||$ is bounded in probability. These conditions are to ensure that the weak convergence of the partial sums onto a tight probability measure in Banach space. Given the above assumptions, we can establish the Banach-valued central limit theorem for the occupation time statistics $\{\psi_{i,T}(\omega,h)\}_{i=1,2,...,}$ under the null hypothesis of no significant impact from corporate finance event such as merger and acquisition. Many extensions on the central limit theorem from real-valued random variables to function spaces have been provided in mathematical and statistical literature such as in Hoffmann-Jorgensen and Pisier (1976), Araujo and Giné (1980), Ledoux and Talagrand (1991), and on weakly dependent Banach-valued random variables in Dehling (1983), and Ermakov and Ostrovskii (1986).[4]

Theorem 5.2.2 Given the results of Theorem 5.2.1 and Assumptions 5.A.3, 5.A.4, 5.A.5, if the null hypothesis holds such that no significant event exists, then for any given threshold $h \in R^{+}$, $0 \leqq h < \infty$, the partial-sum statistic for the occupation time functional will satistify the Banach-valued Central Limit Theorem such that for $S_n(h) = \frac{1}{n}\sum_{i=1}^{n}\left(\psi_{i,T}(\omega,h) - M_1(h)\right)$, as $n \to \infty$, $T \to \infty$, with T grows faster than the growth of n,[5]

$$Z(h) = \frac{\sqrt{n}}{\sigma(h)}\left(S_n(h)\right) \xrightarrow{d} \gamma^{\star}, \tag{5.2.11}$$

where $M_1(h)$, $\sigma(h)$ are the mean and standard deviation of the occupation time$\int_0^1 \delta(\left|B(z)\right| > h)\mathrm{dz}$ of reflected Brownian motion $\left|B(z)\right|$ for any threshold $h \in R$, and $0 \leqq h < \infty$, $\gamma^{*}$is a random variable in the separable Banach space L^2, with a standard normal distribution $N(0,1)$.

Proof of Theorem 5.2.2:
Given the independence among all the firms' mergers and acquisitons, it is easy to show that for any fixed n, and as $T \to$

∞, when all $\int_0^1 \delta(|B(z)| > h)\mathrm{d}z$ are of identical distribution, the finite partial sum as $\sum_{i=1}^{n} \psi_{i,T}(\omega,h)$ will also converge weakly such that

$$\sum_{i=1}^{n} \psi_{i,T}(\omega,h) \xrightarrow{d} \sum_{i=1}^{n} \int_0^1 \delta(|B(z)| > h)\mathrm{d}z.$$

According to Assumptions 5.A.3, 5.A.4, 5.A.5, and by Theorem 10.13 in Ledoux and Talagrand (1991), it is feasible to have the central limit theorem for $\{\xi_{i,T}(\omega,h)\}_{i=1,2,\ldots,}$ such that as $n \to \infty$,

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \xi_{i,T}(\omega,h) \xrightarrow{d} \gamma\,,$$

where γ is a random variable in the separable Banach space $L^2$with Gaussian measure. Under the null hypothesis that no significant events appear and Theorem 5.2.1 holds, we can have $E[\psi_{i,T}(\omega,h)] \xrightarrow{p} M_1(h)$ given that $\frac{1}{n}\sum_{i=1}^{n} \int_0^1 \delta(|B(z)| > h)dz \xrightarrow{p} M_1(h)$ when the weak law of large numbers holds as $n \to \infty$, for $\int_0^1 \delta(|B(z)| > h)dz$, $i = 1,2,\ldots,n$.

Now since all $\{\psi_i(\omega,h)\}_{i=1,2,\ldots,}$ converge to an identically distributed occupation time of reflected Brownian motion denoted as $\int_0^1 \delta(|B(z)| > h)\mathrm{d}z$, for all $i = 1,2,\ldots.$ as $T \to \infty$, by assuming that T grows faster than n, we may also apply the the $\sigma^2(h) = M_2(h) - M_1^2(h)$ for normalization where $M_2(h)$ is the second-order moment of $\int_0^1 \delta(|B(z)| > h)\mathrm{d}z$. Hence, as $n \to \infty$, we have

$$\frac{1}{\sqrt{n}\sigma(h)} \left(\sum_{i=1}^{n} (\psi_{i,T}(\omega,h) - M_1(h)) \right) \xrightarrow{d} \gamma^\star,$$

under the null hypothesis, where $\gamma^\star$ is a random variable in separable Banach space L^2 with a standardized Gaussian measure as $N(0,1)$. That is,

$$\frac{\sqrt{n}}{\sigma(h)} \left(\frac{1}{n} \sum_{i=1}^{n} (\psi_{i,T}(\omega,h) - M_1(h)) \right)$$

$$=\frac{\sqrt{n}}{\sigma(h)}\left(S_n(h)\right) \xrightarrow{d} \gamma^{\star}.$$

Notice that, although Theorem 5.2.2 shows normality in asymptotic distribution for the weak convergence of statistic $Z(h) = \frac{\sqrt{n}}{\sigma(h)}\left(S_n(h)\right)$ under the null hypothesis where $\sigma^2(h) = M_2(h) - M_1^2(h)$, the occupation time statistic $\psi_{i,T}(\omega, h) - M_1(h)$ does not assume normality in distribution for each selected firm i. In particular, even the occupation time of reflected Brownian motion $\int_0^1 \delta(\left|B(z)\right| > h)dz$ does not have a normal distribution either. Therefore, given that the distribution of $\psi_{i,T}(\omega, h) - M_1(h)$ is unknown and possibly nonnormal, the test statistic generalizes the conventional CARs tests and does not rely on the (asymptotic) normality of the (cumulative) abnormal returns. Also, given that the definition of the occupation time statistic covers the entire sample period, there is no need to distinguish (subjectively) the estimation period, event window, and the post-event period.

Table 5.1 taken from Jeng et al. (2014) shows that the average occupation time for testing if the mergers/acquisitions are essential events for corporate finance in the collected sample of 125 firms subject to successful mergers/acquisitions using cash tender offers from year 2000–2006. The data are collected from CRSP data bank. The average occupation time statistics after normalization show that the events are statistically significant under standard normal distribution.

More precisely, the statistics are mostly negative. This implies that the market actually resolves the uncertainty over the noisy information quickly. The occupation time statistics (for mergers and acquisisionts) on average, are less than the mean of occupation time of reflected Brownian motion. In this case, the average occupation time statistics can be applied to assess the significance of corporate-finance event(s) without using either the change of parameters or distributions. Namely, one only needs to calculate all occupation time statistics among the firms of events and the compatible firms of no events. If the events make the differences, these occupation time functionals will show that they are subject to differences across these two sets of firms. It is also noticeable that the average occupation time is

Table 5.1 The Average Occupation Time Statistics $Z(h)$

Threshold h	Average $\psi(\omega, h)$	$M_1(h)$	$M_2(h)$	$\sigma(h)$	$Z(h)$
0.001	0.9977	0.9984	0.9968	0.0012	−6.7840
0.005	0.9891	0.9921	0.9842	0.0059	−5.6781
0.010	0.9785	0.9841	0.9687	0.0115	−5.4491
0.015	0.9676	0.9763	0.9534	0.0169	−5.7378
0.020	0.9570	0.9685	0.9384	0.0219	−5.8602
0.025	0.9467	0.9607	0.9237	0.0267	−5.8922
0.030	0.9368	0.9530	0.9092	0.0311	−5.8359
0.035	0.9271	0.9453	0.8949	0.0352	−5.7974
0.040	0.9168	0.9377	0.8808	0.0390	−5.9953
0.045	0.9071	0.9301	0.8669	0.0424	−6.0576
0.050	0.8969	0.9226	0.8532	0.0455	−6.3100
0.055	0.8869	0.9151	0.8397	0.0481	−6.5590
0.060	0.8770	0.9076	0.8263	0.0502	−6.8075
0.065	0.8670	0.9002	0.8131	0.0519	−7.1503
0.070	0.8574	0.8928	0.8000	0.0531	−7.4479
0.075	0.8482	0.8855	0.7870	0.0537	−7.7547
0.080	0.8393	0.8782	0.7741	0.0537	−8.1046
0.085	0.8295	0.8709	0.7613	0.0529	−8.7556
0.090	0.8210	0.8637	0.7486	0.0513	−9.3149
0.095	0.8115	0.8565	0.7360	0.0486	−10.3638
0.100	0.8026	0.8494	0.7234	0.0445	−11.7496
0.105	0.7935	0.8422	0.7109	0.0386	−14.1333
0.110	0.7841	0.8352	0.6983	0.0294	−19.4286

a decreasing function of thresholds. This leads to the consideration that for each firm with event or without, the occupation time statistic is in general, also a decreasing function of the levels of thresholds.

Moreover, further extensions for the statistics based on (say) Brownian motion can be devised. For instance, if the intensity of cumulative abnormal returns is of concerned, one can consider the statistics as

$$\pi_i(\omega) = \frac{1}{T}\sum_{m=1}^{T} \delta\left\{\left[\frac{1}{\sqrt{T}\tilde{\sigma}_{\epsilon_i}}|\sum_{t=1}^{m}\hat{\epsilon}_{it} - m\bar{\hat{\epsilon}}_i|\right] > h\right\} \xrightarrow{d} \int_0^1 \delta(|\overline{B}(z)| > h)\mathrm{d}z,$$

where $\bar{\hat{\epsilon}}_i = \frac{1}{T}\sum_{t=1}^{T}\hat{\epsilon}_{it}$ and $\overline{B}(z)$ stands for the standard Brownian bridge, as $T \to \infty$. Or,

$$\pi_i^{\star}(\omega) = \frac{1}{T}\sum_{m=1}^{T}\left[\frac{1}{\sqrt{T}\tilde{\sigma}_{\epsilon_i}}|\sum_{t=1}^{m}\hat{\epsilon}_{it}|\right] \xrightarrow{d} \int_0^1 |B(z)|\mathrm{d}z,$$

where $|B(z)|$ is the reflected Brownian motion. In the first case, it shows the occupation time statistic for mean deviations of abnormal returns, and its convergence toward the occupation time of Brownian bridge. And the second statistics will coverge in distribution to the areas underneath the sample path of reflected Brownian motion. Either one can also be applied for assessments on intensity of events. Loosely speaking, for the first extension, the occupation time statistic is to consider the duration of mean deviations of these cumulative abnormal returns. The second extension however, is provided by the condition that reflected Brownian motion can be shown that $|B(z)| \stackrel{d}{=} \sup_{0 \leq t \leq z} B(t) - B(z),\ 0 \leq z \leq 1$. The intensity of cumulative abnormal returns (asymptotically as $T \to \infty$) can be approximated (under the null hypothesis) by the integral of the difference between running maximum and the current value of Brownian motion, which is similar to the drawdown process in equation (13) in the following. In other words, the area beneath the sample path of the drawdown process can be used as a represenative for the intensity of cumulative abnormal returns over the entire sample period.

In essence, the test statistic is not to verify (even asymptotically) the fittedness of reflected Brownian motion for the absolute values of cumulative abnormal returns. Instead, it is for the frequency of hitting the thresholds from the absolute values of cumulative sums of abnormal returns. Besides, the construction of statistics covers the entire sample period that no distinction of estimation period, event window and post-event window is needed. In addition, the above statistics $\pi_i(\omega)$, $\pi_i^\star(\omega)$, and $\psi_{i,T}(\omega, h)$ can all be applied to consider monitoring for significant events when on-line quickest detection is needed

Other time-based statistics for stochastic processes

There are a few similar time-based statistics for diffusion processes such as Brownian motion, Brownian motion with drift, and many others. For instance, Tanré and Vallois (2006) provide the distribution for the range of Brownian motion with drift where the range is defined as

$$R(t) = \sup_{0 \leq u \leq t} B_\delta(u) - \inf_{0 \leq u \leq t} B_\delta(u), \tag{5.2.12}$$

where $B_\delta(t) = B(t) + \delta t$, $\{B(t)\}_{t\geq 0}$ is the standard Brownian motion. Although the range statistic is interesting, it only shows the extent of fluctuations for the underlying process such as Brownian motion with drift. In other words, the statistic may be applied to analyze the responsiveness of the underlying diffusion process when some shocks may occur. In applications to the financial time series, this statistic may serve as a proxy for volatility over time. However, it does not show how persistent the duration of the impacts. Namely, one may find out that an event that has a strong but short-lived impact will have a higher range than the other long-lived event that has less imminent impact with yet more persistent duration. In that case, the event that has a stronger but short-lived impact may resemble an event-oriented noise. While fervent on the impact, the noise quickly resolves itself to be negligible. On the contrary, the event that has less powerful or immediate impact yet with more persistent duration may become a more noticeable issue since the consequence of the event may have more potential influence for the system.

Zhang and Hadjiliadis (2012) introduce the drawdowns and the speed of market crash to assess the extent of market crash. The drawdown process for a one-dimensional diffusion process on an interval (a, b) is set as

$$D_t = \overline{X_t} - X_t, \tag{5.2.13}$$

where $\overline{X_t} = \sup_{s\in[0,t]} X_s$ as the running maximum of X_t, $t \geq 0$. In addition, for a given level K, the first passage time of D_t to hit the level K (that is, the first time when drawdown process hits the assumed level) is defined as

$$\tau_K^D = \inf\{t \geq 0 | D_t = K\}. \tag{5.2.14}$$

Also, the last visit time of maximum of the diffusion process X_t before τ_K^D is defined as

$$\rho = \sup\{t \in [0, \tau_K^D] | \overline{X_t} = X_t\}. \tag{5.2.15}$$

Zhang and Hadjiliadis (2012) thereby define the speed of market crash as

$$S = \tau_K^D - \rho. \tag{5.2.16}$$

Although the statistics are related with sample path and time, the impacts from events are not known to be positive or negative (or both) in advance. The same logic may apply to the drawup process also. On the other hand, if the impacts are negative, the shorter the time span between first passage time of drawdown process D_t for the level K and the last visit time of maximum of process X_t does imply that the impact lasts only a short period of time. This implies how fast the negative news causes the stock returns to fall. And hence, if the information or event is known as negative, the statistic in Equation (5.2.16) can be applied to see how soon the market responds to the news. The diffulty, however, the analysis is based on the diffusion processes assumed. Further extensions in constructing suitable sampled statistics and their (asymptotic) statistical properties are in need.

5.2.2 (ii) Multiple hypothesis tests with statistics of individual firms

In this section, the hypothesis tests are applied to occupation time statistics of each selected firm individually. An advantage of the multiple hypothesis testing with individual test statistics is that one will know which particular firm selected will accept or reject the null hypothesis explicitly. This approach, in turn, helps to identify how the firms may respond differently to the similar event(s) individually. In terms of multiple hypothesis testing, the idea is to perform tests on each and every firm selected (for instance) such as for some arbitrarily small $\tilde{a} > 0$,

$$\begin{aligned} H_{01} &: |\psi_{1,T}(\omega,h) - \psi^{\star}(h)| \le \tilde{a}, \quad H_{a1} :\text{otherwise}, \\ H_{02} &: |\psi_{2,T}(\omega,h) - \psi^{\star}(h)| \le \tilde{a}, \quad H_{a2} :\text{otherwise}, \\ &\vdots \\ H_{0n} &: |\psi_{n,T}(\omega,h) - \psi^{\star}(h)| \le \tilde{a}, \quad H_{an} :\text{otherwise}. \end{aligned} \tag{5.2.17}$$

where $\psi^{\star}(h)$ is defined in Equation (5.2.18).

Hence, the distribution function for the occupation time of reflected Brownian motion will be applied. Notice that the distribution is not normal. In particular, the moments of the statistics also depend on the thresholds prespecified. In other words, the critical values of the statistics depend on the thresholds (the h's) and the significance levels assumed.

Therefore, it is not straightfoward to obtain the critical values (for statistics) for any given level of significance. Nevertheless, for each given threshold h, the p-values of the statistics can be obtained for all selected firms. The statistical tests can be performed to see if these p-values are smaller than the assumed level of significance or not. If so, the test statistic will reject the null hypothesis that the event is not influential. Takács (1998) developed explicit formulas for the distribution function, density function, and the moments of the occupation time for the reflected Brownian motion. Specifically, let the threshold be denoted as $h \geq 0$, the occupation time (or sojourn time) of a reflected Brownian motion in the time interval $(0,1)$ is given as

$$\psi^{\star}(h) = \int_0^1 \delta(|B(z)| > h)\mathrm{d}z, \tag{5.2.18}$$

where $\{B(z),\ z \geq 0\}$ is the standard Brownian motion. For the distribution function $G_h(x)$ of the occupation time $\psi^{\star}(h)$, $G_h(x) = Pr\{\psi^{\star}(h) \leq x\}$, Takács (1998) shows that

$$G_h(x) = 2F_h(x) - 1 + 2\sum_{k=2}^{\infty}\sum_{j=2}^{k} \frac{(-1)^j j!}{(k+j-1)!}\begin{pmatrix} k-2 \\ j-2 \end{pmatrix} \frac{d^{k-1}x^{k-1}[1 - F_{(2j-1)h}(x)]}{dx^{k-1}} \tag{5.2.19}$$

where the notation $\frac{d^{k-1}}{dx^{k-1}}(.)$ represents the $(k-1)$th derivative of the function of interest, and

$$F_h(x) = 1 - \frac{1}{\pi}\int_0^{1-x} \frac{\mathrm{e}^{-h^2/2u}}{\sqrt{u(1-u)}}\mathrm{d}u \tag{5.2.20}$$

for $0 \leq x \leq 1$, and $h \geq 0$; and $F_h(0) = 2\Phi(h) - 1$, for $h \geq 0$, where

$$\Phi(h) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{h} \mathrm{e}^{-u^2/2}\mathrm{d}u. \tag{5.2.21}$$

Therefore, for any given threshold level h, and for any occupation time statistic $\psi_{i,T}(\omega, h)$, it is feasible to find out the

p-value of the statistic based on the distribution function $G_h(x)$ – under the null hypothesis such that the corporate finance event of interest is not influential, and under Assumptions 5.A.1 and 5.A.2. It appears that the distribution function for occupation time (of reflectd Brownian motion) is not a normal distribution even though the Brownian motion is with normally distributed increments. Hence, the p-value of one-sided test for the occupation time statistic (of selected firm i) under the null hypothesis is shown as

$$p_i = 1 - G_h(\psi_{i,T}(\omega,h)) = 1 - Pr\{\psi^{\star}(h) \leq \psi_{i,T}(\omega,h)\}, \tag{5.2.22}$$

where p_i represents the p-value of the occupation time statistic of selected firm i. The hypothesis testing then is to compare the p_i with the prespecified significance level α and reject the null if $p_i < \alpha$. The advantage of this approach is that the specific differences (for the strength of impact from event(s)) in the selected individual firm's occupation time statistics are taken into account.

However, since this involves multiple hypotheses across all different firms, the statistical inferences must consider the possible overrejection rate when the statistics are applied repetitively. Two concepts are usually discussed: family-wise error rate (FWER) and false discovery rate (FDR). The definition of family-wise error rate is defined as FWER = Pr{reject any true H_{oi}}, for $i = 1,2,\cdots,n$.

Hence, according to Efron (2010), for selected n firms' occupation time statistics and if the FWER control is applied, the procedure is to have these p-values $\left(p_1, p_2, \cdots, p_n\right)$ that produce the list of accepted and rejected hypotheses satisfying the constraint

$$\text{FWER} \leq \alpha, \tag{5.2.23}$$

for any preselected value α. For instance, Bonferroni's procedure for controlling the family-wise error rate is: reject the null hypotheses for which

$$p_i \leq \frac{\alpha}{n}, \tag{5.2.24}$$

where n is the number of null hypotheses of interest. In other words, the procedure actually provides the adjusted p-value as

$$\tilde{p}_i = \min[np_i, 1], \tag{5.2.25}$$

where $\min[x, 1]$ stands for the lower value of x and 1, and hence, reject the hypothesis H_{0i} if the adjusted p-value $\tilde{p}_i \leq \alpha$.

According to Benjamini and Hochberg (1995), the family-wise error rate control of multiple hypothesis testing is with lots of faults. One particular disadvantage of the FWER control is that the procedure may lead to loss of power in hypothesis testing. Hence, they provide a different control scheme using false discovery rate. The idea of false discover rate is better explained in the following table.

Actual	**Test Accepted**	**Test Rejected**	**Total**
True null hypotheses	U	V	m_0
Non-true null hypotheses	T	S	$n - m_0$
Total	$n - R$	R	n

Let the total number of hypotheses be n. Suppose there are actually m_0 true null hypotheses and $n - m_0$ non-true null hypotheses within the population. These numbers are unknown. Now let the statistical tests show that there are U accepted and V rejected under the actual information that the null hypotheses are true. On the other hand, there are T accepted and S rejected when the null hyptheses are not true. The false discovery proportion then, is defined as $\text{FDP} = \frac{V}{R}$, and hence, the false discovery rate (FDR) is defined as

$$\text{FDR} = E[\frac{V}{R}]. \tag{5.2.26}$$

In other words, the false discovery rate is the expected value of the proportion of *rejected* hypotheses that are erroneously rejected when the null hypotheses are true. Notice that this false discovery propotion FDP is not equal to the proportion of true null hypotheses that are erroneously rejected, which is $\frac{V}{m_0}$ instead. Hence, the procedure using FDR is to control the erroneous rejection rate among the rejected hypotheses, not all true null hypotheses. The Benjamini and Hochberg (1995) control procedure is shown as follows. Suppose the decision rule for each hypothesis H_{0i} provides the p-value for each $i = 1, 2, \cdots, n$ such that p_i has a uniform distribution when H_{0i} is true

$$H_{0i} : p_i \quad \sim U(0,1). \tag{5.2.27}$$

where $U(0,1)$ denotes as the uniform distribution in the interval $(0,1)$. Order the p-values by

$$p_{(1)} \leq p_{(2)} \leq \cdots \leq p_{(n)}. \tag{5.2.28}$$

Now, for any given (desired) level of significance α, $\alpha \in (0,1)$, let $i_{\max}$ be the larget index i such that

$$p_i \leq \frac{i}{n}\alpha, \tag{5.2.29}$$

then, reject the null hypotheses H_{oi} for any $i \leq i_{\max}$, and accept H_{0i} for the others.

According to Benjamini and Hochberg (1995), if the p-values for the correct null hypotheses are independent of each other, their control procedure will control the FDR such that

$$E[\text{FDP}] = \text{FDR} = \pi_o \alpha \leq \alpha,. \tag{5.2.30}$$

where $\pi_o = \frac{m}{n - m_o}$.

In other words, this actually implies that the procedure will ensure the multiple hypothesis testing will have the error rate of rejection less than or equal to the original (desired) level of significance.[6] For applications of false discovery rate control in occupation time statistics, it is feasible to apply Equation (5.2.19) to obtain the p-values of null hypotheses across all firms of interest, and re-order these hypotheses together with their p-values. Choose the highest indexed hypothesis such that its p-values satisfies Equation (5.2.29). Then, reject all hypotheses with lower indices while accept the null for those with higher indices, instead.

The advantage of the tests is that individual differences from various hypotheses are considered. It is probable to see that the impacts from the same event may differ across selected firms. Firms that have the same announcement(s) or events may not necessarily have the identical nature of their business. For instance, the same type of mergers and acquisitions from different firms may not result from idential operational characteristics, managerial concerns, or financial status. Although average occupation time statistics across selected firms are simpler to obtain, there exists possible loss of information where individual differences are averaged out. The disadvantage of multiple hypothesis testing, however, is that these repetitive assessments for the p-values of applied statistics may become demanding if the sample size is relatively large.

5.2.3 Further Extensions for the Occupation-time Statistics

Similar to the tests for comparing distributions, a new occupation-time-based test for event studies can be devised. The approach is that, instead of purposedly selecting all firms of events of interest, all firms with events and without events are all sampled into two groups. Intuitively, if the event is essential, the occupation-time statistics of the firm with events and without events will differ from each other within the same time period sampled. A particular advantage is that if some econometric models are applied, no selectivity bias (for event studies) will be introduced. This gives us an alternative that one can choose to select some benchmark firms of interest and compare their occupation-time statistics versus the selected firms of events of interest. Technically, as one stacks up the abnormal returns for the firms of interest with the other firms together, and when the occupation-time statistic of each firm can be treated as a functional (of thresholds h within a compact support) in a proper function space, the test for essential event(s) becomes the test for differences in functional data. The advantage is that the test is no longer for the structural change in parameters of regression models alone. Instead, it is for the differences of functionals as occupation time statistics. The studying interest then, is for whether these functionals (of thresholds) differ significantly among the event-related and non-event-related firms.

If the event is essential, the occupation-time functionals for the event-related firms will differ significantly from those of the non-event-related firms. Notice that the occupation time statisitcs $\{\psi_{i,T}(\omega,h)\}_{i=1,2,\cdots,n}$ defined above are some random (decreasing) functions of threshold h. Given certain regularity conditions of these functionals in some function spaces, the tests for event studies can be performed as comparing the differences among these functionals $\{\psi_{i,T}(\omega,h)\}_{i=1,2,\cdots,n}$ where the sample size n may contain some n_1 firms of no events in the first block, and $n-n_1$ firms of events in the second block. If the events are essential, then there exists a significant difference over these occupation time functionals across all firms. This setting then, will be similar to the hypothesis testing for differences of functional data especially for the tests of mean changes. In a sense, this is similar to a structural change test for functional data where the change point is known in advance. For the threshold $h \in [0,1]$, let the occupation time functionals for the sampled firms be donted as

$$\psi_{i,T}(\omega,h) = \mu_1(h)1_{\{1\leq i\leq n_1\}} + \mu_2(h)1_{\{n_1<i\leq n\}} + \Psi_i(h),\ E[\Psi_i(h)] = 0. \tag{5.2.31}$$

The null hypothesis for the occupation time statistics on event studies will be shown as

$$H_0 : E[\psi_{i,T}(\omega,h)] = \mu_1(h), \forall i = 1,2,\cdots,n, \tag{5.2.32}$$

and the alternative hypothesis is shown as

$$H_A : \begin{cases} E[\psi_{i,T}(\omega,h)] = \mu_1(h), & 1\leq i\leq n_1, \\ E[\psi_{i,T}(\omega,h)] = \mu_2(h), & n_1 < i \leq n. \end{cases} \tag{5.2.33}$$

Alternatively, it can also be restated in the generalized two-sample test such as

$$H_0 : \mathscr{L}(\psi_{1,T}(\omega,h)) = \mathscr{L}(\psi_{2,T}(\omega,h)) = \cdots = \mathscr{L}(\psi_{n,T}(\omega,h)), \tag{5.2.34}$$

$$H_A : \mathscr{L}(\psi_{1,T}(\omega,h)) = \mathscr{L}(\psi_{2,T}(\omega,h)) = \cdots \neq \mathscr{L}(\psi_{n_1,T}(\omega,h)) = \cdots = \mathscr{L}(\psi_{n,T}(\omega,h)), \tag{5.2.35}$$

where the function $\mathscr{L}(.)$ stands for the distribution function. For example, a two-sample test such as Ferger (2000) for functional data can be applied to occupation time statistics $\{\psi_{i,T}(\omega,h)\}_{i=1,2,\cdots,n}$. The idea is to apply the difference between estimated change point versus the assumed change point which is given by the samples. For instance, if the setting is given in Equation (5.2.31), let $n_1 = [\theta n]$, $0 \leq \theta \leq 1$, $[x]$ represents the largest integer that is less than x. Since n_1 is known already, it is straightforward to obtain the parameter θ, accordingly. Hence, intuitively, if the estimated change point of these stacked up $\left\{\psi_{i,T}(\omega,h)\right\}_{i=1,2,\cdots,n}$ differs little from the θ, the null hypothesis of no change will be rejected. Ferger (2000) introduces a kernel-based estimator for θ as

$$\theta_n^{(+)} = \frac{1}{n} \underset{1\leq k\leq n-1}{\text{argmax}}\ w\left(\frac{k}{n}\right) \sum_{i=k+1}^{n} \sum_{j=1}^{k} \mathcal{K}\left(\psi_{i,T}(\omega,h), \psi_{j,T}(\omega,h)\right), \tag{5.2.36}$$

where the weight function is chosen as $w(t) = t^{-a}(1-t)^{-b}$, $0 < t < 1$, $0 \leq a, b < \frac{1}{2}$. The kernel function is set as antisymmetric such that $\mathcal{K}(x,y) = -\mathcal{K}(y,x)$. Then, under the other regularity conditions,

$$\theta_n^{(+)} - \theta = O(n^{-1}),\ n \to \infty. \tag{5.2.37}$$

Hence, the test statistic can be formed as

$$\Phi_n^{(+)} = 1_{\left\{|\theta_n^{(+)} - \frac{n_1}{n}| \leq c^{(+)}\right\}}, \tag{5.2.38}$$

where $1(.)$ is an indicator function. Or more specifically, reject the null hypothesis (that is, $\Phi_n^{(+)} = 0$) if the absolute difference between $\theta_n^{(+)}$ and $\frac{n_1}{n}$ is greater than the critical value $c^{(+)}$. Ferger (2000) in showing the weak convergence such as $\theta_n^{(+)} \xrightarrow{d} \tau_w^{(+)}$, where $\tau_w^{(+)} = \underset{0<t<1}{\text{argmax}}\, w(t)\overline{B}(t)$, $\overline{B}(t)$ represents the standard Brownian bridge that provides the critical values $c_\alpha^{(+)}$ as the $\alpha-$quantile of the process $|\tau_w^{(+)} - \theta|$.

The extensions show that many applications of occupation time statistics can be provided to analyze the event studies in corporate finance. Given that the underlying financial time series are subject to possible time-varying parameters, these assessments of level crossings over the time period become some functionals (of certain function spaces) instead. In contrast to relying on changes in parameters to verify the impact of event, justification such as considering differences in shapes, changes in distributions of functional data seems more promising to analyze the event studies in corporate finance. On the other hand, the monitoring test or on-line detection test can also be provided by observing the change of shapes in these occupation time functionals.

In particular, the monitoring (on-line) test for occupation time functionals $\psi_{i,T}(\omega, h)$ is similar to the test using empirical survival functions of h. For instance, given a particular firm i, it is feasible to consider that if the training period is (say)$[0, T]$, where$T \to \infty$, define the test statistic $Q_T(k)$ as

$$Q_T(k) = \left(\sup_{0 \le h \le 1} |\psi_{i,T+k}(\omega, h) - \int_0^1 \delta(\left|\varpi(z)\right| > h)\mathrm{dz}| \right),$$

where $\psi_{i,T+k}(\omega, h) = \frac{1}{T+k} \sum_{m=1}^{T+k} \delta\left[\frac{1}{(\sqrt{T+k})\tilde{\sigma}_{\epsilon_i}} \left|\sum_{t=1}^{m} \hat{\epsilon}_{it}\right| > h \right]$, $k \ge 1$, $\varpi(z)$ is the diffusion process under the null hypothesis, such as standard Brownian motion.

In other words, the statistic $Q_T(k)$ can be applied to see if there is a significant difference between the occupation time statistic (with additional k observations) and the occupation time of reflected diffusion process under the null hypothesis as additional observations flow in. The stopping time for the $Q_T(k)$ statistics is defined as $k^* = \inf\left(k \ge 1, |Q_T(k) > c^\star\right)$, where $c^\star$is the critical value. The advantage is that it does not assume the coefficient(s) of the underlying processes $\left\{\hat{\epsilon}_{it}\right\}_{t=1,2,\cdots,T,\cdots}$are stable before the starting time k for monitoring. Instead, it only assumed that the occupation time statistic under the null does not differ much from that of reflected diffusion process (such as reflected Brownian motion) prior to the starting time k. These are left for elaborations in further research.

Epilogue

Event studies in corporate finance are so critical for verification on the capital market efficiency and the speed of adjustments in stock returns. The contents of this book merely touch the surface of this gigantic territory of intellectual expertise. Although the issues in event studies of corporate finance are not as spectacular as the space wonderment of galaxies, their varieties and depths are enormous. For the purpose of continuing research, certain extended works are required. For instance, the model search procedures can be extended with further works in statistics for long dependence. Given that the concept of long (or strong) dependence in stochastic processes (either for time series or cross-sectional observations) is more extensive than the specification of unit root(s), developments of robust statistics for long dependence is in need to elaborate the model selection (or variable selection) in empirical asset pricing models. Various definitions of strong dependence can be introduced to provide better verifications on the essential feature of nondiversifiable pricing kernels that describe the benchmark normal (or expected) returns of risky securities.

Separation between normal (or expected) returns and abnormal returns should be clearly defined (at least, theoretically) to make event studies in corporate finance with less ambiguities. Further works for broader definitions of diversification will also improve the rigors and thoroughness for the asset-pricing kernels that clearly define this separation of returns. In fact, with more general definitions on diversification applied, the property as nondiversifiable systematic components of stock returns will become even apparent. This, in turn, may provide another vehicle to prevent the joint hypothesis testing problems in event studies of corporate finance.

For the developments of occupation time statistics, many extensions can be provided in further elaborations. Currently, the statistics are based on the invariance principle for Brownian motion when cumulative sums of abnormal returns are applied. The other diffusion processes such as Lévy processes, or jump processes can be introduced to allow even broader possibilities where the statistics can be applied. Specifically, the asymptotic normality assumption will not become the limitation for the occupation time statistics to analyze the duration of impacts from events. The jump processes (or simply the Brownian motion with random drift) will also reduce the possibility of sporadic disturbances.

In addition, the intensity of these cumulative abnormal returns (over the entire sample period) can be used as alternative measures for the event studies as well. Two alternative statistics of diffusion process (in Chapter 5, for example) are provided as

$$\pi_i(\omega) = \frac{1}{T}\sum_{m=1}^{T}\delta\left\{\left[\frac{1}{\sqrt{T}\tilde{\sigma}_{\epsilon_i}}\left|\sum_{t=1}^{m}\hat{\epsilon}_{it} - m\bar{\hat{\epsilon}}_i\right|\right] > h\right\}$$
$$\xrightarrow{d} \int_0^1 \delta(|\overline{B}(z)| > h)\mathrm{d}z,$$

or

$$\pi_i^\star(\omega) = \frac{1}{T}\sum_{m=1}^{T}\left[\frac{1}{\sqrt{T}\tilde{\sigma}_{\epsilon_i}}\left|\sum_{t=1}^{m}\hat{\epsilon}_{it}\right|^2\right] \xrightarrow{d} \int_0^1 |B(z)|^2\mathrm{d}z,$$

where they can also be applied to analyze the intensity of cumulative abnormal returns in event studies.

Furthermore, the development of monitoring test(s) in using occupation time statistics is more promising than the online detection provided in many statistical and econometric literature. Allowing stability of the occupation time statistics under the null is less demanding than the monitoring tests that require no contamination (for parametrization of models) in the estimation period, at least, in financial time series of interest. Given the time-varying nature of financial time series, the notion as testing structural change is perhaps less essential than promptly accommodate the changes in tracking. The real challenge, however, is "how long these impacts may last?" Monitoring tests based on occupation time statistics are to assess the essentiality (of forthcoming event(s)) based on the intensity of the impacts may follow. All these works can be elaborated for extension of event studies in corporate finance.

Notes

1 Data Collection in Long-Run or Short-Run Format?

1. For simplicity, the notation for information signal y_i is not denoted with the time index since the announcement time for the event(s) may not be exact.
2. The correctness here means, the model specification is best-fitted with respect to the systematic information set Π_t and, the inclusion of essential (not just statistically significant) explanatory variables.
3. For instance, $\ln(\prod_{t=1}^{T}[1+r_{it}]-\prod_{t=1}^{T}[1+r_{mt}])= \sum_{t}^{t+T}(r_{it}-r_{mt})+v_{it}$, since $ln(1+r_{it}) \approx r_{it}$ and $ln(1+r_{mt}) \approx r_{mt}$, respectively, using Taylor expansions around zero.
4. These attributes can be extended to multidimensional so that r_{mt} can be a k-dimensional stochastic process or, it can be the return of some non-event compatible firms.
5. Notice that this setting has included some firm-specific attributes for the conditional expectation of excess returns. However, these attributes may not be the non-diversifiable factors of excess returns. In Chapter 2 later on, it will be shown that inclusion of these attributes in the model(s) may not improve the tests of event studies event when t lies in the event period where announcement is available, if these included attributes are associated with the events.
6. Implicitly, this setting implies that $E[\epsilon_{it}|W_{it}] \neq 0$ when t lies in the announced event period.

2 Model Specifications for Normal (or Expected) Returns

1. Another limitation of Bai's (2003) model is that it implicitly assumes the growth rate for the sum of squared factor loadings is bounded by N. This

implies that the factor loadings of each factor belong to the l_2–space, which contains all squared summable (infinite) sequences of real numbers. The assumption, then, limits the scope for discussing the nondiversifiability of the (hidden) factor(s).

2. Notice that when $k = 0$, the model in equation (7) may only consider a drift or a constant term. In this case, the ϵ_{it} may be considered as the mean-adjusted abnormal return for event studies.
3. The factor loadings of fitted model in equation (10) and the factor loading β_i^h can be stochastic and time-varying also. The extension, however, may need to assume that the factors (included and hidden) and the factor loadings are mutually independent. Nevertheless, the conclusion of the nondiversifiability of hidden factor and its implication of strong cross-sectional memory will still hold.
4. Notice that this growth rate is for the cumulative sums of the covariances, not for the partial sums of covariances.
5. Notice that the condition is applied to L^2-norm of the cumulative sums, not just the sums themselves.
6. In fact, even if the mean $E[\epsilon_{it}] \neq 0$, the following definition can be modified to $N^{-\alpha}\left(\sum_{m=1}^{N} \mathrm{Var}\left(\frac{1}{N}\sum_{i=1}^{m}\epsilon_{it}\right)\right) < \infty$,.as $N \to \infty$. And the following assumption A7 will be modified to weak convergence to fractional Brownian motion with drift. Yet, this will incur additional technicality without improving the asymptotic arguments. Hence, this is left for further studies.
7. Notice that the bottom-up sequential model search is not identical to forward search proposed by Atkinson and Riani (2002). Their approach is based on the increasing subsets of all observations to verify the model.
8. One reason for the orthogonalization is to obtain the estimate the factor loadings of newly included proxy - given that these proxies are possibly correlated. This approach is similar to Forsythe et al. (1973) or so-called orthogonal least squares algorithm in system science.
9. It is not surprising to see that some empirical asset pricing models with inclusion of diversifaible factor(s) may still enjoy tentative predictability. Although promising, this predictability do not sustain over different time horizons.

3 Cumulative Abnormal Returns or Structural Change Tests?

1. Apparently, one can set that $\varepsilon_t^i = \lambda_i + \xi_{it}$,where $\lambda_i \neq 0$, when $t \in (T_1, T_2]$.
2. For further extension that considers when $\gamma = \frac{1}{2}$, it is discussed in Huková et al. (2007).

4 Recursive Estimation for Normal (or Expected) Returns

1. Certainly, these explanatory variables may include some lagged dependent/ independent variables. In addition, ψ_t can also be extended to include the functionals of known explanatory variables in some suitable function spaces.
2. The notation $\log^{\alpha}(x)$ is denoted for $(\log(x))^{\alpha}$ for simplicity.
3. Notice that the notations $||.||_{L_{3p}}$and $||.||_{L_{6p}}$ represent the L^{3p}–norm and L^{6p}–norm of the variable of interest, respectively.
4. Although stochastic function is more general, it is far more difficult to obtain asymptotic properties for the system.
5. Grillenzoni (2008) also considers the kernel version of the weighted least squares method. However, since the recursive version for the kernel version is not feasible, the discussions focus on the exponential version only.

5 Time Will Tell! A Method with Occupation Time Statistics

1. It is straightforward to verify that if η_{it} and some of $\left\{\phi_{jt}\right\}_{j=1,2,\cdots,k}$ are not mutually orthogonal, the pricing error can still be absorbed by the time-varying coefficient model. For instance, let $\eta_{it} = \tilde{\delta}_i\left(\sum_{j=1}^{k}\phi_{jt}\right)$, $\tilde{\delta}_i \neq 0$, it is convenient to combine this terms with the other $\left\{\phi_{jt}\right\}_{j=1,2,\cdots,k}$ and the equation will change to $r_{it} = \ddot{\alpha}_{it} + \sum_{j=1}^{k}\ddot{\beta}_{jt}^{i}\phi_{jt} + \epsilon_{it}$.
2. Certainly, weak convergence to other diffusion processes such as Levy process can be assumed to allow more generality where jumps and asymptotic nonnormality are considered.
3. Given that there are many different definitions for the types of Banach space, the definition of Ledoux and Talagrand (1991) is applied for simplicity.
4. For simplicity, the Banach-valued central limit theorem for independent Banach-valued random variables here assuming the occupation time statistics are mutually independent. In addition, the occupation time functionals will converge to some independent identically-distributed occupation times of reflected Brownian motion under the null. Further extensions can be obtained if applying the central limit theorem for dependent Banach-valued random variables.
5. Here, by saying $T \to \infty$ faster than n is to assume that the sample time period expands faster than the numbers of selected firms concerning the event(s) of interest. This seems reasonably feasible, especially for financial time series.
6. The independence is assumed for simplicity in multiple hypothesis testing for occupation time statistics. However, further extensions on allowing dependence of these hypotheses are available in many references.

References

Acharya, S., 1993, Value of latent information: alternative event study methods, *Journal of Finance*, 48, 363–385.

Aktas, N., E. deBodt, and J. Cousin, 2007, Event study under contaminated estimation period, *Journal of Corporate Finance*, 13, 129–145.

Ang, J. and S. Zhang, 2004, An evaluvation of testing procedures for long horizon event studies, *Review of Quantitative Finance and Accounting*, 23, 251–274.

Araujo, A., and E. Giné, 1980, *The central limit theorem for real and Banach valued rando variables*, John Wiley an Sons, Inc.

Asquith, P., R. F. Bruner, D. W. Mullins, 1983, The gains to bidding firms from merger, *Journal of Financial Economics*, 11, 121–139.

Atkinson, A. C., and M. Riani, 2002, Forward search added variable t tests and the effects of masked outliers on model selection, *Biometrika*, 89, 939–946.

Aue, A., and L. Horváth, 2013, Structural breaks in time series, *Journal of Time Series Analysis*, 34, 1–16.

Bai, J., 2003, Inferential theory for factor models of large dimensions, *Econometrica*, 71, 135–171.

Bailey, N., G. Kapetanios, and M. H. Pesaran, 2012, Exponent of cross-sectional dependence: estimation and inference, working paper.

Barber, B. M., and J. D. Lyon, 1997, Detecting long-run abnormal stock returns: the empirical power and specification of test statistics, *Journal of Financial Economics*, 43, 341–372.

Barber, B. M. and J. D. Lyon, 1997, Firm size, book-to-arket ratio, and security returns: A hold-out sample of financial firms, *Journal of Finance*, 52, 875–883.

Benjamini Y., and Y. Hochberg, 1995, Controlling the false discovery rate: A practical and powerful approach to multiple testing, *Journal of the Royal Statistical Society* B, 57, 289–300.

Bernard, V. L., 1987, Cross-sectional dependence and problems in inference in market-based accounting research, *Journal of Accounting Research*, 25, 1–48.

Bohn, N., F. A. Rabhi, D. Kundisch, L. Yao, and T. Mutter, 2013, Toward automated event studies using high frequency news and trading data, *Enterprise Applications and Services in the Financial Industry*, Rabhi, F. A., and P. Gomber (eds.), 20–41.

Bradley, M., 1980, Interfirm tender offers and the market for corporate control, *Journal of Business*, 53, 345–376.

Branch, B., and T. W. Yang, 2003, Predicting successful takeovers and risk arbitrage, *Journal of Business and Economics*, 42, 3–18.

Bremer, R., B. G. Buchanan, and P. C. English II, 2011, The advantage of using quarterly returns for long-term event studies, *Review of Quantitative Finance and Accounting*, 36, 491–516.

Brockett, P. L., H-M. Chen, and J. R. Garven, 1999, A new stochastically flexible methodology with application to Proposition 103, *Insurance: Mathematics and Economics*, 25, 197–217.

Brown, K. C., L. J. Lockwood, and S. Lummer, 1985, An examination of event dependency and structural change in security pricing models, *Journal of Financial and Quantitative Analysis*, 20, 315–334.

Brown, K. C., and M. V. Raymond, 1986, Risk arbitrage and the prediction of successful corporate takeovers, *Financial Management*, Autumn, 54–63.

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

Brown, S. J., and J. B. Warner, 1985, Using daily stock returns—the case of event studies, *Journal of Financial Economics*, 14, 3–31.

Cable, J. and K. Holland, 1999, Modeling normal returns in event studies: a model-selection approach and pilot study, *European Journal of Finance*, 5, 331–341.

Carhart, M. M., 1997, On persistence in mutual fund performance, *Journal of Finance*, 52, 57–82.

Chamberlain, G., 1983, Funds, factors, and diversification in arbitrage pricing models, *Econometrica,* 51, 1305–1323.

Chamberlain, G. and M. Rothschild, 1983, Arbitrage, factor structure, and mean-variance analysis on large asset markets, *Econometrica,* 51, 1281–1304.

Chen, H.-F., and L. Guo, 1991, *Identification and Stochastic Adaptive Control*, Birkhäuser.

Chochola, O., M. Huková, Z. Práková, and J. G. Steinebach, 2013, Robust monitoring of CAPM portfolio betas, *Journal of Multivariate Analysis*, 115, 374–395.

Chudik, A., M. H. Pesaran, and E. Tosetti, 2011, Weak and strong cross-section dependence and estimation of large panels, *Econometrics Journal*, 14, 45–90.

Collins, D. W., and W. T. Dent, 1984, A comparison of alternative testing methodologies used in capital market research, *Journal of Accounting Research*, 22, 48–84.

Connor, G. and R. A. Korajczyk, 1993, A test for the numbers of factors in an approximate factor model, *Journal of Finance*, 48, 1263–1291.

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

Cyree, K. B. and R. P. Degennaro, 2002, A generalized method for detecting abnormal returns and changes in systematic risk, *Review of Quantitative Finance and Accounting*, 19, 399–416.

Dahlhaus, R. and S. S. Rao, 2007, Statistical inference for time-varying ARCH processes, *Annals of Statistics*, 34, 1075–1114.

Dann, L., 1981, Common stock repurchase: an analysis of returns to bondholders and stockholders, *Journal of Financial Economics*, 9, 113–138.

Davidson, W. N., D. Dutia, and L. Cheng, 1989, A re-examination of the market reaction to failed mergers, *Journal of Finance*, 44, 1077–1083.

Debont W. F. M., and R. Thaler, 1985, Does the stock market overreact? *Journal of Finance*, 40, 793–805.

Debont W. F. M., and R. Thaler, 1987, Further evidence on investor overreaction and stock market seasonality, *Journal of Finance*, 42, 557–581.

Dehling, H., 1983, Limit theorems for sums of weakly dependent Banach space valued random variables, *Probability Theory and related fields*, 63, 393–432.

Dennis, D. K., and J. J. McConnell, 1986, Corporate merges and security returns, *Journal of Financial Economics*, 16, 143–187.

Dodd, P., 1980, Merger proposals, management discretion and stockholder wealth, *Journal of Financial Economics*, 8, 105–138.

Dodd, P., and R. Ruback, 1977, Tender offers and stockholder returns: an empirical analysis, *Journal of Financial Economics*, 5, 351–374.

Draper, P., and K. Paudyal, 1999, Corporate takeovers: mode of payment, returns, and trading activity, *Journal of Business Finance and Accounting*, 26, 521–558.

Eberlein, E., 1986, On strong invariance principles under dependence assumptions, *Annals of Probability*, 14, 260–270.

Efron, B., 2010, *Large-scale inference*, Cambridge University Press.

Embrechts, P. and M. Maejima, 2002, *Self-similar Processes*, Princeton University Press.

Ermakov, S.V., and E. I. Ostrovoskii, 1986, The central limit theorem for weakly dependent Banach-valued random variables, *Theory of Probability and its applications*, 30, 391–394.

Fama, E., 1998, Market efficiency, long-term returns, and behavioral finance, *Journal of Financial Economics*, 49, 283–306.

Fama, E., L. Fisher, M. C. Jensen, and R. Roll, 1969, The adjustment of stock prices to new information, *International Economic Review*, 10, 1–21.

Fama, E. F. and K. R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics,* 25, 23–49.

Fatalov, V. R., 2009, Occupation time and exact asymptotics of distributions of L^p−functionals of the Ornstein-Uhlenbeck prosses, $p > 0$, *Theory of Probability and its applications*, 53, 13–36.

Ferger, D., 2000, Optimal tests for the general two-sample problem, *Journal of Multivariate Analysis*, 74, 1–35.

Fitzsimmons, P. J., and R. K. Getoor, 1995, Occupation time distributions for Lévy bridges and excursions, *Stochastic Processes and their applications*, 58, 73–89.

Forsythe, A. B., L. Engelman, R. Jennrich, and P. R. A. May, 1973, A stopping rule for variable selection in multiple regression, *Journal of American Statistical Association*, 68, 75–77.

Fu, F., 2009, Idiosyncratic risk and the cross-section of expected stock returns, *Journal of Financial Economics*, 91, 24–37.

Giraitis, L., P. Kokoszaka, R. Leipus, and G. Teyssiere, 2003, Rescaled variance and related tests for long memory in volatility and levels, *Journal of Econometrics,* 112, 265–294.

Grillenzoni, C. 2000, Time-varying parameters prediction, *Annals of the Institue of Statistical Mathematics*, 52, 108–122.

Grillenzoni, C. 2008, Performance of adaptive estimators in slowly varying parameter models, *Statistical Methods and Applications*, 17, 471–482.

Guégan, D., 2005, How can we define the concept of long memory? An econometric survey, *Econometric Reviews*, 24, 113–149.

Guo, L., 1990, On adaptive stabilization of time-varying stochastic systems, *SIAM journal of Control and Optimization*, 28, 1432–1451.

Guo, L., 1994, Stability of recursive stochastic tracking algorithms, *SIAM Journal of Control and Optimization*, 32, 1195–1225.

Guo, L., and L. Ljung, 1995a, Exponential stability of general tracking algorithms, *IEEE Transactions on Automatic Control*, 40, 1376–1387.

Guo, L., and L. Ljung, 1995b, Performance analysis of general tracking algorithms, *IEEE Transactions on Automatic Control*, 40, 1388–1402.

Guo, L., L. Xia, and J. B. Moore, 1991, Tracking randomly varying parameters analysis of a standard algorithm, *Mathematics of Control, Signals and Systems*, 4, 1–16.

Hall, P., and I.V. Keilegom, 2007, Two-sample tests in functional data analysis starting from discrete data, *Statistica Sinica*, 17, 1511–1531.

Harrington, S. E., and D. G. Shrider, 2007, All event induced variance: analyzing abnormal returns when effects vary across firms, *Journal of Financial and Quantitative Analysis*, 42, 229–256.

Hastie, T., and R. Tibshirani, 1993, Varying-coefficient models, *Journal of Royal Statistical Society B, Methodological*, 55, 757–796.

Heyde, C. C., and Y. Yang, 1997, On defining long-range dependence, *Journal of Applied Probability*, 34, 939–944.

Hoffmann-Jorgensen, J. and G. Pisier, 1976, The law of large numbers and the central limit theorem in Banach spaces, *Annals of Probability*, 4, 587–599.

Horváth, L., M. Huková, P. Kokoszka, and J.G. Steinebach, 2004, Monitoring changes in linear models, *Journal of Statistical Planning and Inference*, 126, 225—251.

Horváth, L., P. Kokoszka, and J. G. Steinebach, 2007, On sequential detection of parameter changes in linear regression, *Statistics and Probability Letters*, 77, 885–895.

Hutson, E., 2000, Takeover targets and the probability of bid success: evidence from the Australian market, *International Review of Financial Analysis*, 9, 45–65.

Jahn, J., 2007, *Introduction to the Theory of Nonlinear Optimization*, 3rd ed., Springer Science & Business Media.

Jakobsen, J. B., and T. Voetmann, 2005, A new approach for interpreting long-run returns, applied to IPO and SEO stocks, *Annals of Economics and Finance*, 6, 337–363.

Jarrell, G. A., and A. B. Poulsen, 1989, The returns to acquiring firms in tender offers: evidence from three decades, *Financial Management*, Autumn, 18, 3–12.

Jegadeesh, N., and J. Karceski, 2009, Long-run performance evaluation: correlation and heteroskedasticity-consistent tests, *Journal of Empirical Finance*, 16, 101–111.

Jeng, J.-L., and W. Liu, 2012, Do idiosyncratic risks really contain a hidden non-diversifiable factor? A diagnostic testing approach, *Journal of Mathematical Finance*, 2, 251–263.

Jeng, J.-L, D. Park, and M. Dewally, 2014, On the occupation times of mergers and acquisitions in event studies, working paper.

Jeng, J.-L., and E. Tobing, 2012, Finding empirical multifactor asset-pricing models - A sequential model search approach, working paper.

Joensen, A. K., H. Madsen, H. A. Nielsen, and T. S. Nielsen, 2000, Tracking time-varying parameters with local regression, *Automatica*, 36, 1199–1204.

Johnson, B. McK. and T. Killeen, 1983, An explicit formula for the C.D. F. of the L_1 norm of the Brownian bridge, *Annals of Probability*, 11, 807–808.

Karafiath, I., and D. E. Spencer ,1991, Statistical inference in multi-period event studies, *Review of Quantitative Finance and Accounting*, 1, 353–371.

Kothari, S.P., and J. B. Warner, 1997, Measuring long-horizon security price performance, *Journal of Financial Economics*, 43, 301–339.

Khotari, S. P., and J. B. Warner, 2007, Econometrics of event studies, *Handbook of Corporate Finance: Empirical Corporate Finance*, Vol. 1, 4–36.

Lahiri, S. N., 2003, *Resampling Methods for Dependent Data*, Springer-Verlag.

Ledoux M., and M. Talagrand, 1991, *Probability in Banach spaces*, Springer–Verlag.

Loughran, T., and J. R. Ritter, 1995, The new issue puzzle, *Journal of Finance*, 50, 23–51.

Loughran, T., and A. M. Vijh, 1997, Do long-term shareholders benefit from corporate acquisitions? *Journal of Finance*, 52, 1765–1790.

Lyon, J. D., B. M. Barber, and C.-L.Tsai, 1999, Improved methods for tests of long-run abnormal stock returns, *Journal of Finance*, 54, 165–201.

Mackinlay, A. C., 1997, Event studies in economics and finance, *Journal of Economic Literature*, 35, 13–39.

Masulis, R. W., 1980, The effects of capital structure change on security prices: a study of exchange offers, *Journal of Financial Economics*, 8, 139–178.

Mei, C., W. Zhang, and Y., Leung, 2001, Statistical inferences for varying-coefficient models based on locally weighted regression technique, *Acta Mathematicae Applicatae Sinica*, 17, 407–417.

Morse, D., 1984, An econometric analysis of the choice of daily versus monthly returns in test of information content, *Journal of Accounting Research*, 22, 605–623.

Nielsen, H. K., T. S. Nielsen, A. K. Joensen, H. Madsen, and J. Holst, 2000, Tracking time-varying-coefficient functions, *International Journal of Adaptive Control and Signal Processing*, 14, 813–828.

Ouysse, R., 2006, Consistent variable selection in large panels when factors are observable, *Journal of Multivariate Analysis*, 97, 946–984.

Parzen, E., 1981, Time series model identification and prediction variance horizon, in *Applied Time Series Analysis* II, 415–447.

Pesaran, M. H., 2006, Estimation and inference in large heterogeneous panels with a multifactor error structure, *Econometrica*, 74, 967–1012.

Pettengill, G. N. and J. M. Clark, 2001, Estimating expected returns in an event study framework: evidence from the Dartboard column, *Quarterly Journal of Business and Economics*, 40, 3–21.

Pötscher, B. M., 2004, Nonlinear functions and convergence to Brownian motion: beyond the continuous mapping theorem, *Econometric Theory*, 20, 1–22.

Rau, P. R., and T. Vermaelen, 1998, Glamour, value and the post-acquisition performance of acquiring firms, *Journal of Financial Economics*, 49, 223–253.

Ritter, J. R., 1991, The long-run performance of initial public offerings, *Journal of Finance*, 46, 3–27.

Saporoschenko, A., 2011, The effect of Santa Ana wind conditions and cloudiness on Southern California stock returns, *Applied Financial Economics*, 21, 683–694.

Sen, P. K., 1982, Invariance principles for recursive residuals, *Annals of Statistics*, 10, 307–312.

Strong, N., 1992, Modelling abnormal returns: a review article, *Journal of Business Finance and Accounting*, 19, 533–553.

Takacs, L., 1998, Sojourn times for the Brownian motion, *Journal of Applied Mathematics and Stochastic Analysis*, 11, 231–246.

Tang, Q.-Y., P. Ľecuyer, and H.-F. Chen, 1999, Asymptotic efficiency of perturbation-analysis-based stochastic approximation with averaging, *SIAM Journal of Control and Optimization*, 37, 1822–1847.

Tanré, E., and P. Vallois, 2006, Range of Brownian motion with drift, *Journal of Theoretical Probability*, 19, 45–69.

Thompson, J. E., 1988, More methods that make little difference in event studies, *Journal of Business Finance and Accounting*, 15, 77–86.

Thompson, J. E., 1989, An alternative control model for event studies, *Journal of Business Finance and Accounting*, 16, 507–513.

White, H., 2001, *Asymptotic Theory for Econometricians*, Academic Press.

Wu, W., 2007, Strong invariance principles for dependent random variables, *Annals of Probability*, 35, 2294–2320.

Zhang, H., and O. Hadjiliadis, 2012, Drawdowns and the speed of market crash, *Methodology and Computing in Applied Probability* 14 (3), 739–752.

Zhang, J. F., L. Guo, and H.-F. Chen, 1991, L_p-stability of estimation errors of Kalman filter for tracking time-varying parameters, *International Journal of Adaptive Control and Signal processing*, 5, 155–174.

Zhou, Y. and M. S. Taqqu, 2006, How complete random permutations affect the dependence structure of stationary sequences with long-range dependence, *Fractals*, 14, 205–222.

Index