https://doi.org/10.26219/heal.aueb.6067

ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

DEPARTMENT OF STATISTICS

POSTGRADUATE PROGRAM

STRUCTURAL BREAK MODELS: A REVIEW AND APPLICATION IN A HEDGE FUND GENERATING PROCESS

By

Ilias K. Mamouras

A THESIS
Submitted to the Department of Statistics
of the Athens University of Economics and Business
in partial fulfilment of the requirements for
the degree of Master of Science in Statistics

Athens, Greece
September 2009

https://doi.org/10.26219/heal.aueb.6067

https://doi.org/10.26219/heal.aueb.6067

ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ

ΜΟΝΤΕΛΑ ΔΟΜΙΚΩΝ ΑΛΛΑΓΩΝ:ΑΝΑΣΚΟΠΗΣΗ ΚΑΙ ΕΦΑΡΜΟΓΗ ΣΕ ΑΠΟΔΟΣΕΙΣ ΕΝΑΛΛΑΚΤΙΚΩΝ ΜΟΡΦΩΝ ΕΠΕΝΔΥΣΗΣ

Ηλίας Κ. Μαμούρας

ΔΙΑΤΡΙΒΗ

Που υποβλήθηκε στο Τμήμα Στατιστικής

του Οικονομικού Πανεπιστημίου Αθηνών

ως μέρος των απαιτήσεων για την απόκτηση

Μεταπτυχιακού Διπλώματος Ειδίκευσης στη Στατιστική

Αθήνα

Σεπτέμβριος 2009

https://doi.org/10.26219/heal.aueb.6067

https://doi.org/10.26219/heal.aueb.6067

DEDICATION

To my parents who have supported and motivated me throughout my life.

https://doi.org/10.26219/heal.aueb.6067

https://doi.org/10.26219/heal.aueb.6067

ACKNOWLEDGEMENTS

I would like to express my appreciation to several people for their support during my thesis work. First of all I would like to thank my supervisor Lecturer Ioannis Vrontos for his valuable guidance, undivided help and understanding throughout this work. Finally, I am grateful to my family for their support, love and patience.

https://doi.org/10.26219/heal.aueb.6067

https://doi.org/10.26219/heal.aueb.6067

VITA

I was born in Volos, Greece in July 28, 1983 and graduated from the High school of Almyros in 2001. In 2001 I entered the Department of Statistics of Athens University of Economics and Business (AUEB) and four and a half years later, in February 2006 received my degree in Statistics. The year 2007 I was accepted by the Department of Statistics of Athens University of Economics and Business (AUEB) for the Full-Time MSc of Statistics.

https://doi.org/10.26219/heal.aueb.6067

https://doi.org/10.26219/heal.aueb.6067

ABSTRACT

Ilias Mamouras

STRUCTURAL BREAK MODELS: A REVIEW AND APPLICATION IN A HEDGE FUND GENERATING PROCESS

September 2009

The aim of the present dissertation is the review and presentation of some of the structural break risk factor models and afterwards the application of them to a generating process that concerns the hedge funds returns. Firstly we present some basic information about hedge funds and their characteristics and next we present various strategies that have been suggested for testing a linear model for structural changes. We state analytically the generalized fluctuation tests, the F-tests (Chow test, supF, aveF and expF tests), the Bai-Perron method, the Rolling Regression Analysis and the Recursive Regression Analysis. Afterwards, these strategies are applied to real data that concern the return generating process of some hedge funds in order to examine the return series for structural breaks. Also we apply these strategies in data that we have simulated. Furthermore, we present the on-line monitoring method that is used in order to predict a data series for structural breaks when incoming data arrive. The on-line monitoring method is applied to the data that concern the hedge fund's returns and the simulated data of the present dissertation.

https://doi.org/10.26219/heal.aueb.6067

https://doi.org/10.26219/heal.aueb.6067

ΠΕΡΙΛΗΨΗ

Ηλίας Μαμούρας

ΜΟΝΤΕΛΑ ΔΟΜΙΚΩΝ ΑΛΛΑΓΩΝ:ΑΝΑΣΚΟΠΗΣΗ ΚΑΙ ΕΦΑΡΜΟΓΗ ΣΕ ΑΠΟΔΟΣΕΙΣ ΕΝΑΛΛΑΚΤΙΚΩΝ ΜΟΡΦΩΝ ΕΠΕΝΔΥΣΗΣ

Σεπτέμβριος 2009

Σκοπός της παρούσας διπλωματικής εργασίας είναι η αναφορά και παρουσίαση μερικών μοντέλων δομικών αλλαγών και παραγόντων ρίσκου και αργότερα η εφαρμογή αυτών σε μια διαδικασία παραγωγής αποδόσεων που αφορά τα hedge funds. Αρχικά παρουσιάζουμε κάποιες βασικές πληροφορίες σχετικά με τα hedge funds και τα χαρακτηριστικά αυτών και μετά παρουσιάζουμε διάφορες στρατηγικές που έχουν προταθεί για τον έλεγχο ενός γραμμικού μοντέλου για δομικές αλλαγές. Αναφέρουμε αναλυτικά τους γενικευμένους ελέγχους διακύμανσης, τους ελέγχους F (έλεγχος Chow, supF, aveF, και expF), την μέθοδο των Bai-Perron, την κυλιόμενη ανάλυση παλινδρόμησης και την επαναλαμβανόμενη ανάλυση παλινδρόμησης. Μετέπειτα, αυτές οι στρατηγικές εφαρμόζονται σε πραγματικά δεδομένα που αφορούν την διαδικασία παραγωγής αποδόσεων κάποιων hedge funds με σκοπό να ελέγξουμε αυτή την χρονολογική σειρά αποδόσεων για δομικές αλλαγές. Επίσης εφαρμόζουμε αυτές τις στρατηγικές σε δεδομένα τα οποία έχουμε προσομοιώσει. Επιπροσθέτως, παρουσιάζουμε την μέθοδο της πραγματικής παρακολούθησης η οποία χρησιμοποιείται για να προβλέψουμε δομικές αλλαγές σε ένα σετ δεδομένων όταν ανανεώνεται στο σετ αυτό με νέα δεδομένα. Η μέθοδος της πραγματικής παρακολούθησης εφαρμόζεται στα δεδομένα που αφορούν τις αποδόσεις των hedge funds και στα προσομοιωμένα δεδομένα της παρούσας διπλωματικής εργασίας.

https://doi.org/10.26219/heal.aueb.6067

https://doi.org/10.26219/heal.aueb.6067

TABLE OF CONTENTS

https://doi.org/10.26219/heal.aueb.6067

https://doi.org/10.26219/heal.aueb.6067

Chapter 6

https://doi.org/10.26219/heal.aueb.6067

LIST OF TABLES

https://doi.org/10.26219/heal.aueb.6067

https://doi.org/10.26219/heal.aueb.6067

 https://doi.org/10.26219/heal.aueb.6067

LIST OF FIGURES

https://doi.org/10.26219/heal.aueb.6067

https://doi.org/10.26219/heal.aueb.6067

https://doi.org/10.26219/heal.aueb.6067

https://doi.org/10.26219/heal.aueb.6067

Introduction

In this dissertation we are dealing with the presentation and discussion of the breakpoint models as well as with their application in data that concern hedge fund returns. We are going to describe these models and the corresponding structural break tests which we apply in these models in order to test them for structural changes. After presenting these models and the tests we are going to analyze a time series of hedge funds returns using the appropriate model and the appropriate tests. Before proceeding in the presentation of each of the chapters of the present dissertation we shall state and discuss the whole subject that we are up to in this dissertation.

The last decade, hedge funds have grown and have received a lot of attention in the global markets. The hedge funds, as alternative investment vehicles with great trading flexibility, pre-existed since 1949 but the last years are developing with high rates and are considered to be the epitome of active management. The most important reasons for the huge development of the hedge funds are the whole global economic development and the need for alternative investments. As we shall see in the following chapters of this dissertation, a hedge fund, as a private pool of capital which is not being regulated and is administered by a manager, sets as aim the highest returns and profits based on the money that have been invested on it by the different investors. As someone can understand from the above, the manager of the hedge fund can choose from a wide variety of strategies in order to success the above aims. Due to the fact that these movements are complicated, the hedge fund's structure is also complicated and that assigns a lot of properties and characteristics to them. As a consequence, the estimation and evaluation of the hedge fund's returns becomes difficult and complex. Another reason that affects the evaluation of the hedge fund's returns is the fact that there exist many risk factors that affect the invested capital. This also makes the procedure of the exact and correct evaluation of the returns more difficult and increases the possibility of a wrong evaluation. Additionally, as research has shown, the time series data of the returns of the hedge funds is characterized by some special properties such as the non-normality, the fat tails and the skewness (Amin and Kat, 2003, Kouwenberg and Ziemba, 2007). Also the time series of the returns of the hedge funds might be characterized by nonlinearities (Fung and Hsieh, 2004). So the need of finding models which take into account the above special characteristics and the different market condition is imperative. Like this, there were proposed the models that take into account the risk factors that affect the hedge funds return generating process. Especially these models take into account the factors that affect the returns of the hedge funds and also model in a way the ability of the manager of the hedge fund of taking profitable decisions. It is obvious that these models are an attempt of fully modelling the properties and the characteristics of the hedge funds so that the evaluation of the return generating process becomes easier and the manager of the hedge fund chooses the correct strategy that is going to be the most profitable and the one with the less risk. A category of these risk factor models are the break-point risk factor models. These are the models that we are dealing with in the present dissertation.

The break-point risk factor models have many applications in statistics and econometrics and are exceptional in the case of modelling the return generating process of a hedge fund. These models can recognize the changes that happen in the process due to the market conditions. These changes concern changes in the mean, the variance or in the mean and the variance at the same time, of the return generating process. Also another case that these models can recognize the change is the case in which the slope of the linear model changes at a particular point. In other words we can say that the break-point risk factor models are suitable for the case of modelling the hedge fund's return generating process.

Having concisely stated above the subject of the present dissertation we are now going to specify the contents of each chapter. Firstly, in the 1st chapter, we are dealing with the hedge funds. Especially we are giving the definition of the hedge funds and we comment on some basic but important characteristics of them. These characteristics concern the fees that a manager of a hedge fund or someone who invests capital on a hedge fund receives. Also another characteristic that we are commenting on the 1st chapter are the possible strategies that a manager can choose in order to accomplish the

https://doi.org/10.26219/heal.aueb.6067

hedge fund's aim. Furthermore, we are stating the characteristic that concerns the reasons that there exists high risk in the hedge fund's investments, the characteristic that concerns the whole structure of a hedge fund and the characteristic that concerns the benefits of the use of the hedge funds in the global economy. Lastly, we mention some history information about the hedge funds and present the top 10 hedge funds for the previous years as well as some of the most notable hedge fund management companies. The 1st chapter contains some basic information about hedge funds. We decided to state only the above basic and introductory information and not to fully describe the hedge funds due to the fact that our main scope in the present dissertation is the application of the break-point risk factor models in data that concern the return generating process of a hedge fund.

In the 2nd chapter, we are dealing with the tests and the statistical methods that are used in statistics and econometrics in order to test a linear regression model for structural changes. When someone uses a break-point risk factor model in order to analyze time series data he needs the appropriate tests in order to test this series for possible structural breaks. In this chapter, firstly we give a brief summary of the simple linear regression model and then we state the different structural break tests. At this point we have to mention that we classify the structural change tests according to their use. That is to say that in the first subsection of the 2nd chapter we state the tests that are used in order to backtest a data series for structural changes (historical tests) and in the second subsection we state the tests that are used in order to monitor a data series when incoming data arrive (on-line monitoring tests). In the historical tests we state the generalized fluctuation tests, the F-tests, the Bai-Perron method, the Rolling Regression Analysis and the Recursive Regression Analysis for detecting structural breaks. Especially, in the generalized fluctuation tests we state the CUSUM and MOSUM tests and the fluctuation test and in the F-tests we state the Chow, the supF, the aveF and the expF tests. In the subsection which concerns the on-line monitoring tests we state the estimates based processes, which cover the recursive estimates process (RE) and the Moving Estimates process (ME), and the residual based processes which include the OLS CUSUM and the OLS MOSUM processes.

 https://doi.org/10.26219/heal.aueb.6067

In the 3rd chapter we present a review of some of the most important break-point risk factor models that we used and analyzed in the present dissertation. Especially in each case we present the corresponding model and its characteristics and we also simulate data for each model and describe the results that come out by estimating these models and testing them for structural changes. The models that we are up to in this chapter are the models that describe data in which there is a change in the mean of the data generating process, in the variance, in the mean and the variance simultaneously and lastly in the slope of the model. Also we apply some different methods for estimating the above models such as the Simple Linear Regression analysis, the Rolling Regression analysis and the Recursive Regression analysis.

In the 4th chapter we deal with the simulation of data and test for break-points. Specifically, we simulate data and using the tests that are described in the 2nd chapter we try to find and estimate the possible break-points. We also apply the on-line monitoring to the simulated data in order to test for possible breaks having already some of the initial data. As in the application of the 5th chapter, the application of the realistic data, in the on-line monitoring method that we apply to the simulated data we use both the Moving Estimates and the Recursive Estimates tests as monitoring methods. Also in each case we use the bandwidth values h=1 and h=0.5.

In the 5th chapter we deal with the application of a break-point risk factor model in real data that concern the generating return process of a hedge fund. For the real data, we use the model that Fung and Hsieh (2004) proposed. They proposed a model for hedge fund returns which is linear to seven asset-based style (ABS) factors and which has dynamic coefficients. Using this model we try to analyze the HFRCI index (Hedge Fund Research Composite Index) and the CSTHFCI index (Credit Suisse/Tremont Hedge Fund Composite Index) using the tests that are described in the 2nd chapter. A manager can count on these indices and according to their values he can decide the appropriate investment strategy for the hedge fund. It is worthy to mention at this point that apart from applying the Fung and Hsieh' s model in the realistic data that concern a hedge fund's returns and trying to detect the possible breaks we also deal with the on-line monitoring process. We apply

https://doi.org/10.26219/heal.aueb.6067

the on-line monitoring method to the realistic data using again the Fung and Hsieh's model. The on-line monitoring method assumes that we already have a part of the full dataset and using the model that we have chosen we try to predict if there exists any break for the rest of the data and if so, we try to distinguish how fast and in which time point will this break-point be observed. This is the main idea of the on-line monitoring method, the idea of trying to detect and estimate the possible breaks in later observations using the initial data that we already have. It is made clear from the above statements that the aim of this method is to help the manager that takes the decisions about the hedge funds investments to be in position to predict the future breakpoints because the future is what is of prime interest for the persons that are busy with investments and generally with the financial data. A manager, by predicting accurately and early enough the possible breakpoints, can choose the most profitable and beneficial strategy for the future of the hedge fund. In the application of this chapter we use the cases of bandwidth values h=1 and h=0.5 and we compare the corresponding results. Lastly we have to mention that in the on-line monitoring process we use the Moving Estimates and the Recursive Estimates tests as monitoring methods.

In the 6th chapter of the present dissertation, which is the last chapter, we put out all the conclusions that we reached to, using the models in the case of the realistic data and the case of the simulated data.

https://doi.org/10.26219/heal.aueb.6067

https://doi.org/10.26219/heal.aueb.6067

Chapter 1

Hedge Funds

1.1 Introduction

Our aim in this chapter is to present the Hedge Funds and their operational characteristics. We are going to state their definition and describe some of their most important features. At this point we have to say that we are going to give some basic information and a synopsis of these features and not to fully describe the hedge fund's characteristics because our main aim in this dissertation is the application of the break-point risk factor models in data that concern the return generating process of a hedge fund.

1.2 Hedge Funds and their characteristics

Nowadays it is widely known that due to the fact of globalization and distracted global economic development, the current investment dogmas and techniques (asset allocation) are about to be old fashioned and fail to fulfil their diversification promises, particularly in turbulent times. As a consequence, investors all around the world tried and achieved to broaden their investment plans and techniques by seeking new and alternative asset classes, the Hedge Funds. A Hedge Fund is a private pool of capital which is not being regulated and whose manager has the right to buy or sell any possessions (assets). Also, a hedge fund is open to a limited number of investors and can undertake a wide range of economic activities. Although its hedge fund has each own strategy which characterizes the type and methods of investments it undertakes, hedge funds invest in a wide range of investments. For the history, we have to say that the term hedge fund was used by Carol Loomis in 1966 in order to describe the investment philosophy

https://doi.org/10.26219/heal.aueb.6067

of Alfred Winslow Jones. Jones, in 1949, had the opinion that the movement of the price of an asset could be analyzed in a factor which measures the overall market and a factor which measures the performance of the asset. In order to inactivate the movement of the market he enriched his portfolio by buying assets whose price was expected to rise up and by selling assets whose price was expected to fall. The fact in this idea is that the loss of the assets that he sold was neutralized by the benefit of the assets that he bought. That is to say that the risk involved in the market price movements is «hedged». As a consequence, someone could allege that Alfred Winslow Jones is the founder of the Hedge Funds and generally of the investment strategies that characterize whole the economic mechanism behind the Hedge Funds. So Hedge Funds existed in economic history since 1949 regardless the fact that the last decade the term Hedge Fund is so famous and someone could believe that they have been invented in the 1990's. Having given the definition and stated the history of the hedge funds we are now going to describe and state some details about their operation and their characteristics.

Firstly, we are going to give some details about the fees of the hedge funds. The Hedge Fund manager is surely going to take on both a management fee and a performance fee. A management fee is the percentage of the net asset value of the fund which is being calculated right at the time when the fee gets payable. Usually these fees are assessed annually and paid to the managers monthly and range from 1% to 4% per annum according to the agreement that has been achieved between the administrations of the hedge funds. On the contrary, performance fees is a percentage of the hedge funds's profits. These fees have hypostasis due to the fact that people who invest in a hedge fund and at the end make money from this investment, want to give some of their money to the manager of the hedge fund as a token of gratitude and as expressing their contentment for the money that they made. It is being intelligible that if the investors make much money then the manager has greater probabilities of being payed a large amount of money. At the moment that this dissertation is being written the prevalent opinion is that the performance fee of a hedge fund is calculated as the 20% of gross returns. But it is sure that some managers charge higher fees and some others charge lower fees. The 20% performance fee is something like the mean value that prevails

 https://doi.org/10.26219/heal.aueb.6067

in the global economic market place. Also we shall mention at this point that some managers charge investors a withdrawal fee if the investors withdraw money from the fund portfolio before a specific period that has been agreed since the investor invested his money for the first time. Furthermore, it is worthy to mention that the performance fees have been criticised for allowing the managers to gain huge profits in a short period of time instead of planning a lucrative long strategy for the hedge fund. Due to this fact the fees of the hedge funds are sometimes limited by a high water mark and sometimes by a hurdle rate. Another way of limiting the profits that managers make is to constrain them to return the performance fees or a percentage of them if the value of the hedge fund drops. It is obvious that these methods apply a maximum value for the fees ensuring that the managers cannot gain extremely high profits.

In this paragraph we are going to mention some details about the strategies that can be applied in hedge funds. To begin with, we can say that there is no standard strategy used in hedge fund investments and someone who wants to invest in a hedge fund can plan his strategy based on elements such as the market, the style, the instrument, the exposure, the sector of the market, the method and the diversification. The primary aim of the different strategies that a hedge fund manager chooses to adopt is to reduce the volatility and the risk and to attempt to preserve capital and deliver constantly positive returns under all the economic conditions that dominate in the market. So recapitulating about hedge fund strategies we could say that it is extremely important for a manager to understand the different strategies that exist between the hedge funds. The return quantities, the variance and the risk exposure are different among the different hedge fund strategies. Some strategies may return as result low risk exposure and stable returns while others might be more risky and have as result high risk and variance and also larger possibilities of capital loss. However, the most common strategy for a hedge fund is the long or short equity strategy which means that the fund can take long or short positions in shares traded on stock exchanges. The managers of a hedge fund go long the undervalued shares and short the overvalued shares. In other words they buy the undervalued shares and sell the overvalued shares.

 https://doi.org/10.26219/heal.aueb.6067

We are now going to describe the high risk that takes someone when he takes the decision to invest in hedge funds. There are many reasons for the increased risk in hedge funds. Some of these reasons are the following:

1) Leverage: This is the case where except the money that have been invested in the hedge fund by the investors the hedge fund borrows money from other sources. Sometimes the hedge funds borrow much more money than the initial capital that has been invested by the investors and a loss of some of the final investment capital might have catastrophic results for the whole initial capital of the investors if the creditors have called in their loans.
2) Short Selling: Short Selling is the case where the hedge fund sells not owned securities hoping to buy them back at a future time at a lower price expecting that their price is going to decrease. Also the hedge fund must pledge to the lender other securities or cash as guarantee for the shorted securities at a price at least equal to the price of the borrowed securities. It is obvious that if a manager of a hedge fund uses short selling as an investment strategy rather than as a hedging strategy the economic results for the hedge fund itself and the investors might be catastrophic if the market turns against it.
3) Appetite for risk: Hedge funds have much more possibilities of taking on investments that have a lot of risk than other types of funds. Such investments are the high yield bonds, the distressed securities etc.
4) Lack of transparency: In hedge funds the whole ambience is too secretive and it is difficult for an investor to be informed about the strategies, the portfolios and the decisions about the different hedge fund's investments. As a consequence, if the investors are not fully and correctly informed about the situations that take place in the hedge funds, the possibility of taking wrong decisions that might yield to loss of capital increases.
5) Lack of regulation: In the hedge funds there is no regulation and some hedge funds might involve risks that are not exposed to the investors.

It is being intelligible that someone has to be a sceptical investor and think of the risk factors and the consequences that his/hers move might involve.

https://doi.org/10.26219/heal.aueb.6067

In the current paragraph we are going to give some details about the structure and the setup of a hedge fund. As it can be concluded from the definition of the Hedge Funds, a Hedge Fund is not a company or a business. We could say that a Hedge Fund is a financial mechanism but not a business. Also a Hedge Fund has no possessions (assets) except from the money that an investor has invested on it. We could say that there exists a portfolio (a pool of capital) consisting of the money that different people invest on the Hedge Fund. Responsible for this portfolio is the manager of the fund. He is the person who manages the whole capital which the investors invested on the Hedge Fund and this person, on the contrary with the Hedge Fund, has employees and also has possessions, which possessions are the business that he is up to. Something that is important to state and concerns the manager's rights is that a Hedge Fund manager has the right to manage more than one Hedge Fund. He is not being restricted on the number of Hedge Funds that he can manage.

In this paragraph we are going to give some details about the benefits of the hedge funds. The use of hedge funds as a private pool of capital which is not being regulated and whose manager has the right to buy or sell any possessions (assets) has revealed many benefits to the world wide economy. First of all, as research has shown, hedge funds have higher returns and lower risk than traditional investment funds. Additively to this first benefit is the fact that the inclusion of hedge funds in a balanced portfolio reduces overall portfolio risk and volatility and increases returns. Someone could say that the fact that the risk is lower in the hedge funds than in other investment funds is too important and guarantees that the probability of loosing a percentage or whole the invested capital decreases and on the contrary the probability of achieving higher returns increases. Also many hedge fund strategies have the ability to generate positive returns in both rising and falling equity and bond markets. Another benefit is that there exist too many different hedge fund investment styles that an investor can choose his strategy in order to meet his investment objectives. Also due to the fact that hedge funds provide an ideal long-term investment solution, someone doesn't have to entry or exit at the correct time from the market.

 https://doi.org/10.26219/heal.aueb.6067

Having finished with the presentation of some basic and introductory information about the Hedge Funds and its operation we shall mention some history-statistical information about the Hedge Funds. Based on the global statistical data about hedge funds it is notable and commonly known that hedge fund activity in the public securities markets has grown substantially, accounting for approximately 10% of all U.S. fixed-income security transactions, 35% of U.S. activity in derivatives with investment-grade ratings, 55% of the trading volume for emerging-markets bonds, and 30% of equity trades. Someone could say that all the affair about hedge funds has become an industry due to the large effect and application that hedge funds have. This hedge fund «industry» is estimated, based on statistical data, to be a 2$ trillion business and year by year grows extremely with high rates of development. There are about 10000 active hedge funds around the world which manage the capital that many people have invested on them. The top 10 hedge funds for the previous years as they were ranked in October 2007 are the ones that follow. Each of the hedge funds is considered in this rank to have as average returns of over 35%.

1. RAB Special Situations Fund (RAB Capital, London): 47.69%
2. The Children's Investment Fund (TCI), (The Children's Investment Fund Management, London): 44.27%
3. Highland CDO Opportunity Fund (Highland Capital Management, Dallas): 43.98%
4. BTR Global Opportunity Fund, Class D (Salida Capital, Toronto): 43.42%
5. SR Phoenicia Fund (Sloane Robinson, London): 43.10%
6. Atticus European Fund (Atticus Management, New York): 40.76%
7. Gradient European Fund A (Gradient Capital Partners, London): 39.18%
8. Polar Capital Paragon Absolute Return Fund (Polar Capital Partners, London): 38.00%
9. Paulson Enhanced Partners Fund (Paulson & Co., New York): 37.97%
10. Firebird Global Fund (Firebird Management, New York): 37.18%

Also based on statistical data found on the web, in the following lines we present some of the most notable hedge fund management companies:

- Amaranth Advisors
- Bridgewater Associates
- Centaurus Energy
- Citadel Investment Group
- D. E. Shaw & Co.
- Fortress Investment Group
- Goldman Sachs Asset Management
- Harbert Management Corporation
- Long Term Capital Management
- Man Group
- Marshall Wace
- Maverick Research Group
- Pirate Capital LLC
- Renaissance Technologies
- SAC Capital Advisors
- Soros Fund Management
- The Children's Investment Fund

Furthermore in the following lines we present the largest hedge funds by assets under management as of March 2008:

- JP Morgan $44.7bn
- Farallon Capital $36bn
- Bridgewater Associates $36bn
- Renaissance Technologies $34bn
- Och-Ziff Capital Management $33.2bn
- Goldman Sachs Asset Management $32.5bn
- DE Shaw $32.2bn
- Paulson and Company $29bn
- Barclays Global Investors $18.9bn
- Man Investments $18.8bn

https://doi.org/10.26219/heal.aueb.6067

- ESL Investments $17.5bn

We could easily say, and the hedge funds investments in the market affirm this, that investors view hedge funds as the phenomenon of the 1990s. From a few hundred players ten years ago, the hedge fund industry has expanded to several thousand today with double-digit annual growth. It is likely and clear that the development of the hedge funds and the interest that many investors show in them will continue in the near future, increasing by this way the role of hedge funds within traditional asset-allocation models.

https://doi.org/10.26219/heal.aueb.6067

Chapter 2

Structural Change Tests

2.1 Introduction

In this chapter we deal with the tests and the statistical methods that are used in statistics and econometrics in order to test a linear regression model for structural changes. These tests are the generalized fluctuation tests, the F-tests and the Bai-Perron method. The generalized fluctuation tests include in particular the CUSUM and MOSUM tests and the fluctuation test while the F-statistic tests include the Chow and the supF tests. The BP method for testing a linear regression model for structural changes is a new statistical technique and its statistical background differs from the fluctuation and the F-tests. The statistical methods that are used for the structural break detection procedure include the Recursive Regression Analysis and the Rolling Regression Analysis. In the following subsections of this chapter we are going to present the linear regression model as well as the tests and methods that are used for the retrospective detection of structural changes in given data sets (historical data) and the tests that are used for the procedure of monitoring a linear regression model when new data arrive over time (future data).

2.2 Linear Model

In this subsection we illustrate and describe the linear regression model in which the structural change tests are applied. The standard linear regression is defined by

$$y_t = x_t^{\ T} \beta_t + u_t \quad (t=1,\ldots n) \qquad (2.1) \quad ,$$

https://doi.org/10.26219/heal.aueb.6067

where y_t is the observation of the dependent variable at time t, $x_t = (1, x_{t2}, ..., x_{tk})^T$ is a k x1 vector of observations of the independent variables at time t, u_t are the residuals that are normally distributed with mean zero and σ^2 variance, and β_t is the k x1 vector of regression coefficients at time t. The null hypothesis of the structural change tests states that there is no structural change in the model while the alternative hypothesis states that there exist structural changes. This statement can be defined by $H_0 : \beta_t = \beta_0$ and $H_1 : \beta_t \neq \beta_0$. An assumption that has to be valid in the linear model is that the regressors must be non-stochastic.

2.3 Retrospective tests for testing historical data for structural breaks

In this subsection we deal with the classical approach of testing for structural changes, which is the application of retrospective tests using historical data set of a given length. That is to say that someone can use historical tests trying to find a structural change ex post. In other words that means that the sample of the problem that is being investigated is fully available in the researcher and he/she has to apply the retrospective tests described below in order to find any possible structural changes in the under testing dataset. Also, in this subsection, except the retrospective structural change tests, we are going to illustrate the rolling regression and the recursive regression analysis which are also used in order to detect any possible structural breaks in the already given full data set.

https://doi.org/10.26219/heal.aueb.6067

2.3.1 CUSUM processes

These processes contain cumulative sums of standardized residuals. As Brown, Durbin, and Evans (1975) have suggested, the cumulative sums of recursive residuals should be computed using the quantity

$$W_n(t) = \frac{1}{\tilde{\sigma}\sqrt{\eta}} \sum_{i=k+1}^{k+[t\eta]} \tilde{u}_i \quad (2.2)$$

where $\eta = n - k$ is the number of recursive residuals and $[t\eta]$ is the integer part of $t\eta$. Recursive residuals are another type of residuals that are used in tests on structural change problems and are defined as

$$\tilde{u}_i = \frac{y_i - x_i^T \hat{\beta}^{(i-1)}}{\sqrt{1 + x_i^T \left(X^{(i-1)^T} X^{(i-1)} \right)^{-1} x_i}} \quad (2.3).$$

For the limiting process of the empirical fluctuation process $W_n(t)$ under the null hypothesis, someone can read the paper of Kramer, Ploberger and Alt (1988). Under the alternative hypothesis, of the existence of a single structural change point t_0, the recursive residuals are going to have zero mean up to t_0 point. Henceforth, the CUSUM process should have zero mean up to the t_0 point and this zero mean is going to be constant in the continuity of the process. As Kramer, Ploberger, and Alt (1988) have shown, the main properties of the CUSUM process remain constant even if the assumptions of the under testing model are weaker, in particular in dynamic models. The boundaries for the Recursive CUSUM process are proportional to the standard deviation function of the corresponding theoretic process. So the boundaries are of the form

$$b(t) = \lambda \cdot \sqrt{t} \quad (2.4)$$

https://doi.org/10.26219/heal.aueb.6067

where λ determines the confidence level. But the boundaries that are commonly used are linear because a closed form solution for the crossing probability is known, so the boundaries for the Recursive-CUSUM process are of type

$$b(t) = \lambda \cdot (1+2t) \quad (2.5).$$

Another alternant of the CUSUM process is the OLS-CUSUM process which is defined by

$$W_n^{\,0}(t) = \frac{1}{\hat{\sigma}\sqrt{n}} \sum_{i=1}^{[tn]} \hat{u}_i \quad (2.6).$$

This process is based on cumulative sums of the common OLS residuals. For the limiting process of the OLS-CUSUM process someone can see the paper of Ploberger and Kramer (1992). One main characteristic of the above process is that it starts in 0 at t=0 and returns to 0 for t=1. Also, calculations of the OLS-CUSUM process have shown that when there is a single structural change in the model, the whole process should have a peak near the point t_0. Lastly, the boundaries for the OLS CUSUM process are proportional to the standard deviation function of the corresponding theoretic process. So the boundaries are of the form

$$b(t) = \lambda \cdot \sqrt{t(1-t)} \quad (2.7)$$

where λ determines the confidence level. But, as in the case of the Recursive CUSUM process, the boundaries that are commonly used are linear because a closed form solution for the crossing probability is known, so the boundaries for the OLS-CUSUM process are of type

$$b(t) = \lambda \quad (2.8).$$

https://doi.org/10.26219/heal.aueb.6067

2.3.2 MOSUM processes

The MOSUM process is a process that is also used in order to test a linear regression model for structural changes. In this process we analyze the moving sums of residuals instead of using cumulative sums of the same residuals. As a result, the process contains the sum of fixed number of residuals in a data window whose size is determined by the bandwidth parameter $h \in (0,1)$ and which is moved over the whole sample period. The Recursive MOSUM process is then defined by

$$M_n(t/h) = \frac{1}{\tilde{\sigma}\sqrt{\eta}} \sum_{i=k+[N_\eta t]+1}^{k+[N_\eta t]+[\eta h]} \tilde{u}_i = W_n\left(\frac{[N_\eta t]+\eta h}{\eta}\right) - W_n\left(\frac{[N_\eta t]}{\eta}\right) \quad (2.9)$$

where

$$N_\eta = \frac{(\eta - [\eta h])}{(1-h)} \quad (2.10).$$

In the same way, calculations have shown that the OLS MOSUM process is defined by

$$M_n{}^0(t/h) = \frac{1}{\hat{\sigma}\sqrt{n}} \sum_{i=[N_n t+1]}^{[N_n t]+[nh]} \hat{u}_i = W_n{}^0\left(\frac{[N_n t]+nh}{n}\right) - W_n^0\left(\frac{[N_n t]}{n}\right) \quad (2.11)$$

where

$$N_n = \frac{(n - [nh])}{(1-h)} \quad (2.12).$$

The limiting process for the empirical MOSUM processes is described in the paper of Chu, Hornik and Kuan (1995b). The technique that the Recursive and the OLS MOSUM process identify the possible structural breaks is that if there exists a structural break at point t_0 there is going to be a shift in the

https://doi.org/10.26219/heal.aueb.6067

processes around this point t_0. Furthermore, the boundaries for the MOSUM process are constants, that is to say that they are of type

$$b(t) = \lambda \quad (2.13),$$

due to the fact that the limiting processes are stationary.

2.3.3 Fluctuation test

Another way of identifying the structural breaks in the under testing model is to use the fluctuation test. Both the Recursive and the OLS MOSUM processes used the residuals in order to test for structural breaks while the simple fluctuation test is based on estimates of the unknown regression coefficients. In the case of the fluctuation test the vector β of the regression coefficients is estimated recursively with a number of observations that increases or with a moving window which has constant bandwidth h. Then these estimates are compared to the estimates that have resulted from the full sample. Ploberger, Kramer, and Kontrus (1989) have shown that the fluctuation process described above is defined by

$$Y_n(t) = \frac{\sqrt{i}}{\hat{\sigma}\sqrt{n}}\left(X^{(i)T}X^{(i)}\right)^{1/2}\left(\hat{\beta}^{(i)} - \hat{\beta}^{(n)}\right) \quad (2.14)$$

where

$$i = [k + t(n-k)] \quad (2.15)$$

with $t \in [0,1]$. The process that was described above is the fluctuation process. Chu, Hornik and Kuan (1995a) were based on the fluctuation process and introduced another method of testing for structural breaks, the Moving Estimates process. The ME process is defined by

https://doi.org/10.26219/heal.aueb.6067

$$Z_n(t/h) = \frac{\sqrt{[nh]}}{\hat{\sigma}\sqrt{n}} \left(X^{([nt],[nh])T} X^{([nt],nh])} \right)^{1/2} \left(\hat{\beta}^{([nt],[nh])} - \hat{\beta}^{n} \right) \quad (2.16)$$

where $0 \le t \le 1-h$. The limiting process of the Fluctuation process and the Moving Estimates process is described in the papers that we mentioned above. In the Recursive Estimates processes if there exists a structural break in the linear regression model then the process has a peak and the Moving Estimates process has a shift around the structural break point t_0.

2.3.4 F-tests

Another way of testing a linear regression model for structural breaks is the test that is based on the F-test statistics. A considerable difference between the F-tests and the generalized fluctuation tests is that the generalized fluctuation tests can be used in order to test a model for more than one single structural breaks while the F-tests are used in order to test the under testing model for the presence or not of only a single structural break. In the F-test the null hypothesis states that there is no structural break in the model while the alternative hypothesis can be defined by

$$\beta_i = \left\{ \begin{matrix} \beta_A, \ 1 \le \mathrm{i} \le i_0 \\ \beta_B, \ \mathrm{i_o} \le \mathrm{i} \le \mathrm{n} \end{matrix} \right\} \quad (2.17)$$

where i_0 is a structural break point in the interval $(k, n-k)$. Chow (1960) suggested such an F-test in 1960 for the case where the structural break point i_0 is known. His idea was to split the full sample in two subsamples and then for each subsample to fit two linear regression models according to the position of the structural break point i_0. So one model is calculated for the sample before the point i_0 and the other model is calculated for the sample after the point i_0. The next step, according to his suggestion, is to reject the null hypothesis of no structural break whenever

https://doi.org/10.26219/heal.aueb.6067

$$F_{i0} = \frac{\hat{u}^T \hat{u} - \hat{e}^T \hat{e}}{\hat{e}^T \hat{e} / (n-2k)} \quad (2.18)$$

is extremely large, where $\hat{e} = (\hat{u}_A, \hat{u}_B)^T$ are the residuals from the full model where the coefficients in the two sub samples are estimated separately, $\hat{u}$ are the residuals from the model that is fitted for all the observations of the sample and k is the number of parameters (regressors). The Chow test states that the structural break point has to be known in advance but there exist tests based on F-statistics that do not require the structural break point to be known. The test statistic F_{i0} has an asymptotic Chi-Square distribution with k degrees of freedom and if the assumption of normality holds for the data F_{i0} / k has an exact F distribution with k and $n-2k$ degrees of freedom. For the case that there exists more than one structural break in the model the methodology is to calculate the F-statistics for all possible breaks and reject the null hypothesis if any of those statistics exceeded some critical values. Some years later, in 1993 and 1994, Andrews (1993) and Andrews and Ploberger (1994) created F-tests which were based on mathematical quantities obtained from the sequence of F statistics. According to these F-tests the estimation of the break-date is the time at which the F-statistic takes its maximum value ($F - stat = \max_{i=1,\dots,t} F_i$).

2.3.5 supF, aveF and expF tests

Andrews (1993) and Andrews and Ploberger (1994) suggested three types of test statistics that are based on different kinds of F statistics—Wald, LM or LR statistics—in models fitted by Generalized Method of Moments. These tests are the supF, aveF and expF statistics. The above tests are considered for the case of pure structural change , in which the entire parameter vector is subject to change under the alternative, and for the case of partial structural change , in which only a component of the parameter vector is subject to change under the alternative. Due to the fact that these test statistics are easy to interpret and possess certain optimality properties

https://doi.org/10.26219/heal.aueb.6067

against single shift alternatives, they are very popular and are probably the most used in practice. Although the asymptotic behaviour for the Wald, LM and LR statistics is the same, only the test based on LM statistics can be embedded into the framework above because this is the only test statistic which is only based on the full sample estimate $\hat{\theta}$. The other two require partial sample estimates before and after a hypothetical breakpoint which is moved over a subset of the sample Π, a closed subset of (0, 1). This supLM statistic according to Andrews (1993) can be transformed as follows:

$$\begin{aligned} \sup F &= \sup_{t\in\Pi} LM(t) = \sup_{t\in\Pi}\left(t\left(1-t\right)\right)^{-1} W_{n}\left(t,\hat{\theta}\right)\hat{J}^{-1} W_{n}\left(t,\hat{\theta}\right) \\ &= \sup_{t\in\Pi}\left(t\left(1-t\right)\right)^{-1} efp(t)^{T} efp(t) \qquad (2.19) \\ &= \sup_{t\in\Pi}\frac{\left\|efp(t)\right\|_2^2}{t\left(1-t\right)}. \end{aligned}$$

where the empirical fluctuation process is formally defined as

$$efp\left(t\right) = \hat{J}^{-1/2} n^{-1/2} \sum_{i=1}^{nt} \psi\left(y_i, x_i, \theta\right) \quad (2.20)$$

where $\hat{J}$ is some suitable consistent estimate of the covariance matrix of the scores $\psi\left(Y_i,\theta\right)$. Also the dependent variable y_i follow some distribution F with k-dimensional parameter θ_i, conditional on the regressors x_i. To define a test statistic based on the empirical fluctuation process, a scalar functional is required that captures the fluctuations in the process. The corresponding limiting distribution is then determined by applying the functional to the limiting process. Closed form solutions exist for the distributions implied by certain functionals. The corresponding critical values can be found by simulations for any kind of functional. As the empirical process is essentially a matrix with n observations over time and k components (usually corresponding to parameters), this functional can typically be split up into a functional λ_{comp} which aggregates over the k components and a functional λ_{time} which aggregates over time. If λ_{comp} is applied first, a univariate process is

https://doi.org/10.26219/heal.aueb.6067

obtained which can be inspected for changes over time. However, applying λ_{time} first results in k independent test statistics such that the component/parameter that causes the instability can be identified. Some common values for λ_{time} are the absolute maximum, the mean or the range and typical functionals λ_{comp} include the maximum norm (or L_∞ norm, denoted as $\|\cdot\|_\infty$) or the squared Euclidean norm (or L_2 norm, denoted as $\|\cdot\|_2^2$). In the present dissertation, in the supF tests the test statistics are all of the form

$$\lambda_{tme}\left(\frac{\lambda_{comp}\left(efp\left(t\right)\right)}{d\left(t\right)}\right) \quad (2.21)$$

where $d(\cdot)$ is a weighting function. Hence, statistics based on LM scores, F statistics and OLS residuals are all based on the same empirical fluctuation process and to only differ in the choice of the functionals λ_{time}, λ_{comp} and the function d. Continuing the presentation of the supLM statistic, we observe in this test statistic, according to (2.20) that the empirical fluctuation process is first aggregated over the components using the squared L_2 norm, weighted by the variance of the Brownian bridge and then aggregated over time using the supremum over the interval Π. That is to say that the null hypothesis is rejected when the L_2 aggregated process crosses the limit

$$b(t) = c \cdot d(t) \quad (2.22)$$

where c determines the significance level. Exactly we have that λ_{comp} is equal to $\|\bullet\|_2^2$, λ_{time} is equal to $\sup_{t\in\Pi}$, and

$$d(t) = t(1-t) \quad (2.23).$$

Consequently the limiting distribution is

https://doi.org/10.26219/heal.aueb.6067

$$\sup_{t\in\Pi} \left(t(1-t)\right)^{-1} \left\|W^0(t)\right\|_2^2 \quad (2.24)$$

The aveLM and expLM can be derived analogously, with the same λ_{comp} and d and replacing only λ_{time} by the average and the exp functional respectively. Another view on the same statistic could be to not use the process efp but

$$\tilde{efp} = \lambda_{LM} efp \quad (2.25)$$

where λ_{LM} is a transformation functional λ_{traf} defined as

$$\left(t(1-t)\right)^{-1} \| \ \|_2^2 \quad (2.26).$$

By doing this transformation we obtain the univariate process of LM statistics which just has to be aggregated over time using the supremum. Andrews (1993) imported this univariate process of LM statistics by establishing the Functional Central Limit Theorem (FCLT) not using the cumulative scores but using the F statistics. For the Wald and LR-based statistics, the same aggregation functionals are used and the limiting distribution is identical, but on the basis of a fluctuation process that requires estimation of the model on various sub-samples. Eventually, according to the theory described above and in the case of the linear models that we are up to in this dissertation the above F tests take the following forms. The supF test is defined by the quantity

$$\sup F = \sup_{\underline{i} \le i \le \bar{i}} F_i \quad (2.27),$$

the aveF by the quantity

$$aveF = \frac{1}{\bar{i} - \underline{i} + 1} \sum_{i=\underline{i}}^{\bar{i}} F_i \quad (2.28)$$

and the expF by the quantity

$$\exp F = \log\left(\frac{1}{\bar{i} - \underline{i} + 1} \sum_{i=\underline{i}}^{i=\bar{i}} \exp\left(0.5 \cdot F_i\right) \right) \quad (2.29).$$

2.3.6 BP Method

The method that Bai and Perron (1998, 2003) suggested is a dynamic programming procedure that helps the researcher to estimate and test for multiple structural breaks in linear regression models. This algorithm is based on the principle of dynamic programming that allows the computation of estimates of the break points as global minimizers of the sum of squared residuals. This algorithm uses at most least-squares operations of order $O(T^2)$ for any number of structural changes m, unlike a standard grid search procedure which would require least squares operations of order $O(T^m)$. The original method works only for pure structural change models but here we propose an alternative that allows estimating more general partial structural change models. The basic idea of the approach becomes fairly intuitive once it is realized that, with a sample of size T, the total number of possible segments is at most $T(T+1)/2$ and is therefore of order $O(T^2)$. In the case of a pure structural change model the computation of the estimates of the model and the estimate $S_T(T_1,...,T_m)$ can be done applying OLS segment by segment without constraints among them. The computation of the triangular matrix of sums of squared residuals can be achieved using standard updating formulae to calculate recursive residuals. Indeed, all the relevant information can be calculated from $T - hm + 1$ sets of recursive residuals. Let $v(i, j)$ be the recursive residual at time j obtained using a sample that starts at date i, and let $SSR(i, j)$ be the sum of squared residuals obtained by applying least-squares to a segment that starts at date i and ends at date j. We have the following recursive relation

$$SSR(i, j) = SSR(i, j-1) + v(i, j)^2 \quad (2.30).$$

https://doi.org/10.26219/heal.aueb.6067

All the relevant information is contained in the values $SSR(i,j)$ for the relevant combinations (i,j). Note that the number of matrix inversions needed is simply of order $O(T)$. Once the sums of squared residuals of the relevant segments have been computed and stored, a dynamic programming approach can be used to evaluate which partition achieves a global minimization of the overall sum of squared residuals. This method essentially proceeds via a sequential examination of optimal one-break (or two segments) partitions. Let $SSR(\{T_{r,n}\})$ be the sum of squared residuals associated with the optimal partition containing r breaks using the first n observations. The optimal partition solves the following recursive problem:

$$SSR(\{T_{m,T}\}) = \min_{mh \le j \le T-h}\left[SSR(\{T_{m-1,j}\}) + SSR(j+1,T)\right] \quad (2.31).$$

The procedure starts by evaluating the optimal one-break partition for all subsamples that allow a possible break ranging from observations h to $T-mh$. Hence, the first step is to store a set of T-(m+1)h+1 optimal one-break partitions along with their associated sum of squared residuals. Each of the optimal partitions correspond to subsamples ending at dates ranging from $2h$ to T-(m-1)h. Consider now the next step which proceeds in a search for optimal partitions with two breaks. Such partitions have ending dates ranging from $3h$ to T-(m-2)h. For each of these possible ending dates, the procedure looks at which one-break partition (saved earlier) can be inserted to achieve a minimal sum of squared residuals. The outcome is a set of T-(m+1)h+1 optimal two breaks (or three segments) partitions. The method continues sequentially until a set of T-(m+1)h+1 optimal $(m-1)$ breaks partitions are obtained with ending dates ranging from $(m-1)h$ to $T-2h$. The final step is to see which of these optimal $(m-1)$ breaks partitions yields an overall minimal sum of squared residuals when combined with an additional segment. This procedure is being done using significance tests or using some information based criteria such as the Akaike Information Criteria (AIC) or the Bayesian Information Criteria (BIC). By this procedure that we described above, the

https://doi.org/10.26219/heal.aueb.6067

optimal numbers of breaks are being selected. In this dissertation we decide the number of breaks using the BP method based on the BIC criterion which, by definition, states that the number of breaks is the number in which the BIC criterion takes its smaller value.

2.3.7 Rolling Regression Analysis

In this method, the Rolling Regression analysis, we are able to estimate the model's stability over time. This fact about the stability of the model coefficients over time is a basic assumption when analyzing data with such models. We can paraphrase this assumption and say that the parameters of the testing model should be constant over time. However, it is obvious and clear that such an assumption is hard to be true because the data of each case that we deal with often change dramatically and fast. So in this case the assumption about constant parameters is violated. In order to avoid this problem and assess the constancy of the parameters of the testing models the rolling regression method implies that we compute the parameters estimates over a rolling window of a fixed sized through our sample. If there is going to exist any structural change in the sample the rolling estimates should be different while if there is not going to exist any structural change, so the parameters are constant over the sample, the rolling estimates should not be statistically different. Here, in our case, we are going to use the rolling regression analysis to backtest the different structural change models on historical data. Such an application, that is to say on testing the historical data, is of the most common uses of the rolling regression. The algorithm that describes this idea and application is the following. The first step is to split the sample, the historical data, in the estimation sample and the prediction sample. The first sample, the estimation sample, is going to be used in order to estimate the parameters of the under testing model and the second sample, the prediction sample, contains the observations that are used in the under testing estimated model in order to compute-predict the new parameter estimates. In other words we fit the model using the estimation sample and later we make h future predictions for the prediction sample and also compute

https://doi.org/10.26219/heal.aueb.6067

the corresponding prediction errors. Then we roll again the estimation sample after we have made an increment and this procedure is iterated until we cannot make any h future predictions. We have to notice at this point that the collection of the prediction errors that were computed above by the algorithm are summarized and used to evaluate the statistical model. In other words that means that all the rolling regression estimates for the residual standard error of the under testing model are computed and this data series of the standard errors is being tested for structural breaks. If there occurs a structural change in the series of the standard errors, the under testing model is definitely going to be unstable over time. If there isn't any break in the standard error data series, then the model is stable over time and there is no structural change in the data.

All the above statements can be summarized and defined by the following expressions. For a given window $n < T$ the rolling linear regression model is defined by

$$y_t(n) = X_t(n)\beta_t(n) + \varepsilon_t(n) \quad (2.32)$$

for $t = n,...,T$ where $y_t(n)$ is an $(nx1)$ vector of observations of the dependent variable, $X_t(n)$ is an (nxk) matrix of the explanatory variables, $\beta_t(n)$ is an $(kx1)$ vector of the regression coefficients and $\varepsilon_t(n)$ is an $(nx1)$ vector of the error terms. The n observations in $y_t(n)$ and $X_t(n)$ are the n most recent observations from $t-n+1$ to t. Also in the above model we assume that $n > k$. The rolling least squares estimates are the following:

- $\hat{\beta}_t(n) = \left[X_t(n)' X_t(n)\right]^{-1} X_t(n)' y_t(n)$ (2.33)

- $\hat{\sigma}_t^2 = \frac{1}{n-k}\hat{\varepsilon}_t(n)'\hat{\varepsilon}_t(n) = \frac{1}{n-k}\left[y_t(n) - X_t(n)\hat{\beta}_t(n)\right]'\left[y_t(n) - X_t(n)\hat{\beta}_t(n)\right]$ (2.34)

 https://doi.org/10.26219/heal.aueb.6067

2.3.8 Recursive Regression Analysis

In the case of the recursive regression analysis, the method for testing for structural breakpoints states that we have to estimate the model

$$y_t = \beta_t^{'} x_t + \varepsilon_t \quad (2.35)$$

by least squares recursively for $t = k+1,...,T$ giving $T-k$ recursive least squares (RLS) estimates $\left(\hat{\beta}_{k+1},...,\hat{\beta}_T\right)$. If the parameter β is constant then the recursive estimates should tend to a specific value rapidly. If some of the components in β are not constant the RLS estimates are about to show instability. A simple way of testing if the parameter β is not constant is to plot the Recursive Least Squares estimates $\hat{\beta}_{it}$ ($i = 0,...,k$) and examine them for stability or instability. Brown, Durbin and Evans (1976) proposed two tests for parameter instability in the case of the recursive regression analysis. These tests are the CUSUM and CUSUMSQ tests. Both tests have as common quantity the standardized 1-step ahead recursive residuals

$$\hat{w}_t = \frac{\hat{\varepsilon}_t}{\hat{f}_t} = \frac{y_t - \hat{\beta}'_{t-1} x_t}{\hat{f}_t} \quad (2.36) ,$$

where $\hat{f}_t^2$ is an estimate of the recursive error variance

$$\hat{f}_t^2 = \sigma^2 \left[1 + x_t^{'} \left(X_{t-1}^{'} X_{t-1}\right)^{-1} x_t\right] \quad (2.37)$$

and X_t is the txk matrix of observations on x_s where $s = 1,...,t$. The CUSUM test is defined by

$$CUSUM_t = \sum_{j=k+1}^{t} \frac{\hat{w}_j}{\hat{\sigma}_w} \quad (2.38).$$

 https://doi.org/10.26219/heal.aueb.6067

It is clear that the CUSUM test is based on the cumulated sum of the standardized recursive residuals where $\hat{\sigma}_w$ is the sample standard deviation of $\hat{w}_j$. Under the null hypothesis the test statistic $CUSUM_t$ has mean zero and variance that is proportional to the quantity t-k-1. As Brown, Durbin and Evans (1976) have shown, the approximate 95% intervals for the $CUSUM_t$ statistic are the ones that are given by the lines which connect the points

$$\left(k, \pm 0.948\sqrt{T-k-1}\right) \quad (2.39)$$

and

$$\left(T, \pm 0.948 \cdot 3\sqrt{T-k-1}\right) \quad (2.40).$$

If some of the values or all the values of the test statistic $CUSUM_t$ lay outside of these intervals then the null hypothesis is rejected. Additionally, the CUSUMSQ test is based on the quantity of the cumulative sum of the squared standardized recursive residuals. This test is defined by

$$CUSUMSQ_t = \frac{\sum_{j=k+1}^{t} \hat{w}_j^2}{\sum_{j=k+1}^{T} \hat{w}_j^2} \quad (2.41).$$

Brown, Durbin and Evans (1976), as with previous test, have shown that the 95% confidence intervals for the $CUSUMSQ_t$ test statistic have the form

$$c \pm t / \sqrt{T-k-1} \quad (2.42).$$

The null hypothesis is rejected if the values of the $CUSUMSQ_t$ statistic lay outside the confidence intervals.

https://doi.org/10.26219/heal.aueb.6067

2.4 The Generalized Fluctuation Tests for Monitoring Data Series for Structural Breaks

In the procedure of monitoring structural changes we deal with the detection of parameter instabilities online in incoming data. Especially, this means that after the history period of observations $1,\ldots,n$ (corresponding to $t \in [0,1]$) where the parameters are assumed to be stable $\theta_i = \theta_0$, we test whether they remain stable for any other incoming observations $i > n$ (the on-line monitoring period, corresponding to $t > 1$). The end of this monitoring period might be infinity, but also might be limited to some finite $T > 1$ or $N = \lfloor nt \rfloor$. The theory of monitoring structural changes in linear regression models was introduced by Chu et al. (1996), who used fluctuation processes based on recursive residuals and recursive estimates. Their test was extended by Leisch et al. (2000) to general estimates-based processes. Also, Zeileis et al. (2005) discuss several extensions in the context of dynamic econometric models including processes based on OLS residuals and new boundary functions. Zeileis et al. (2005) showed that there are different approaches to the application of monitoring a dataset. The most intuitive is probably in a policy intervention setting where it should be assessed if and when a known intervention becomes effective. In such a situation, it is plausible to establish a fitted model once before the intervention and then compare the incoming data with this fitted model. Another application might be diagnostic checking of a model which is actively used for data analysis during the monitoring period. Here, the practitioner typically wants to update the model with every incoming observation which leads naturally to the recursive/moving estimates monitoring tests that can be carried out with virtually no additional computations. In our case we deal with the second pre-described application, the one that states that there exists a model during the monitoring period which describes our data and we update this model with new incoming observations trying to predict the position of the possible structural break.

https://doi.org/10.26219/heal.aueb.6067

2.4.1 Estimates-based processes

Chu et al. (1996) extended the fluctuation test, especially the RE (recursive estimates) test, to the monitoring case. They suggested using the recursive estimates process

$$Y_n(t) = \frac{i}{\hat{\sigma}\sqrt{n}}\left(\frac{\left(X_{(n)}{}^T X_{(n)}\right)}{n}\right)^{1/2}\left(\hat{\beta}^{(i)} - \hat{\beta}^{(n)}\right) \quad (2.43)$$

where

$$i = \lfloor k + t(n-k) \rfloor \quad (2.44)$$

and $t \geq 0$ is standardized time relative to the history sample (i.e., $t = 1$ corresponds to $i = n$), and to reject the null hypothesis whenever (one component of) the process $Y_n(t)$ comes over the limits $\pm b_1(t)$ where

$$b_1(t) = \sqrt{t(t-1)\left[\lambda^2 + \log\left(\frac{t}{t-1}\right)\right]} \quad (2.45)$$

in the monitoring period $1 < t < T$ and λ determines the significance level of this procedure, or equivalently when $\max_i \left|Y_{in}(t)\right|$, $i = 1,...,k$, comes over $b_1(t)$. Both $1 < t < T$ and $n < i < Tn$ are said to be the monitoring period as they coincide to the same observations. Another special case of this class of tests is the ME (moving estimates) test which uses estimates from a moving data window of fixed width, i.e.

$$Z_n(t/h) = \frac{[nh]}{\hat{\sigma}\sqrt{n}}\left(\frac{\left(X_{(n)}{}^T X_{(n)}\right)}{n}\right)^{1/2}\left(\hat{\beta}^{(\lfloor nt \rfloor - \lfloor nh \rfloor, \lfloor nh \rfloor)} - \hat{\beta}^{(n)}\right) \quad (2.46)$$

https://doi.org/10.26219/heal.aueb.6067

and rejects the null hypothesis if (one component of) the process comes over the limit $\pm c(t)$, where

$$c(t) = \lambda \cdot \sqrt{\log_{+} t} \quad (2.47)$$

in the monitoring period $1 < t < T$. Here we have that $\log_{+}(t)$ is equal to 1 for $t \leq e$ and equal to $\log t$ otherwise. In theory the end of the monitoring period T may be infinity but sometimes using a finite T is more natural because the monitoring period is known in advance. In that way no size is lost for an infinite monitoring period on $[t, \infty)$.

2.4.2 Residual-based processes

Fluctuation tests for monitoring can also be based on residuals. Chu et al. (1996) considered this case although they focused on the recursive estimates approach. In this subsection we are going to illustrate monitoring processes based on the OLS residuals. The OLS residual- and estimates-based types of tests are equivalent in the case where there is only a constant regressor. The idea is as following. The regression coefficients are estimated once for the history period and based on these estimates the residuals of the observations in the monitoring period are computed. If there is a structural change in the monitoring period the residuals are going to deviate methodically from their zero mean. Thus, we introduce monitoring processes based on the OLS residuals

$$\hat{u}_i^{(n)} = y_i - x_i^T \hat{\beta}^{(n)} \quad (2.48).$$

The OLS-based CUSUM process for monitoring is then defined as

$$W_n^0(t) = \frac{1}{\hat{\sigma}\sqrt{n}} \sum_{i=1}^{\lfloor nt \rfloor} \hat{u}_i^{(n)} \quad t \geq 0 \quad (2.49).$$

https://doi.org/10.26219/heal.aueb.6067

The OLS-based MOSUM process for monitoring is defined analogously as

$$M_n^{\;0}(t/h) = \frac{1}{\hat{\sigma}\sqrt{n}} \sum_{i=\lfloor \eta t \rfloor - \lfloor nh \rfloor + 1}^{\lfloor \eta t \rfloor} \hat{u}_i = W_n^{\;0}\left(\frac{\lfloor \eta t \rfloor}{n}\right) - W_n^{0}\left(\frac{\lfloor \eta t \rfloor - \lfloor nh \rfloor}{n}\right) \quad (2.50)$$

where

$$\eta = \frac{\left(n - \lfloor nh \rfloor\right)}{\left(1-h\right)} \quad (2.51).$$

The limiting process for the OLS-based CUSUM and MOSUM process is the 1-dimensional special case of the k-dimensional recursive and moving estimates process. The respective empirical processes are in fact equivalent if $x_i = 1$ for all i. Thus, the limits given in the previous subsection can be used as well for the OLS-based processes. The advantage of estimates-based processes is that there is a process for each regression coefficient. Consequently, it can be determined which coefficient(s) is (are) responsible for the rejection of the null hypothesis. On the contrary, the OLS based processes are much easier to compute because a linear model has to be fit only once for the whole process and not in every single step and then just residuals have to be computed.

https://doi.org/10.26219/heal.aueb.6067

 https://doi.org/10.26219/heal.aueb.6067

Chapter 3

Application of the Simple, the Recursive and the Rolling Regression Analysis to simulated data

3.1 Introduction

In this 3rd chapter of the present dissertation we are going to discuss the cases and so the corresponding models, in which structural breaks occur. Especially we are going to note briefly the different cases where a structural break occurs in a time series and later, in each case, we are going to simulate the corresponding structural break model. Also we are going to apply some different methods for estimating and testing for breaks in the above models, such as the F-test in the Simple Regression analysis, the Recursive Regression analysis method and the Rolling Regression analysis method. At this point we shall say that we are going to apply the above techniques in the cases that there occurs only one structural break in a time series. The data that we are going to use in order to show the above structural break cases are going to be simulated.

3.2 Introduction in the structural breaks

First of all, before continuing in the presentation of the different structural break cases we are going to give a full definition about the structural breaks. A structural break is a change that occurs in a time series in a certain time period. This change could be due to the change in the mean, in the variance, in the mean and the variance at the same time of the generating process of the time series that we are dealing with. Also another case of a

https://doi.org/10.26219/heal.aueb.6067

structural break is when the slope of a linear model changes somewhere in a certain time period. The above cases of the structural change presence are depicted, in order to be fully understandable, in the following plots.

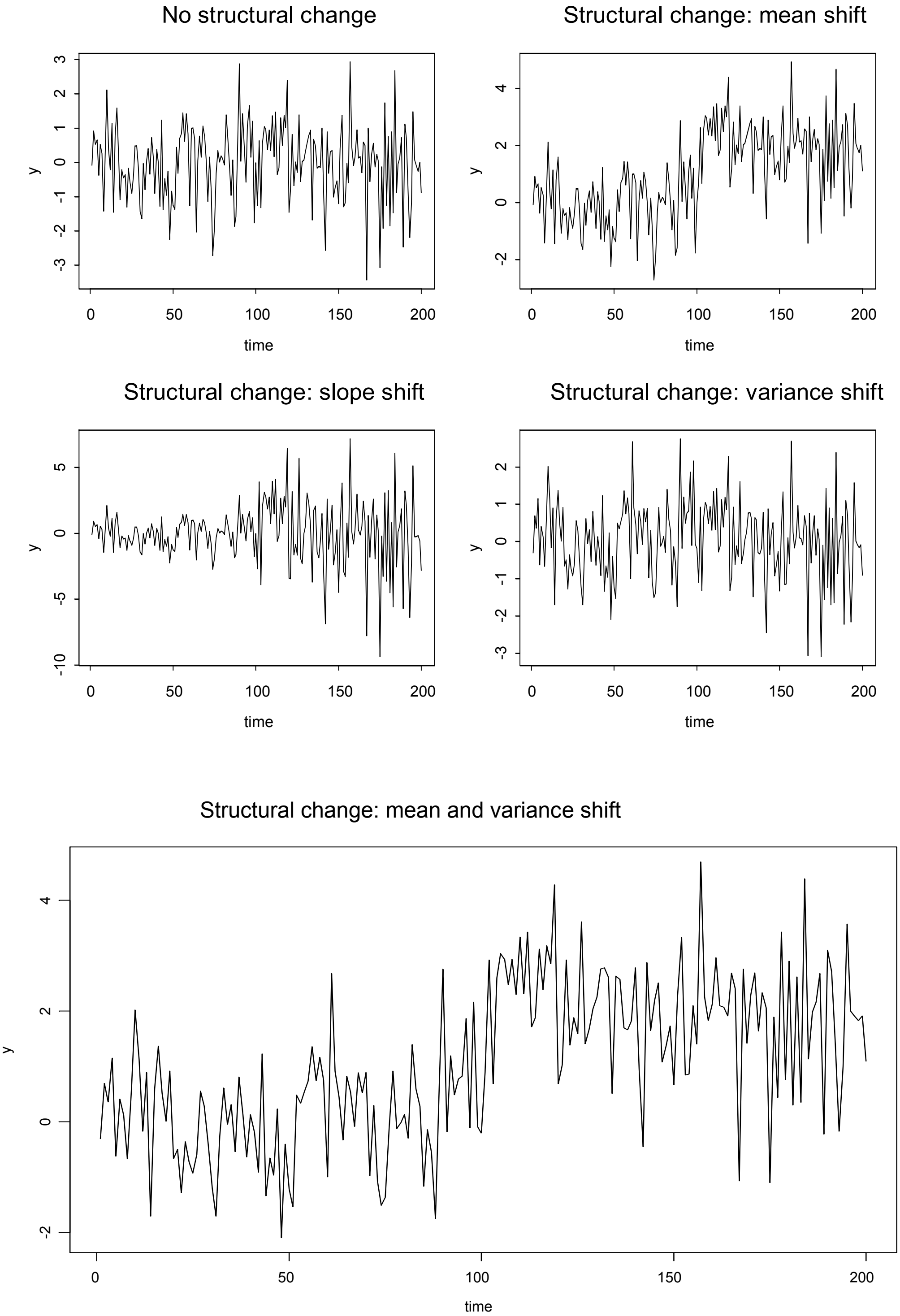


Figure 3.1 Cases of structural changes

https://doi.org/10.26219/heal.aueb.6067

After showing the possible cases of the structural breaks we are going to consider these cases by simulating data for each case and then to interpret the corresponding models. We shall use three different methods for estimating these models for every case of a structural break presence. These methods are the Simple Regression method (Ordinary Least Squares), the Rolling Regression analysis and the Recursive Regression analysis.

3.3 Simple Regression Method

Regression analysis is a statistical tool that utilizes the relation between two or more quantitative variables so that one variable can be predicted from another. This relation is statistical, which means that this relation has error associated with it. The simple linear regression model is defined by

$$Y_t = a + \beta X_t + \varepsilon_t \quad (3.1)$$

where α, β and ε_i are parameters and Y_t and X_t are measured variables. We estimate the simple regression model via the Ordinary Least Squares and the corresponding estimates are the following:

- $\hat{\alpha} = \overline{Y} - \beta \overline{X}$ (3.2)
- $\hat{\beta} = \dfrac{\sum\left(X_i - \overline{X}\right)\left(Y_i - \overline{Y}\right)}{\sum\left(X_i - \overline{X}\right)^2}$ (3.3)

These estimates are unbiased and have minimum variance among all other linear estimators. In the following cases of structural changes, wherever there exists a structural change we estimate the corresponding breakdate using the F-test. Hence, we estimate the parameters of the linear model for the segments that are generated due to the possible existence of the structural change.

https://doi.org/10.26219/heal.aueb.6067

3.3.1 No structural change data

In this case we are going to apply the simple linear regression analysis in data in which there is no structural change. The data consist of the variable y and the variable x. The variable y is defined by

$$y_t = x_t + \varepsilon_t \quad (3.4)$$

where x is the variable which is constituted of 200 observations from the normal distribution with zero mean and variance equal to one and ε_t are the residuals which come from the normal distribution with zero mean and standard deviation equal to 0.5. The whole analysis is applied to the variable y, as it is defined above, and the variable x. The results of this application are shown in the following table and plot.

	Full Sample
α	0.0467 (0.0348)
β	1.0023 (0.0364)*
σ^2	0.2402

* Statistically significant

Table 3.1 Results of the simple linear regression model in the case of no structural change

In the above table we see the estimates of the intercept, the slope and the variance of the simple linear model that we are up to in this case.

 https://doi.org/10.26219/heal.aueb.6067

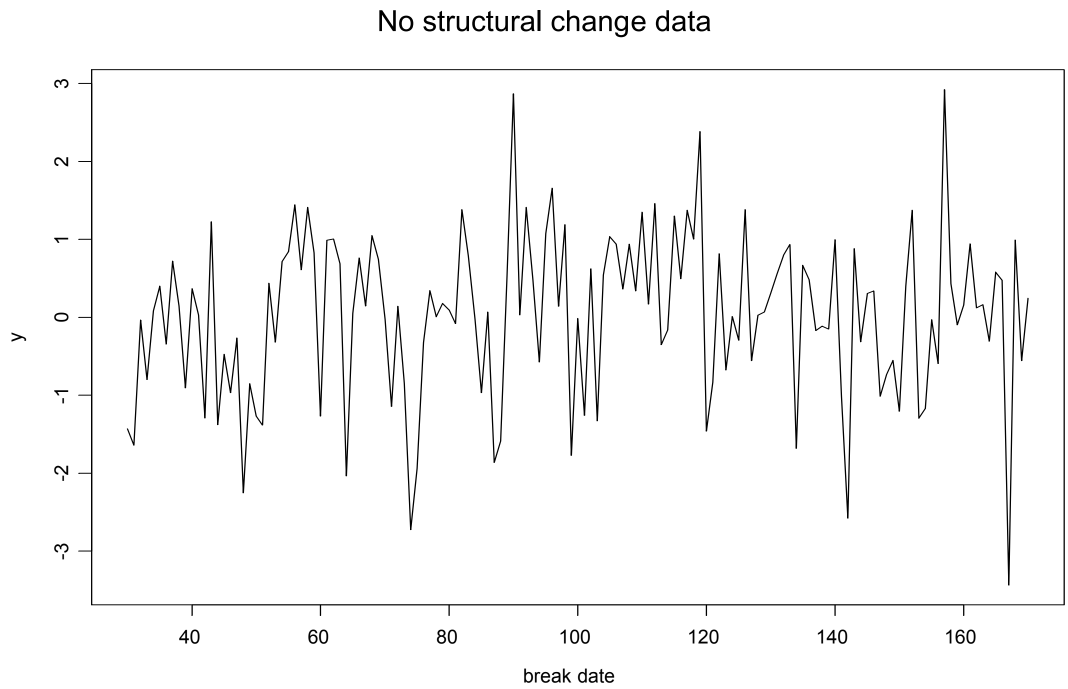


Figure 3.2 Plot of the case where there is no structural change

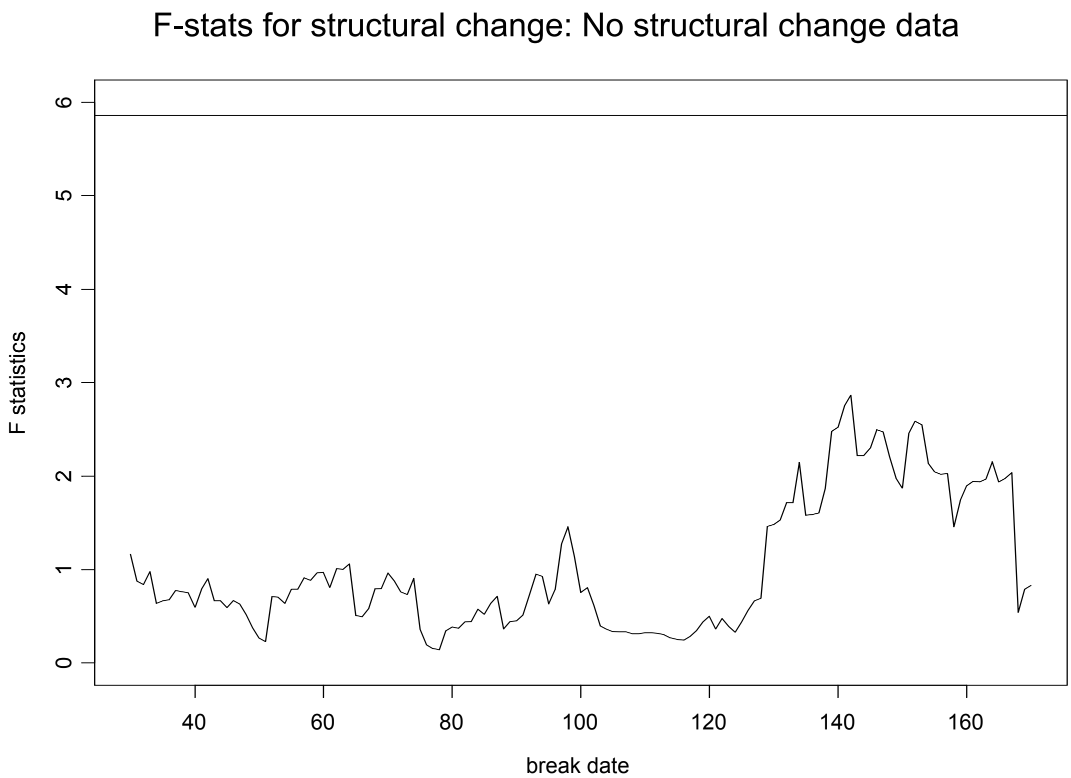


Figure 3.3 Plot of the F-statistics of the data in which there is no structural change

 https://doi.org/10.26219/heal.aueb.6067

Furthermore in the above figures we see that there is no structural change in the data that we examine in this subsection because the F-statistics for each observation are smaller than the critical values depicted by the straight line around the area of the value 6.

3.3.2 Structural change in mean

In this case we deal with the application of the simple linear regression analysis in data in which there exists a structural change in the mean. The data consist of the variable y and the variable x. The variable y is defined by

$$y_t = x_t + \varepsilon_t + 2 \cdot I(t > 100) \quad (3.5)$$

where x is the variable which is constituted of 200 observations from the normal distribution with zero mean and variance equal to one and ε_t are the residuals which come from the normal distribution with zero mean and standard deviation equal to 0.5. The simple linear regression analysis is applied to the above variable y and the variable x. The results of this application are shown in the following plots and table.

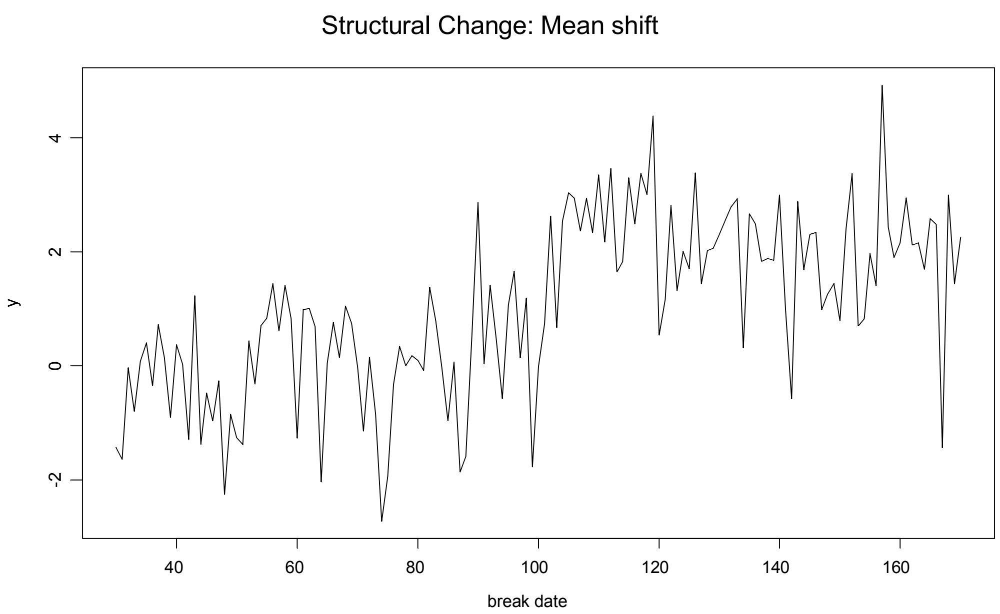


Figure 3.4 Plot of the case where there is structural change in the mean

https://doi.org/10.26219/heal.aueb.6067

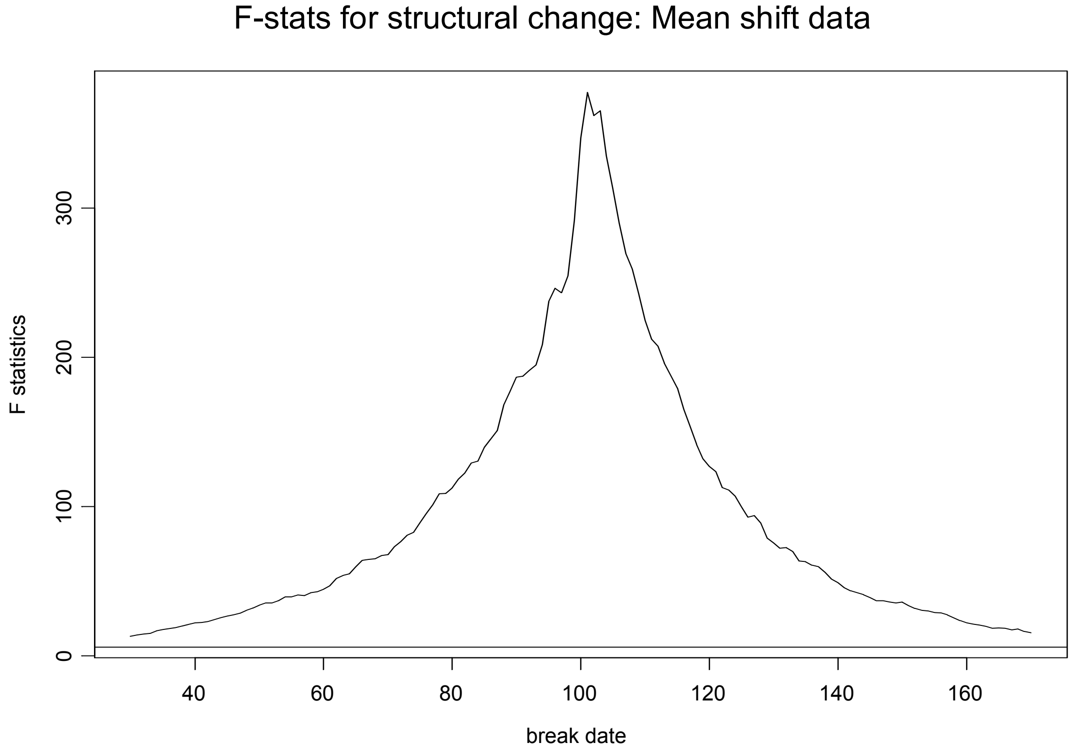


Figure 3.5 Plot of the F-statistics of the data in which there is structural change in the mean

In the above plots we see the plot of the time series that we examine in the present case, the case of the structural change in the mean of the data, and the plot of the F-statistics. As we have said in the introduction, we use the F-test in order to test the data for structural breaks and if there exist such breaks, to identify the position in which they occur. From the plot of the F-statistics we conclude that there exists a break in the 101st observation of the data because in this observation there is observed the maximum value of the F-statistics series, assumption which according to the theory of the F-statistics shows that at this point there is observed a structural break. In the same way and using the assumption of the maximum value of the F-statistics series we are going to analyze the plots of the F-statistics that follow in the following subsections of the present chapter. Having identified that there exists a break we shall estimate the parameters of the model for the two segments that are created due to this break. These parameters are shown in the following table.

https://doi.org/10.26219/heal.aueb.6067

	First Segment	Second Segment
α	0.05512 (0.04769)	1.97129 (0.04956)*
β	1.00800 (0.05232)*	1.03298 (0.04497)*
σ^2	0.2271	0.2443
Break detected at observation #101		

* Statistically significant

Table 3.2 Results of the simple linear regression model in the case of structural change in the mean

Inspecting the above table we conclude that all the parameters of the model are higher in the second segment than in the first segment.

3.3.3 Structural change in slope

In this case we are going to apply the simple regression analysis in data in which there exists a structural change in the slope. The variable y of the data is defined by

$$y_t = x_t + \varepsilon_t + 2 \cdot I(t > 100) \cdot x_t \quad (3.6)$$

where x is the variable which is constituted of 200 observations from the normal distribution with zero mean and variance equal to one and ε_t are the residuals which come from the normal distribution with zero mean and standard deviation equal to 0.5. The results of this application to the variables y and x are shown in the following plots and table.

https://doi.org/10.26219/heal.aueb.6067

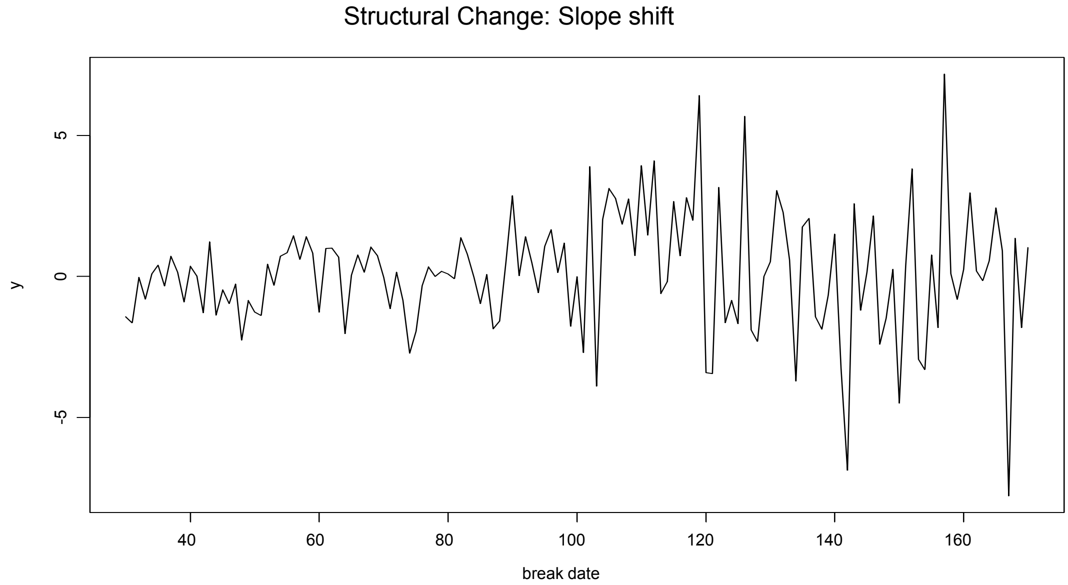


Figure 3.6 Plot of the case where there is structural change in the slope

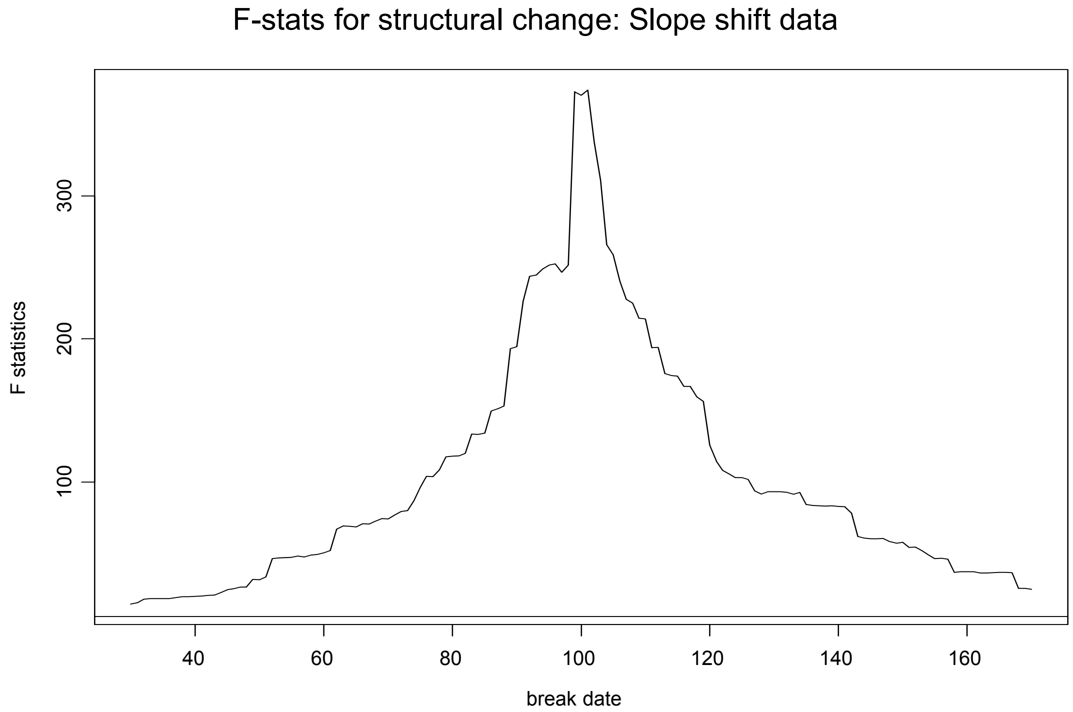


Figure 3.7 Plot of the F-statistics of the data in which there is structural change in the slope

In the above plots we see the plot of the time series that we examine in the present case, the case of the structural change in the slope, and the plot of the F-statistics. From the plot of the F-statistics we conclude that there exists a break in the 101^{st} observation of the data. The estimates of the parameters of

 https://doi.org/10.26219/heal.aueb.6067

the model for the two segments that are created because of the structural change are shown in the following table.

	First Segment	Second Segment
α	0.0404 (0.0487)	-0.0526 (0.0504)
β	0.9900 (0.0488)*	2.9307 (0.0482)*
σ^2	0.2373	0.2529
Break detected at observation #101		

* Statistically significant

Table 3.3 Results of the simple linear regression model in the case of structural change in the slope

It is obvious from the above table that there are substantial differences between the regression coefficients. There is a significant decrease in the values of intercept and a significant increase in the values of slope. Also there is an increase in the values of the variance between the two segments.

3.3.4 Structural change in variance

In this case we concern with the application of the simple linear regression analysis in data in which there exists a structural change in the variance of the error term. The data consist of the variable y and the variable x. The variable y is computed according to the transformation

$$y_t = x_t + \varepsilon_t^{\,2} \cdot I(t < 100) + \sqrt{0.5} \cdot \varepsilon_t \cdot I(t > 100) \quad (3.7)$$

where x is the variable which is constituted of 200 observations from the normal distribution with zero mean and variance equal to one and ε_t are the residuals which come from the normal distribution with zero mean and standard deviation equal to 0.5. The results of this application in the under testing variables y and x are shown in the following plots and table.

https://doi.org/10.26219/heal.aueb.6067

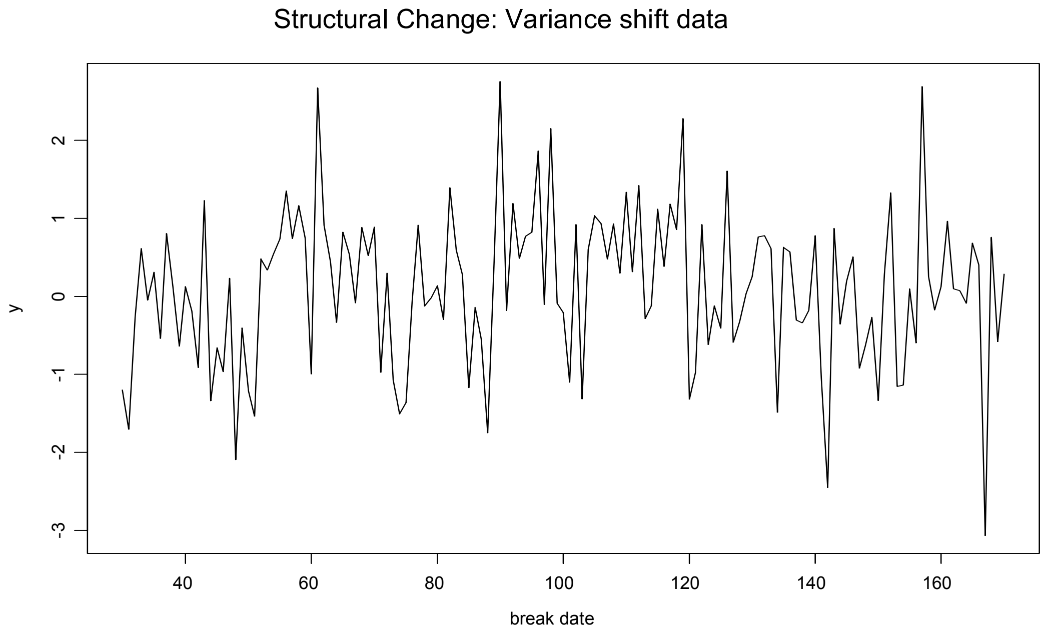


Figure 3.8 Plot of the case where there is structural change in the error variance

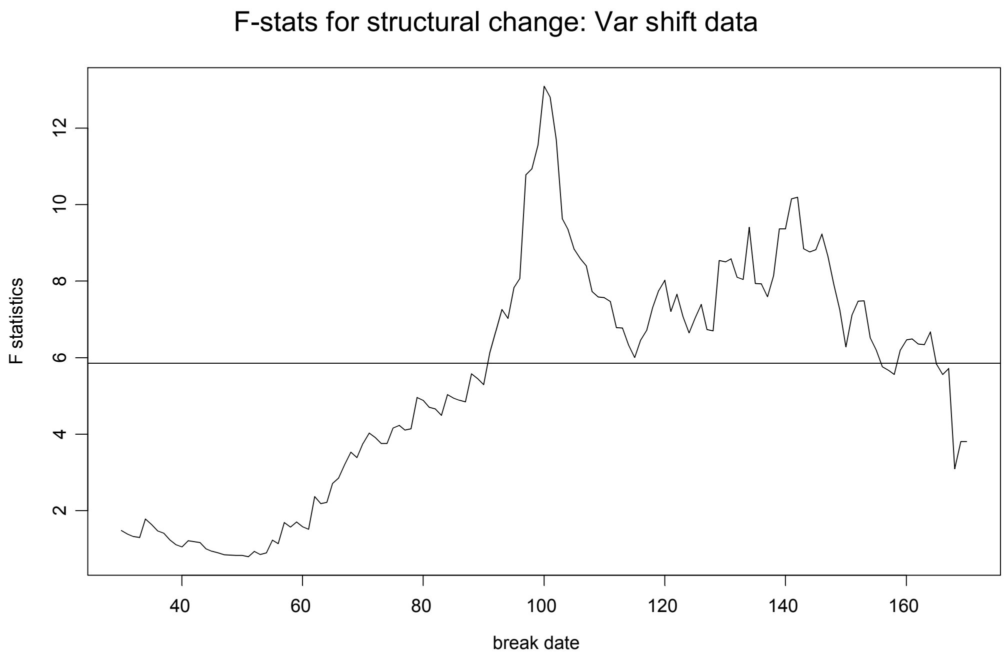


Figure 3.9 Plot of the F-statistics of the data in which there is structural change in the error variance

In the above plots we see the plot of the time series that we examine in the present case, the case of the structural change in error variance, and the plot of the F-statistics. From the plot of the F-statistics we conclude that there

https://doi.org/10.26219/heal.aueb.6067

exists a break in the 100^{th} observation of the data. The estimates of the parameters of the model for the two segments that are created due to the existence of the structural change are shown in the following table.

	First Segment	Second Segment
α	0.3121 (0.0462)*	-0.0474 (0.0341)
β	1.0580 (0.0504)*	0.9982 (0.0373)*
σ^2	0.2114	0.1296
Break detected at observation #100		

* Statistically significant

Table 3.4 Results of the simple linear regression model in the case of structural change in the variance

Going through the above table we conclude that there is a decrease in the intercept and the slope values between the two segments. Also there is a decrease in the variance between the two segments.

3.3.5 Structural change in the mean and the variance

In this case we are going to apply the simple linear regression analysis in data in which there exists a structural change in the mean of the data and the variance of the error term. The variable y of the data is defined by

$$y_t = x_t + 2 \cdot I(t > 100) + \varepsilon_t^{\,2} \cdot I(t < 100) + \sqrt{0.5} \cdot \varepsilon_t \cdot I(t > 100) \quad (3.8)$$

where x is the variable which is constituted of 200 observations from the normal distribution with zero mean and variance equal to one and ε_t are the residuals which come from the normal distribution with zero mean and standard deviation equal to 0.5. The simple linear regression analysis is applied to the above variable y and the variable x. The results of this application are shown in the following plots and table.

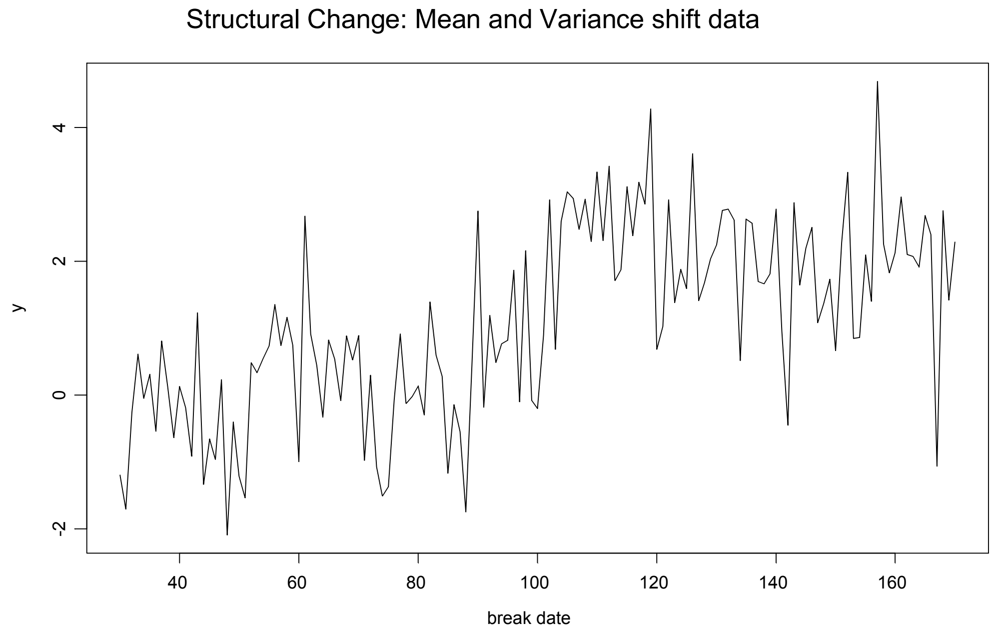


Figure 3.10 Plot of the case where there is structural change in the mean and the error variance

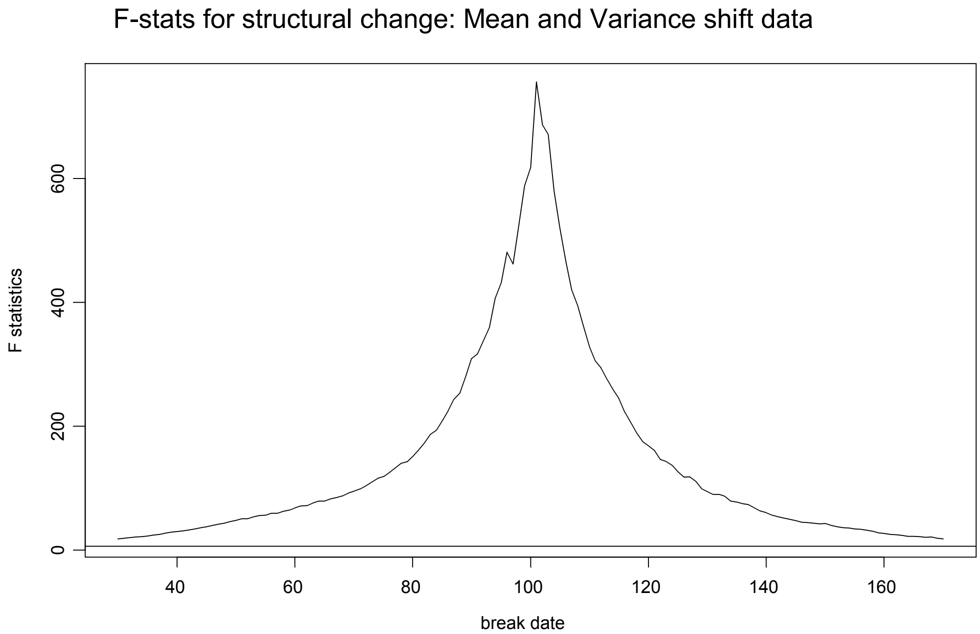


Figure 3.11 Plot of the F-statistics of the data in which there is structural change in the mean and the error variance

In the above plots we see the plot of the time series that we examine in the present case, the case of the structural change in the mean of the data and the error variance, as well as the plot of the F-statistics. From the plot of the F-

https://doi.org/10.26219/heal.aueb.6067

statistics we conclude that there exists a break in the 101st observation of the data. The estimates of the parameters of the model for the two segments that are created because of the existence of this structural change are shown in the following table.

	First Segment	**Second Segment**
α	0.2331 (0.0284)*	2.0019 (0.0357)*
β	1.0241 (0.0315)*	1.0031 (0.0355)*
σ^2	0.07868	0.1275
Break detected at observation #101		

* Statistically significant

Table 3.5 Results of the simple linear regression model in the case of structural change in the mean and the error variance

Inspecting the above results we conclude that there is a significant increase in the values of the intercept between the two segments and a small decrease in the values of the slope. Also there is an increase in the variance of the two models between the two segments.

3.4 Rolling Regression Analysis

In this subchapter of the third chapter we deal with the application of the rolling regression analysis in the different cases of the existence of structural breaks that we have described in the beginning of the chapter. Also, at this point we have to mention that in the following cases the rolling regression estimates are computed for 176 rolling windows of width 25 and increment 1.

https://doi.org/10.26219/heal.aueb.6067

3.4.1 No structural change data

In this case we are going to apply the rolling regression analysis in data in which there is no structural change. The data consist of the variable y and the variable x. The variable y is defined by

$$y_t = x_t + \varepsilon_t \quad (3.9)$$

where x is the variable which is constituted of 200 observations from the normal distribution with zero mean and variance equal to one and ε_t are the residuals which come from the normal distribution with zero mean and standard deviation equal to 0.5. The rolling regression analysis is applied to the above variable y and the variable x. In the following lines we present the plots with the rolling coefficients and the rolling scale. From the graphs of the rolling estimates it is clear that $\hat{\alpha}$ is not significantly different from zero and the $\hat{\beta}$ values are also significantly different from the value zero. Inspecting the graphs more accurately is obvious that the $\hat{\alpha}$ and $\hat{\beta}$ values are quite stable and indicate that the quantities measured by the $\hat{\alpha}$ and $\hat{\beta}$ coefficients are constant over time. The residuals scale estimates $\hat{\sigma}$ increase between the 30th and the 70th observation and then decrease until the 100th observation. After the 100th observation there is taking place an increase in the residuals scale estimates $\hat{\sigma}$ until almost the end of the sample where there is noticed a small decrease. So, in other words, inspecting the plot of the residuals scale estimates $\hat{\sigma}$ and especially the range of the values that these increases and decreases take place we can conclude that the residuals scale estimates $\hat{\sigma}$ are stable and this fact indicates, as well as the $\hat{\alpha}$ and $\hat{\beta}$ coefficients do, that there is no structural change in the under testing variables.

https://doi.org/10.26219/heal.aueb.6067

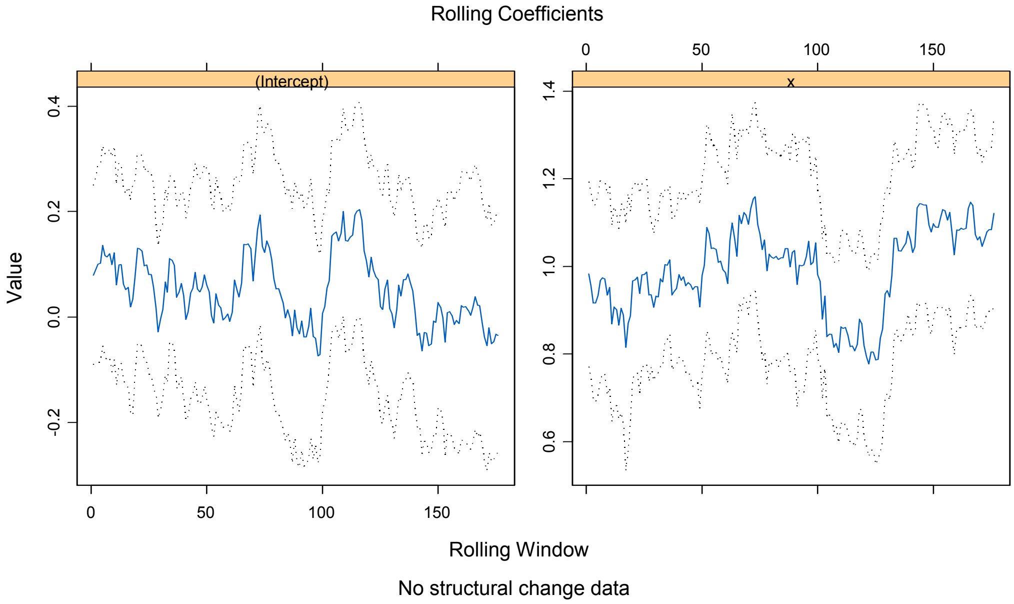


Figure 3.12 Rolling Regression coefficient estimates for the model with no structural change

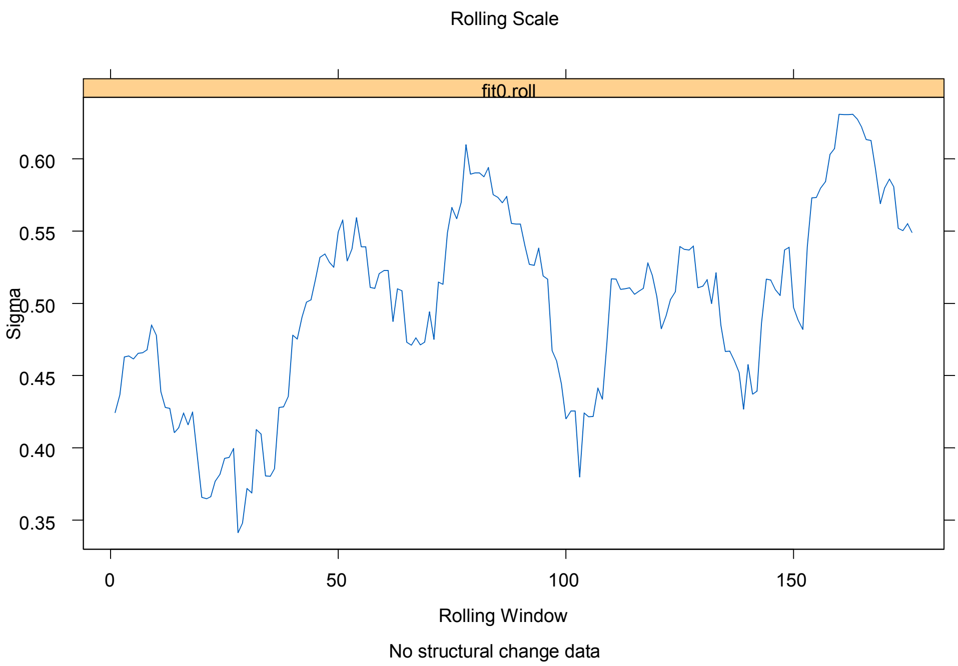


Figure 3.13 Rolling Regression estimates for the residual standard error of the model with no structural change

https://doi.org/10.26219/heal.aueb.6067

3.4.2 Structural change in mean

In this case we are going to apply the rolling regression analysis in data in which there exists a structural change in the mean. The data consist of the variable y and the variable x. The variable y is defined by

$$y_t = x_t + \varepsilon_t + 2 \cdot I(t > 100) \quad (3.10)$$

where x is the variable which is constituted of 200 observations from the normal distribution with zero mean and variance equal to one and ε_t are the residuals which come from the normal distribution with zero mean and standard deviation equal to 0.5. The rolling regression analysis is applied to the above variable y and the variable x. In the following lines we present the plots with the rolling coefficients and the rolling scale. From the graphs of the rolling estimates it is clear that $\hat{\alpha}$ is near the value zero for almost half of the windows while in the middle of the sample there takes place a noticeable increase and later the values are stabilized in a specific area of values until the end of the sample. Also the $\hat{\beta}$ values are near unity for most windows and decrease noticeably between the 100th and the 130th observations and increase again later and balance near unity for the rest observations. So the $\hat{\alpha}$ and $\hat{\beta}$ values are quite variable and indicate that the quantities measured by the $\hat{\alpha}$ and $\hat{\beta}$ coefficients are not constant over time. The residuals scale estimates $\hat{\sigma}$ are stabilized in a specific area of values from the beginning of the sample until the 70th observation where there takes place a rapid increase until the 90th observation. Then the residuals scale estimates $\hat{\sigma}$ decrease abruptly until the 100th observation and then, after some increases, are about to be stabilized in a specific area of values until the end of the sample. So we can conclude that the residuals scale estimates $\hat{\sigma}$ are quite variable and this fact indicates, as well as the $\hat{\alpha}$ and $\hat{\beta}$ coefficients do, that there is a structural change in the under testing variables.

https://doi.org/10.26219/heal.aueb.6067

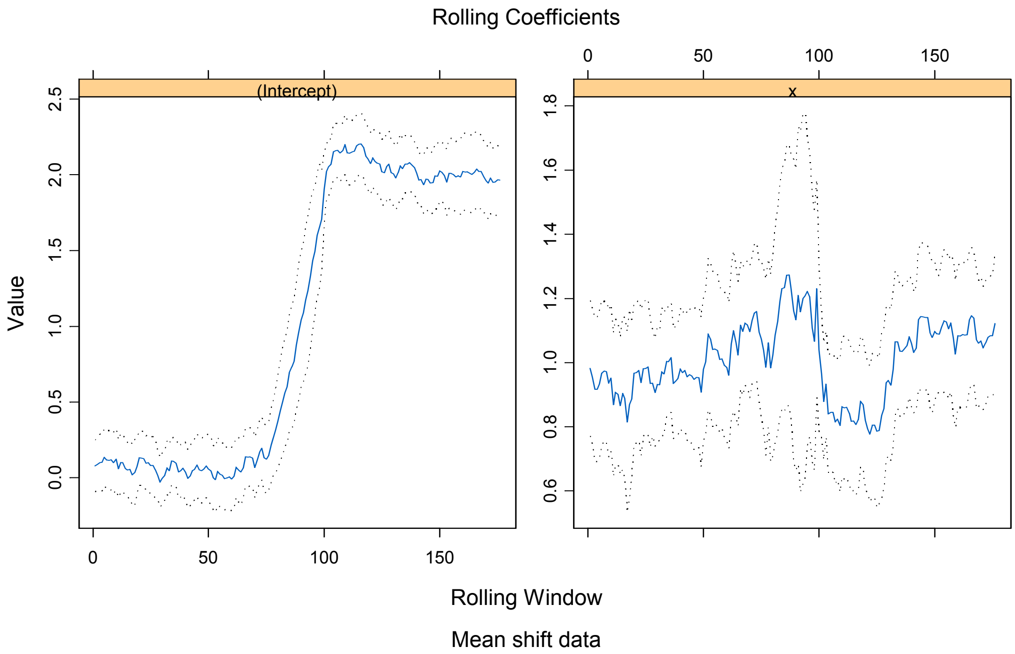


Figure 3.14 Rolling Regression coefficient estimates for the model with structural change in the mean

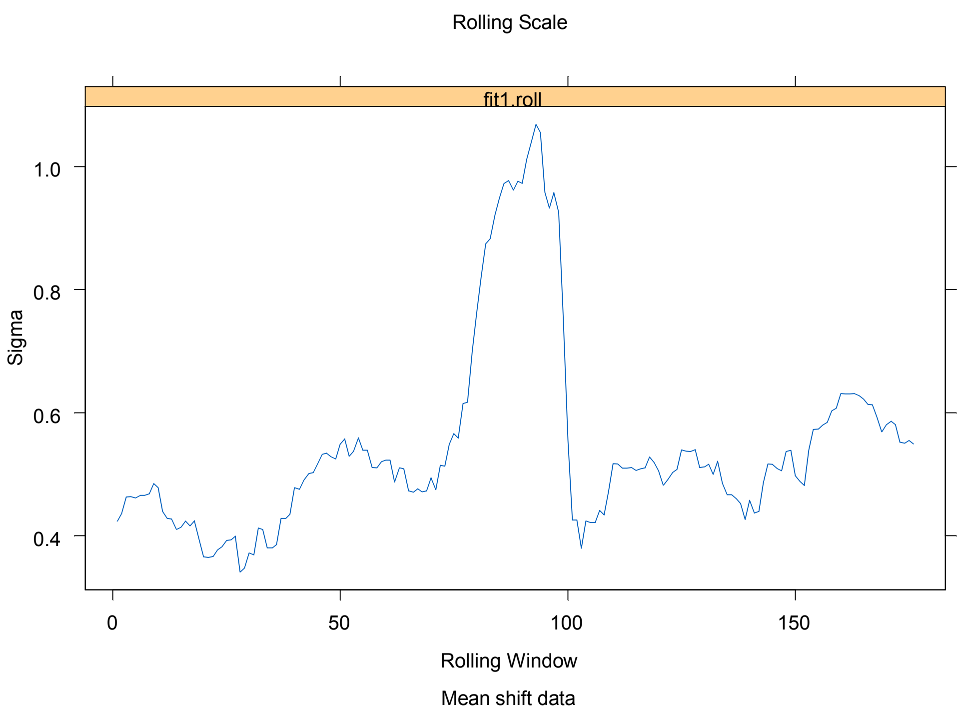


Figure 3.15 Rolling Regression estimates for the residual standard error of the model with structural change in the mean

 https://doi.org/10.26219/heal.aueb.6067

3.4.3 Structural change in slope

In this case we are going to apply the rolling regression analysis in data in which there exists a structural change in the slope. The data consist of the variable y and the variable x. The variable y is defined by

$$y_t = x_t + \varepsilon_t + 2 \cdot I(t > 100) \cdot x_t \quad (3.11)$$

where x is the variable which is constituted of 200 observations from the normal distribution with zero mean and variance equal to one and ε_t are the residuals which come from the normal distribution with zero mean and standard deviation equal to 0.5. The rolling regression analysis is applied to the above variable y and the variable x. In the following lines we present the plots with the rolling coefficients and the rolling scale. From the graphs of the rolling estimates it is clear that $\hat{a}$ is not statistically significantly different from the value zero while the $\hat{\beta}$ values are near unity for about the first 70 observations and increase noticeably between the 70^{th} and the 100^{th} observations and balance later near the value 2.9 for the rest observations. So the $\hat{\beta}$ values are quite variable and this fact indicates that the quantity measured by the $\hat{\beta}$ coefficient is not constant over time. The residuals scale estimates $\hat{\sigma}$ are stabilized in a specific area of values from the beginning of the sample until the 70^{th} observation where there takes place a rapid increase until the 90^{th} observation. Then the residuals scale estimates $\hat{\sigma}$ decrease abruptly until the 100^{th} observation and from the 100^{th} observation and onwards are stabilized in a specific area of values until the end of the sample. So we can conclude that the residuals scale estimates $\hat{\sigma}$ are quite variable and this fact indicates that there is a structural change in the under testing variables.

https://doi.org/10.26219/heal.aueb.6067

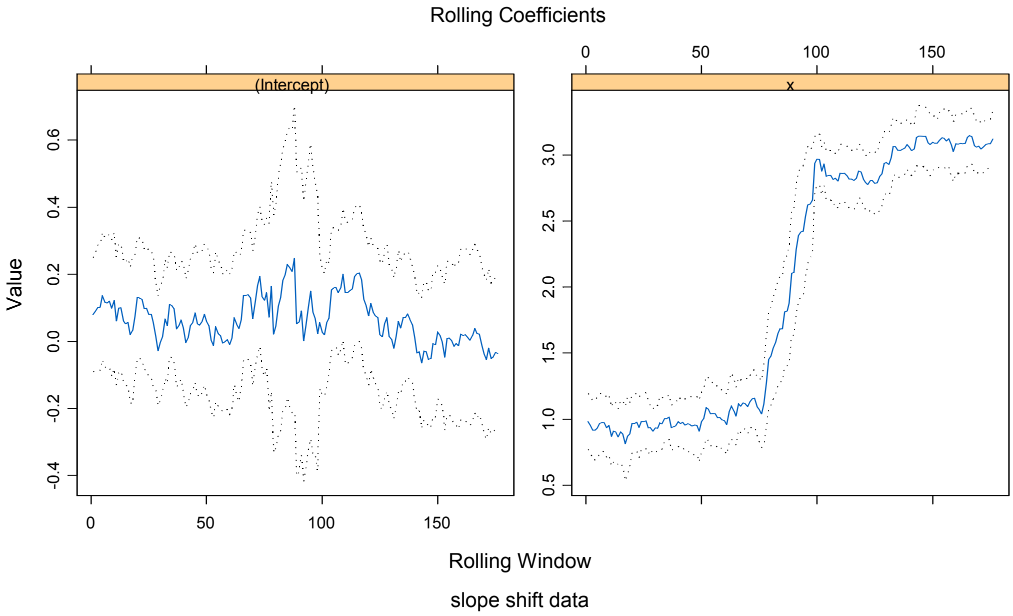


Figure 3.16 Rolling Regression coefficient estimates for the model with structural change in the slope

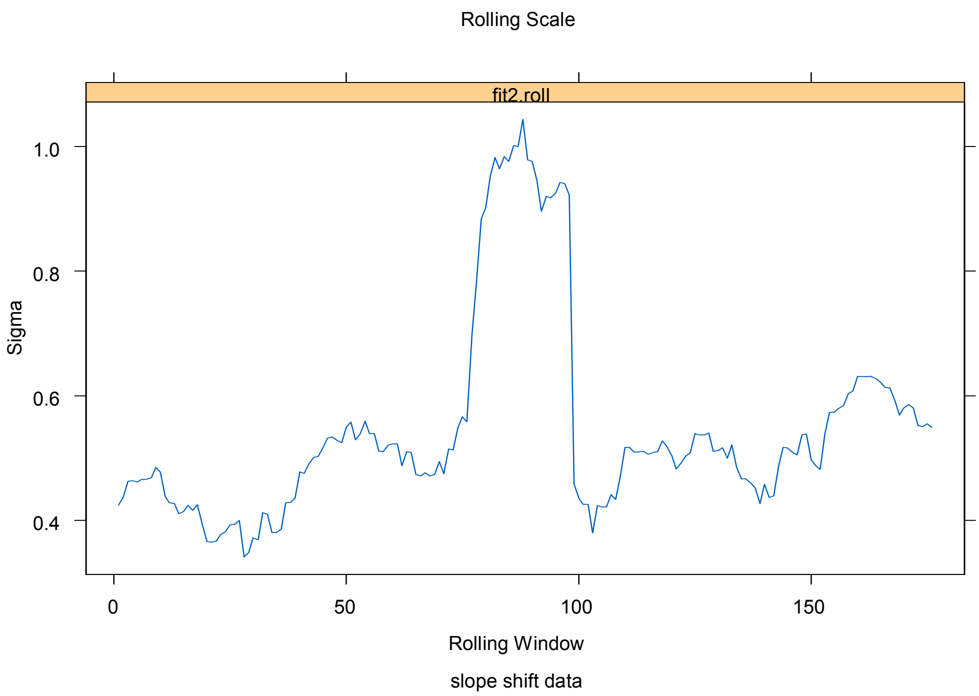


Figure 3.17 Rolling Regression estimates for the residual standard error of the model with structural change in the slope

https://doi.org/10.26219/heal.aueb.6067

3.4.4 Structural change in variance

In this case we are going to apply the rolling regression analysis in data in which there exists a structural change in the variance of the error term. The data consist of the variable y and the variable x. The variable y is defined by

$$y_t = x_t + \varepsilon_t^{\,2} \cdot I(t < 100) + \sqrt{0.5} \cdot \varepsilon_t \cdot I(t > 100) \quad (3.12)$$

where x is the variable which is constituted of 200 observations from the normal distribution with zero mean and variance equal to one and ε_t are the residuals which come from the normal distribution with zero mean and standard deviation equal to 0.5. The rolling regression analysis is applied to the above variable y and the variable x. In the following lines we present the plots with the rolling coefficients and the rolling scale. From the graphs of the rolling estimates it is clear that $\hat{a}$ is significantly different from zero for the first 110 observations while later is not significantly different from zero and the $\hat{\beta}$ values are stabilized in a specific are of values in all the sample except the interval between the 100th and the 130th observations where there is observed a decrease in the values of $\hat{\beta}$. So the $\hat{a}$ and $\hat{\beta}$ values are quite variable and indicate that the quantities measured by the $\hat{a}$ and $\hat{\beta}$ coefficients are not constant over time. The residuals scale estimates $\hat{\sigma}$ are decreasing in the first 30 observations but then takes place an increase until the 85th observation. In the area of values around the 85th observation it is observed the highest value of the residuals scale estimates $\hat{\sigma}$. After the 85th observation there takes place an abrupt decrease in the values of the residuals scale estimates $\hat{\sigma}$ until the 100th observation. From the 100th observation and onwards the values of the residuals scale estimates $\hat{\sigma}$ increase until the 130th observation where there takes place a small decrease but later there takes place an increase again until almost the end of the sample. So we can conclude that the residuals scale estimates $\hat{\sigma}$ are quite variable and this fact indicates that there is a structural change in the under testing variables.

https://doi.org/10.26219/heal.aueb.6067

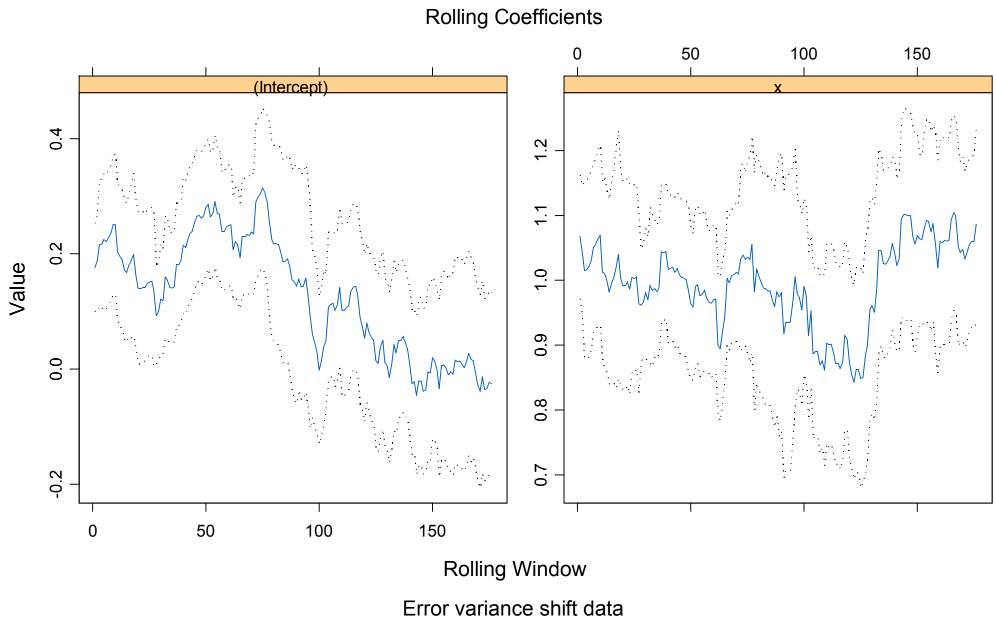


Figure 3.18 Rolling Regression coefficient estimates for the model with structural change in the error variance

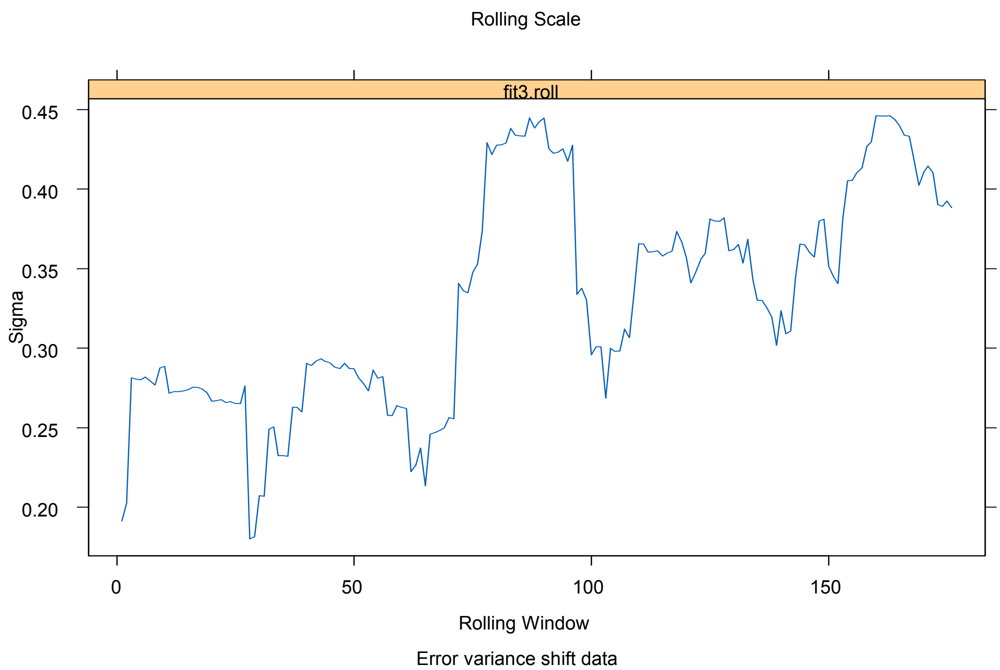


Figure 3.19 Rolling Regression estimates for the residual standard error of the model with structural change in the error variance

https://doi.org/10.26219/heal.aueb.6067

3.4.5 Structural change in the mean and the variance

In this case we are going to apply the rolling regression analysis in data in which there exists a structural change in the mean of the data and the variance of the error term. The data consist of the variable y and the variable x. The variable y is computed by the quantity

$$y_t = x_t + 2 \cdot I(t > 100) + \varepsilon_t^{\,2} \cdot I(t < 100) + \sqrt{0.5} \cdot \varepsilon_t \cdot I(t > 100) \quad (3.13)$$

where x is the variable which is constituted of 200 observations from the normal distribution with zero mean and variance equal to one and ε_t are the residuals which come from the normal distribution with zero mean and standard deviation equal to 0.5. The rolling regression analysis is applied to the above variable y and the variable x. In the following lines we present the plots with the rolling coefficients and the rolling scale. From the graphs of the rolling estimates it is clear that $\hat{a}$ is significantly different from the value zero. Also for about the first 70 observations $\hat{a}$ is stabilized in a specific area of values and later there is observed a noticeable increase in the values of $\hat{a}$ until the middle of the sample. For the rest observations $\hat{a}$ is stabilized in the value 2. Furthermore, the $\hat{\beta}$ values are comparably variable. There takes place a small but steadily increase in the values of $\hat{\beta}$ until the 100^{th} observation where the values of $\hat{\beta}$ decrease abruptly until the 130^{th} observation. From this observation and onwards a steadily increase in the values is observed for about 10 observations. Also from the observations around the area of the observation 135 and onwards the values of $\hat{\beta}$ are constant around a specific area of values. So the $\hat{a}$ and $\hat{\beta}$ values are quite variable and this fact indicates that the quantities measured by the $\hat{a}$ and $\hat{\beta}$ coefficients are not constant over time. The residuals scale estimates $\hat{\sigma}$ are stabilized in a specific value from the beginning of the sample until the 70^{th} observation where there takes place a rapid increase until the 90^{th} observation. Then the residuals scale estimates $\hat{\sigma}$ decrease until the 100^{th} observation and later are stabilized in a specific area of values until the end of

https://doi.org/10.26219/heal.aueb.6067

the sample. So we can conclude that the residuals scale estimates $\hat{\sigma}$ are quite variable and this fact also indicates that there is a structural change in the under testing variables.

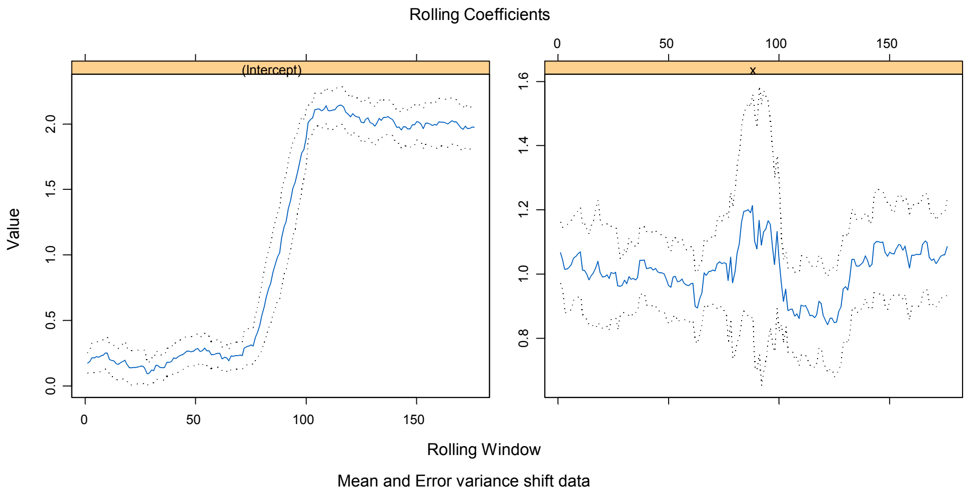


Figure 3.20 Rolling Regression coefficient estimates for the model with structural change in the mean of the data and the error variance

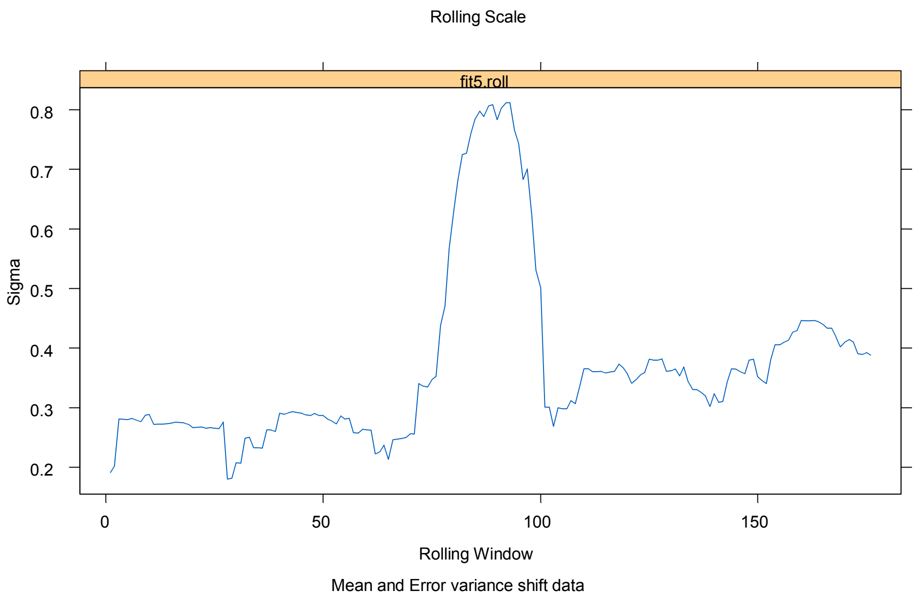


Figure 3.21 Rolling Regression estimates for the residual standard error of the model with structural change in the mean of the data and the error variance

https://doi.org/10.26219/heal.aueb.6067

3.5 Recursive Regression Analysis

In this subchapter of the third chapter we deal with the application of the recursive regression analysis in the different cases of the existence of structural breaks that we have described in the beginning of the chapter.

3.5.1 No structural change data

In this case we are going to apply the recursive analysis in data in which there is no structural change. The data consist of the variable y and the variable x. The variable y is defined by

$$y_t = x_t + \varepsilon_t \quad (3.14)$$

where x is the variable which is constituted of 200 observations from the normal distribution with zero mean and variance equal to one and ε_t are the residuals which come from the normal distribution with zero mean and standard deviation equal to 0.5. The recursive regression analysis is applied to the above variable y and the variable x. The results of this application are shown in the following table and plots.

	Coefficients		**Recursive Residuals Summary**
	Intercept	β	
Mean	0.0858	0.9691	-0.0491
Standard Deviation	0.0444	0.0594	0.4890

Table 3.6 Results of the recursive regression analysis in the case of no structural change

https://doi.org/10.26219/heal.aueb.6067

From the above results of the recursive regression analysis we can see that the average and standard deviation of the $\hat{a}$ values are 0.0858 and 0.0444 respectively and the average and standard deviation of the $\hat{\beta}$ values are 0.9691 and 0.0594 respectively. In the following lines we present the plots with the recursive coefficients, the CUSUM of residuals and the CUSUM of squared residuals.

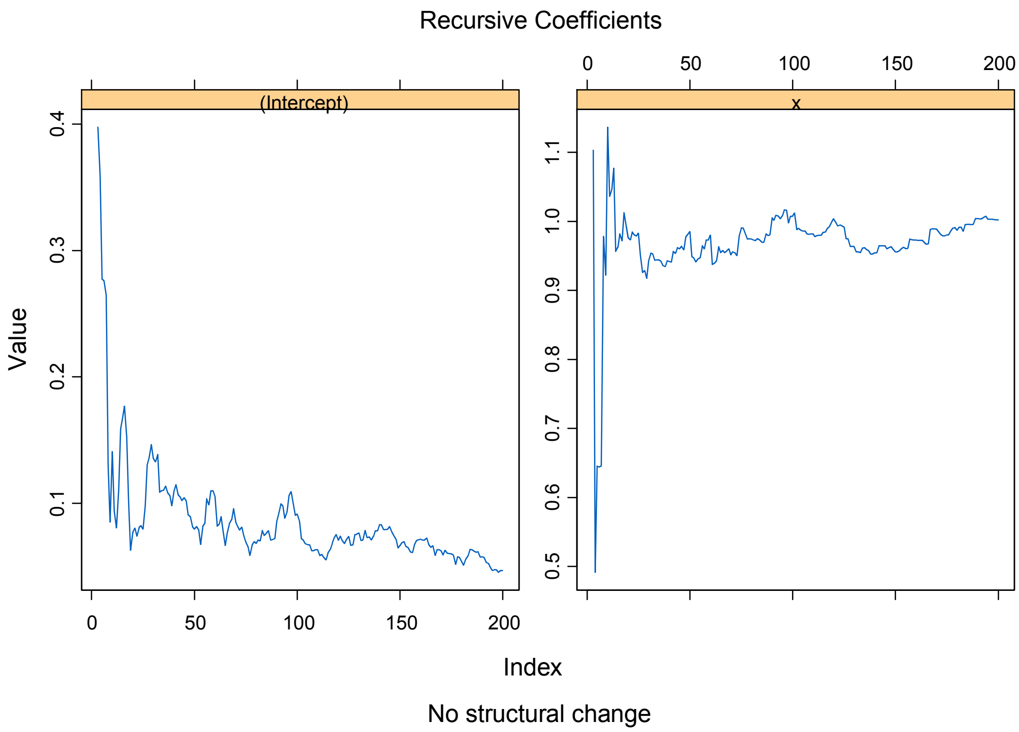


Figure 3.22 RLS coefficient estimates for the model with no structural change

From the above plot of the recursive coefficients it is obvious that the recursive intercept and slope estimates do not seem to vary too much. We can see that the RLS estimates of α and β tend to stabilize to specific values around the middle of the sample and onwards.

https://doi.org/10.26219/heal.aueb.6067

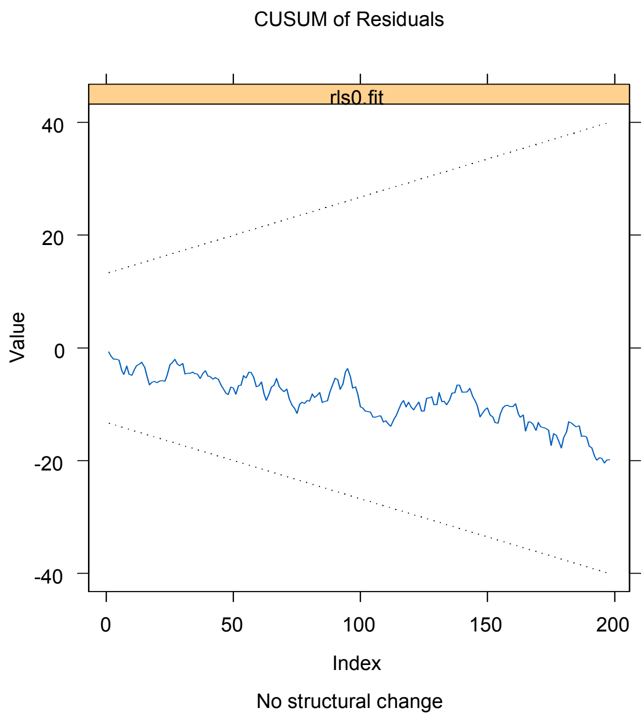


Figure 3.23 CUSUM of residuals for the model with no structural change

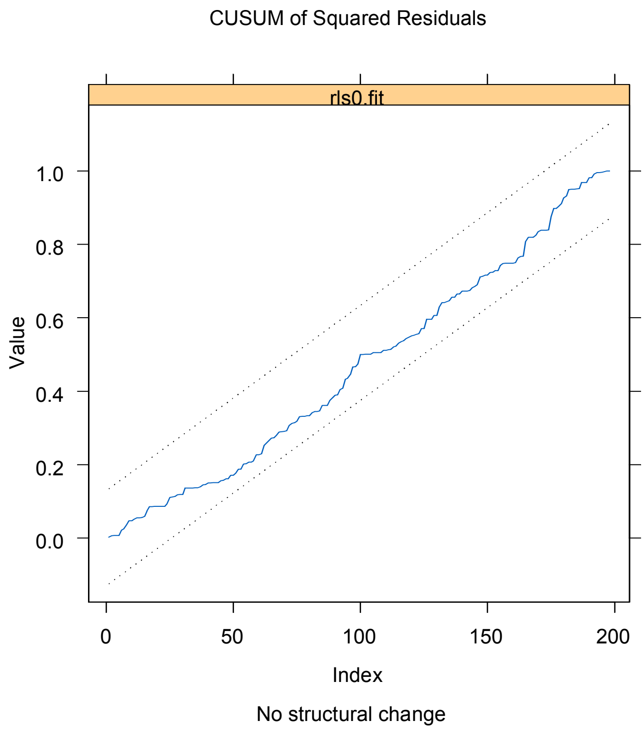


Figure 3.24 CUSUMSQ of residuals for the model with no structural change

https://doi.org/10.26219/heal.aueb.6067

From the above figures it is obvious that the $CUSUM_t$ and the $CUSUMSQ_t$ statistics stay within the 95% confidence bands. Hence, this is some evidence for stability in the coefficients of the under testing linear model.

3.5.2 Structural change in mean

In this case we are going to apply the recursive regression analysis in data in which there exists a structural change in the mean. The data consist of the variable y and the variable x. The variable y is defined by

$$y_t = x_t + \varepsilon_t + 2 \cdot I(t > 100) \quad (3.15)$$

where x is the variable which is constituted of 200 observations from the normal distribution with zero mean and variance equal to one and ε_t are the residuals which come from the normal distribution with zero mean and standard deviation equal to 0.5. The recursive regression analysis is applied to the above variable y and the variable x. The results of this application are shown in the following table and plots.

	Coefficients		Recursive Residuals Summary
	Intercept	β	
Mean	0.4013	1.0368	0.6480
Standard Deviation	0.3534	0.1021	0.8604

Table 3.7 Results of the recursive regression analysis in the case of structural change in the mean

From the above results of the recursive regression analysis we can see that the average and standard deviation of the $\hat{a}$ values are 0.4013 and 0.3534

https://doi.org/10.26219/heal.aueb.6067

respectively and the average and standard deviation of the $\hat{\beta}$ values are 1.0368 and 0.1021 respectively. In the following lines we present the plots with the recursive coefficients, the CUSUM of residuals and the CUSUM of squared residuals.

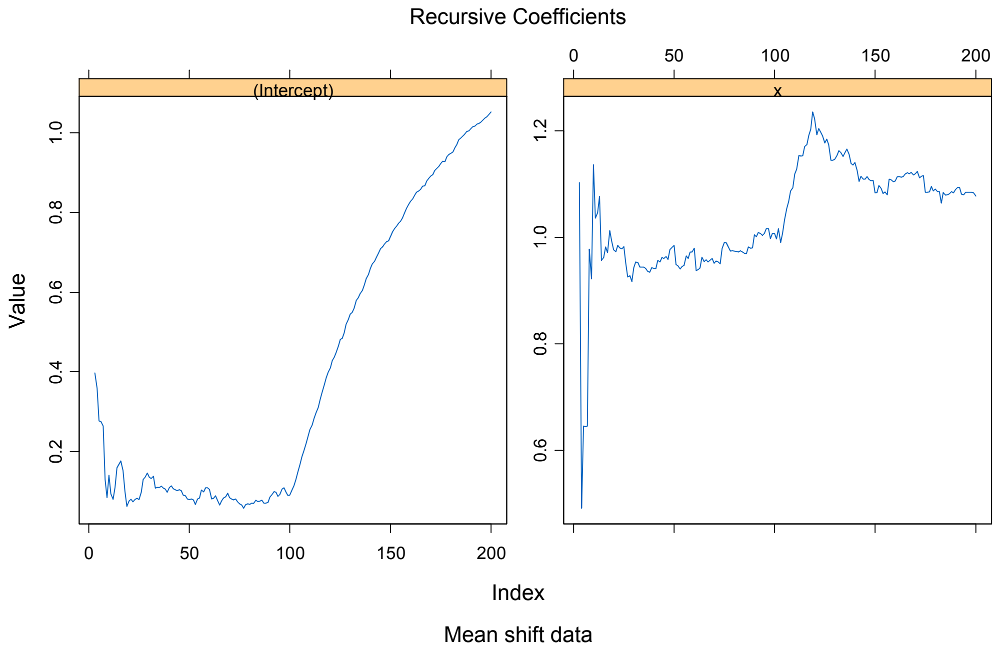


Figure 3.25 RLS coefficient estimates for the model with structural change in the mean

From the above plot of the recursive coefficients it is obvious that the recursive intercept and slope estimates seem to vary too much. The RLS estimate of α is stabilized in a specific area of values from the beginning of the sample until the middle of the sample but then it increases rapidly and reaches its maximum value in the end of the sample. The RLS estimate of β settles down around the middle of the sample but then increases rapidly for some observations and later decreases until it stabilizes in a specific area of values.

https://doi.org/10.26219/heal.aueb.6067

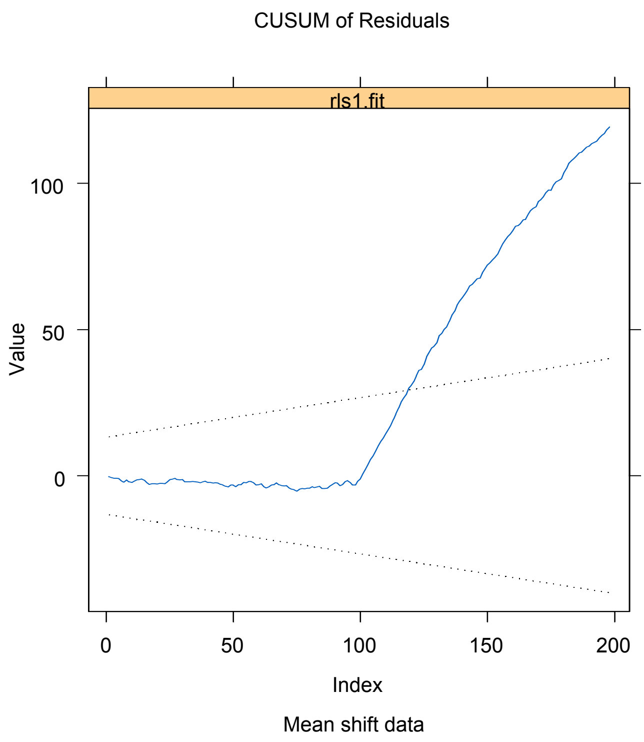


Figure 3.26 CUSUM of residuals for the model with structural change in the mean

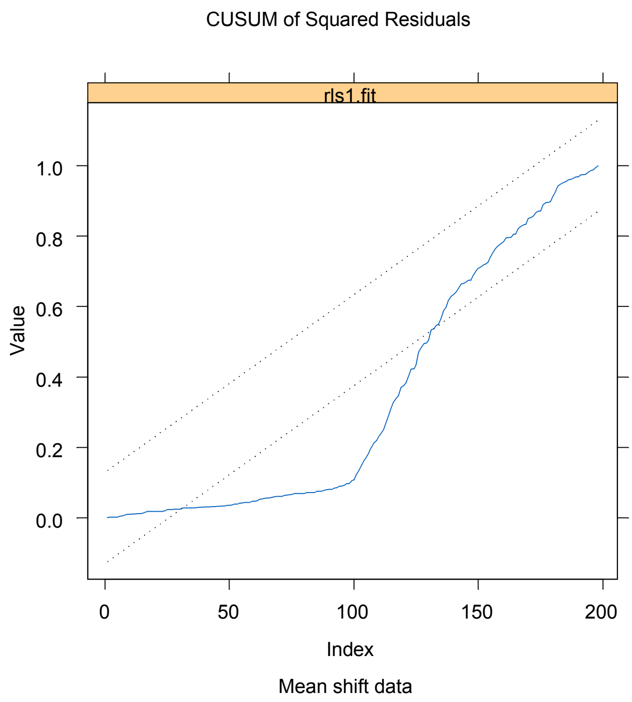


Figure 3.27 CUSUMSQ of residuals for the model with structural change in the mean

https://doi.org/10.26219/heal.aueb.6067

From the above figures it is obvious that the $CUSUM_t$ and the $CUSUMSQ_t$ statistics do not stay within the 95% confidence bands. Hence, this is some evidence for instability in the coefficients of the under testing linear model.

3.5.3 Structural change in slope

In this case we are going to apply the recursive regression analysis in data in which there exists a structural change in the slope. The data consist of the variable y and the variable x. The variable y is defined by

$$y_t = x_t + \varepsilon_t + 2 \cdot I(t > 100) \cdot x_t \quad (3.16)$$

where x is the variable which is constituted of 200 observations from the normal distribution with zero mean and variance equal to one and ε_t are the residuals which come from the normal distribution with zero mean and standard deviation equal to 0.5. The recursive regression analysis is applied to the above variable y and the variable x. The results of this application are shown in the following table and plots.

	Coefficients		**Recursive Residuals Summary**
	Intercept	β	
Mean	0.1451	1.3167	-0.0760
Standard Deviation	0.0629	0.4197	1.0714

Table 3.8 Results of the recursive regression analysis in the case of structural change in the slope

From the above results of the recursive regression analysis we can see that the average and standard deviation of the $\hat{a}$ values are 0.1451 and 0.0629

https://doi.org/10.26219/heal.aueb.6067

respectively and the average and standard deviation of the $\hat{\beta}$ values are 1.3167 and 0.4197 respectively. In the following lines we present the plots with the recursive coefficients, the CUSUM of residuals and the CUSUM of squared residuals.

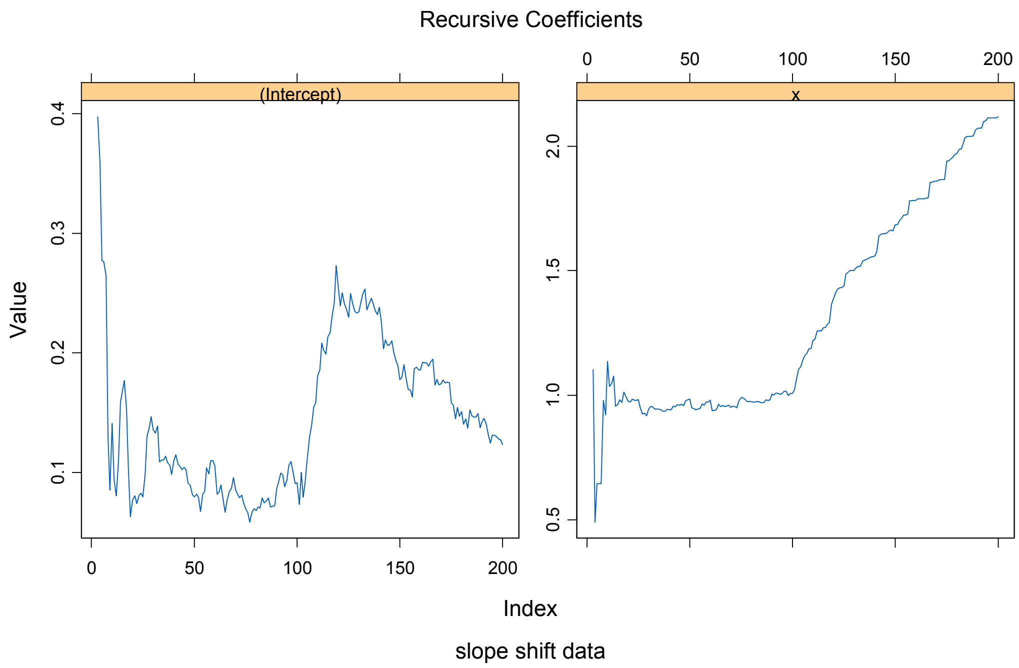


Figure 3.28 RLS coefficient estimates for the model with structural change in the slope

From the above plot of the recursive coefficients it is obvious that the recursive intercept and slope estimates vary too much. The RLS estimate of α settles down in a specific value around the middle of the sample but then it increases rapidly for the next 20 observations. After this increase, the estimate of α decreases slowly until the end of the sample. The RLS estimate of β settles down in a specific value from the beginning of the sample until the middle of the sample but then increases rapidly for the rest observations and reaches its maximum value at the end of the sample.

https://doi.org/10.26219/heal.aueb.6067

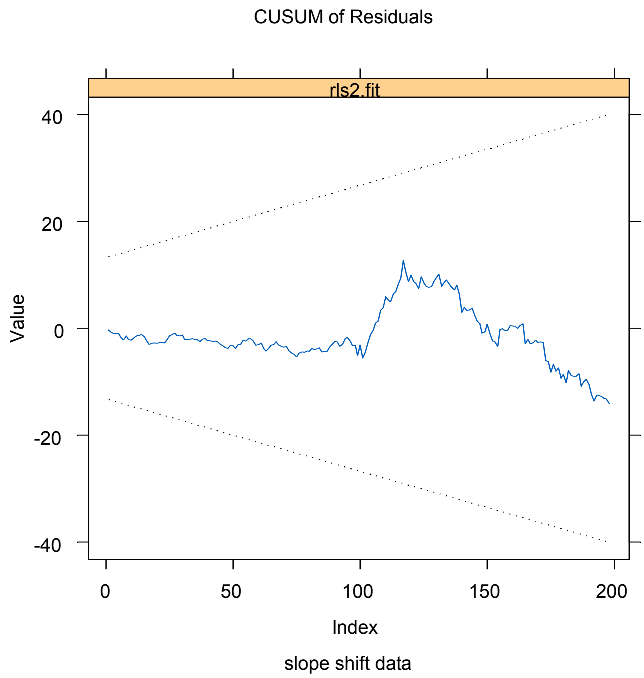


Figure 3.29 CUSUM of residuals for the model with structural change in the slope

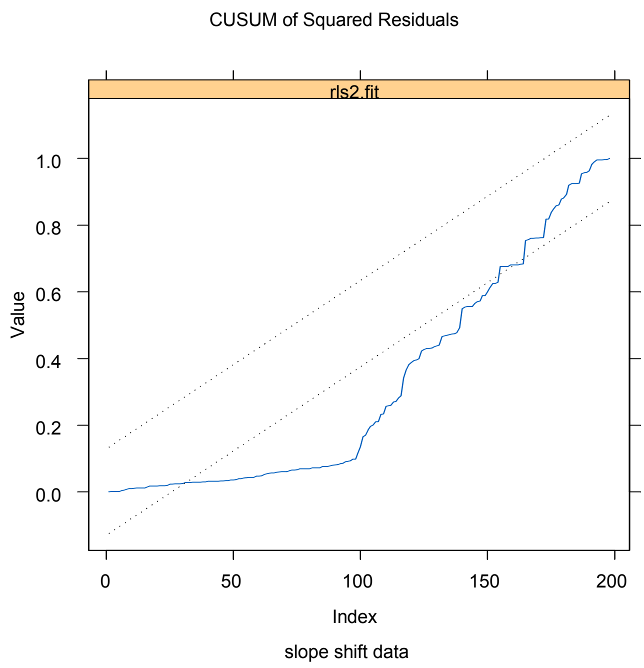


Figure 3.30 CUSUMSQ of residuals for the model with structural change in the slope

https://doi.org/10.26219/heal.aueb.6067

From the above figures it is obvious that the $CUSUM_t$ statistics stay within the 95% confidence bands but the $CUSUMSQ_t$ statistics do not stay within the 95% confidence bands. Hence, this is some evidence for instability in the coefficients of the under testing linear model.

3.5.4 Structural change in variance

In this case we are going to apply the recursive regression analysis in data in which there exists a structural change in the variance of the error term. The data consist of the variable y and the variable x. The variable y is defined by

$$y_t = x_t + \varepsilon_t^{\,2} \cdot I(t < 100) + \sqrt{0.5} \cdot \varepsilon_t \cdot I(t > 100) \quad (3.17)$$

where x is the variable which is constituted of 200 observations from the normal distribution with zero mean and variance equal to one and ε_t are the residuals which come from the normal distribution with zero mean and standard deviation equal to 0.5. The recursive regression analysis is applied to the above variable y and the variable x. The results of this application are shown in the following table and plots.

	Coefficients		**Recursive Residuals Summary**
	Intercept	β	
Mean	0.1783	1.0064	-0.0634
Standard Deviation	0.0291	0.0302	0.3346

Table 3.9 Results of the recursive regression analysis in the case of structural change in the error variance

https://doi.org/10.26219/heal.aueb.6067

From the above results of the recursive regression analysis we can see that the average and standard deviation of the $\hat{a}$ values are 0.1783 and 0.0291 respectively and the average and standard deviation of the $\hat{\beta}$ values are 1.0064 and 0.0302 respectively. In the following lines we present the plots with the recursive coefficients, the CUSUM of residuals and the CUSUM of squared residuals.

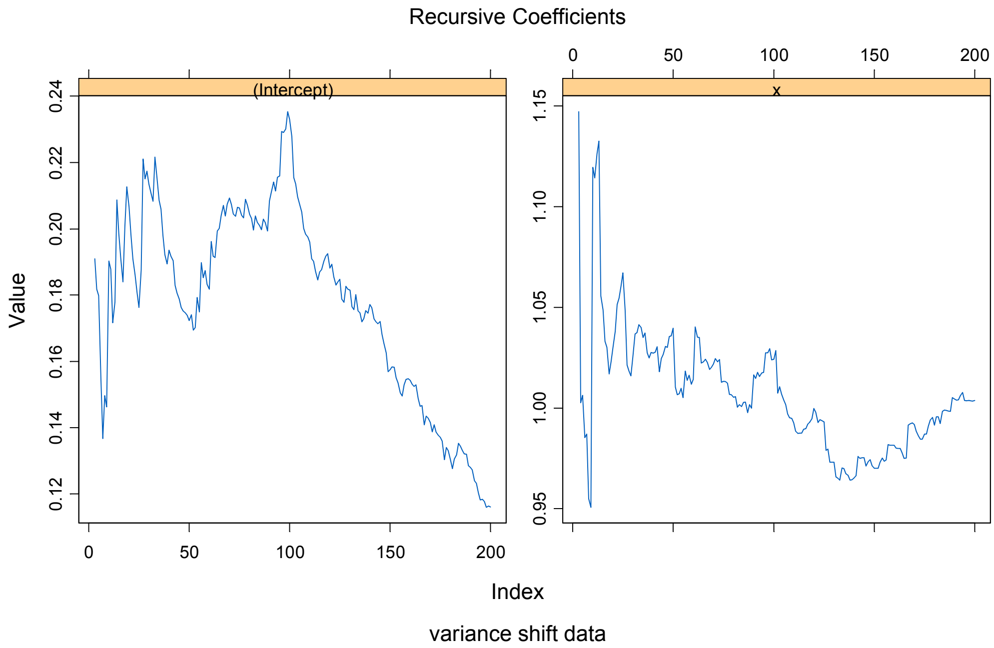


Figure 3.31 RLS coefficient estimates for the model with structural change in the variance

Inspecting the above plot of the recursive coefficients we can say that the recursive intercept and slope estimates vary a little. The RLS estimate of α is increasing in the first 50 observations and later decreases for about 10 observations. Later it increases until the 100th observation while there takes place a huge decrease until the end of the sample. The RLS estimate of β settles down in a specific value in the beginning of the sample until the 110th observation. There takes place a decrease until the 130th observation and later the RLS estimate of β is slightly increasing until the end of the sample.

https://doi.org/10.26219/heal.aueb.6067

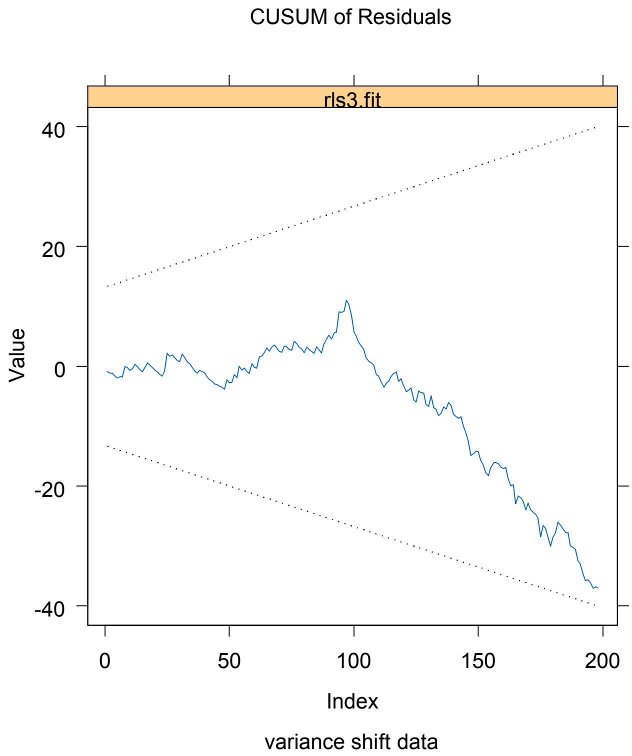


Figure 3.32 CUSUM of residuals for the model with structural change in the variance

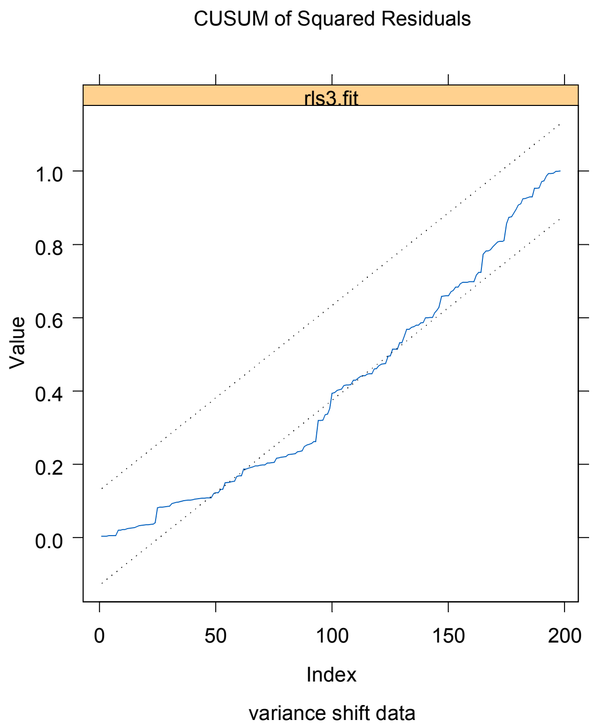


Figure 3.33 CUSUMSQ of residuals for the model with structural change in the variance

https://doi.org/10.26219/heal.aueb.6067

From the above figures it is obvious that the $CUSUM_t$ statistics stay within the 95% confidence bands but some of the $CUSUMSQ_t$ statistics do not stay within the 95% confidence bands. However, this is some evidence for instability in the coefficients of the under testing linear model.

3.5.5 Structural change in the mean and the variance

In this case we are going to apply the recursive regression analysis in data in which there exists a structural change in the mean of the data and the variance of the error term. The data consist of the variable y and the variable x. The variable y is computed by the quantity

$$y_t = x_t + 2 \cdot I(t > 100) + \varepsilon_t^2 \cdot I(t < 100) + \sqrt{0.5} \cdot \varepsilon_t \cdot I(t > 100) \quad (3.18)$$

where x is the variable which is constituted of 200 observations from the normal distribution with zero mean and variance equal to one and ε_t are the residuals which come from the normal distribution with zero mean and standard deviation equal to 0.5. The recursive regression analysis is applied to the above variable y and the variable x. The results of this application are shown in the following table and plots.

	Coefficients		**Recursive Residuals Summary**
	Intercept	β	
Mean	0.4939	1.0741	0.6337
Standard Deviation	0.3457	0.0603	0.7004

Table 3.10 Results of the recursive regression analysis in the case of structural change in the mean and the error variance

https://doi.org/10.26219/heal.aueb.6067

From the above results of the recursive regression analysis we can see that the average and standard deviation of the $\hat{a}$ values are 0.4939 and 0.3457 respectively and the average and standard deviation of the $\hat{\beta}$ values are 1.0741 and 0.0603 respectively. In the following lines we present the plots with the recursive coefficients, the CUSUM of residuals and the CUSUM of squared residuals.

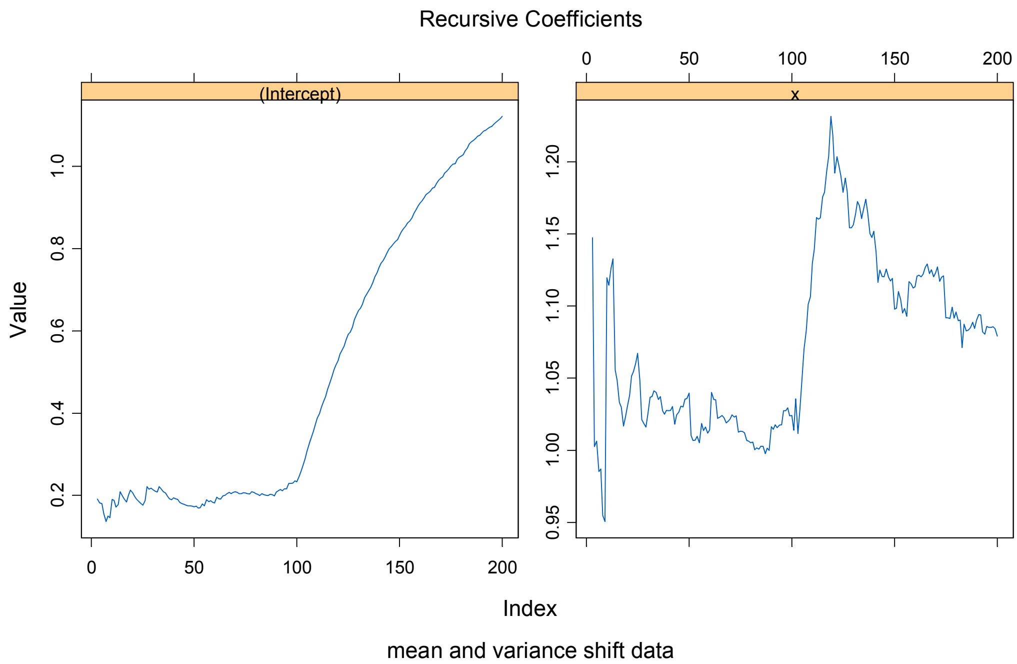


Figure 3.34 RLS coefficient estimates for the model with structural change in the mean of the data and the error variance

From the above plot of the recursive coefficients it is obvious that the recursive intercept and slope estimates seem to vary too much. The RLS estimate of α is stabilized in a specific area of values for the first 100 observations but then it increases rapidly and reaches its maximum value in the end of the sample. The RLS estimate of β settles down in the first 20 observations and remains stabilized until the middle of the sample but then increases rapidly for some observations and later decreases until it stabilizes in a specific value in the last 50 observations of the sample.

https://doi.org/10.26219/heal.aueb.6067

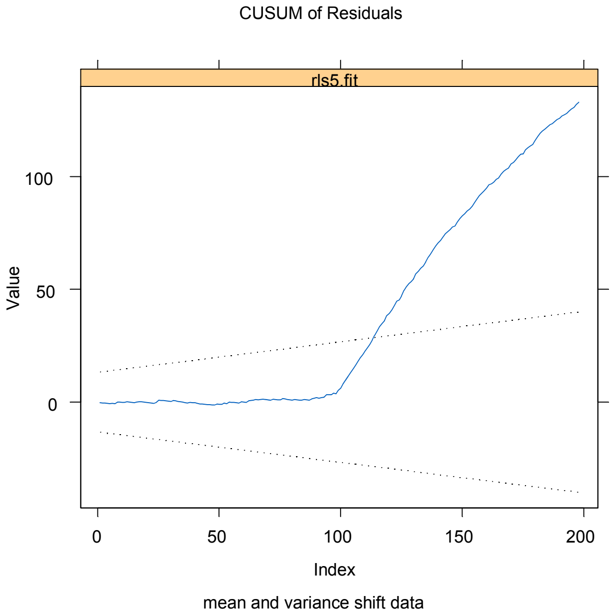


Figure 3.35 CUSUM of residuals for the model with structural change in the mean and the error variance

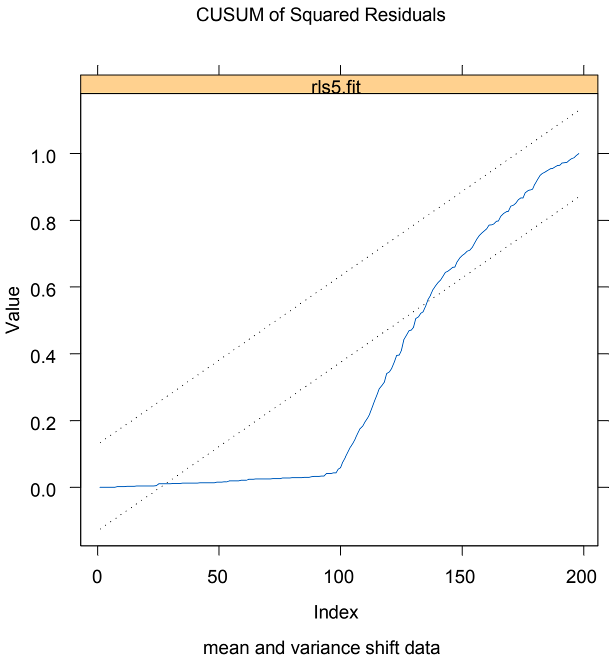


Figure 3.36 CUSUMSQ of residuals for the model with structural change in the mean and the error variance

From the above figures it is obvious that the $CUSUM_t$ and the $CUSUMSQ_t$ statistics do not stay within the 95% confidence bands. As a consequence, this

https://doi.org/10.26219/heal.aueb.6067

is some evidence for instability in the coefficients of the under testing linear model.

https://doi.org/10.26219/heal.aueb.6067

Chapter 4

Application of the Fluctuation tests, the F-tests and the BP method in the simulated data

4.1 Introduction

In this chapter we are going to apply the methods and the tests that we discussed in the 2nd chapter in data that we are going to simulate. Especially we are going to simulate data and later test these data for breaks. Our thought is that we shall simulate data that there exists a break in a known position and we are going to act as if we didn't know the exact position by applying the structural break methods that we described in the 2nd chapter. These data that we are going to simulate could be data from a generating process of a hedge fund's returns or a financial time series but we are going to face these data from a statistical point of view. So, firstly we simulated 70 observations from a normal distribution with mean μ=150 and standard deviation σ=2. Then we simulated another 70 observations with mean μ=159 and standard deviation σ=4. Lastly, we mixed these 140 simulated observations and got the final sample in which we are going to apply the structural break methods. Also we assigned these data to be a time series from January 1960 to August 1971. Before continuing in the application of these methods we shall first note the methods that we are going to use in this data series. As we know from the 2nd chapter and generally as it is known from the scientific research there are three kinds of classical tests in order to test a time series model or a linear regression model for structural breaks. These tests are the fluctuation tests, the F-tests and some tests which are based on least squares estimation. The fluctuation tests that we are going to use in our application are the Recursive

https://doi.org/10.26219/heal.aueb.6067

CUSUM test, the Recursive MOSUM test, the OLS-based CUSUM test and the OLS-based MOSUM test. In these tests we know that if in any point of the corresponding plot of the test the line exceeds the limits of the plot this is an indication that there might be a structural break around this position because the under testing parameter is not constant. If the line lies among the two plot limits (boundaries) there is no structural break in our series and so the under testing parameter is constant. The null hypothesis states that the under testing parameter is constant and the alternative hypothesis states that the parameter is not constant and so there is a structural break. Continuing to the F-tests we already know that the null hypothesis states that the under testing parameter is constant and correspondingly the alternative hypothesis states that the parameter is not constant, which is an indication and a suggestion that there exists a structural break. According to the F-tests the estimation of the break-date is the time at which the F-statistic takes its maximum value. Also we are going to apply the Bai-Perron method in our data. So lastly these are the methods that we are going to apply in our data set in order to test for structural breaks. We have to mention at this point that all the pre-described methods are methods that we use in order to test a data series for structural breaks in the case that we have all the sample, in other words when the full dataset is available. It would be much more interesting if we used a sub sample of our dataset as the history sample and try to predict the break for the rest sub sample, as the second sub sample would be somewhere in the future. This is the method that we have already mentioned in the 2^{nd} chapter as the on-line monitoring method. We are also going to apply this method in our case, in the dataset that we simulated. We shall keep some observations as the history sample and with the rest observations updating this history sample we are going to try to estimate the possible break. This is more realistic and more interesting than using the entire sample as the history sample because someone is interested in identifying structural breaks in the future. So having this in mind we are also going to monitor our dataset with the aim to find some future structural breaks.

https://doi.org/10.26219/heal.aueb.6067

4.2 Application of the Fluctuation tests, the F-tests and the BP method in the simulated data

In the following figures we present the plots of the results of the structural break methods that we applied in the data that we simulated. Especially we can see the plot of the F-statistics, the plots of the Recursive CUSUM and MOSUM tests and the plots of the OLS-based MOSUM and CUSUM tests. These plots are demonstrative of the conclusions that we reach about the position of the structural break. The whole analysis of the results of the above tests and the comments are reported after the following plot.

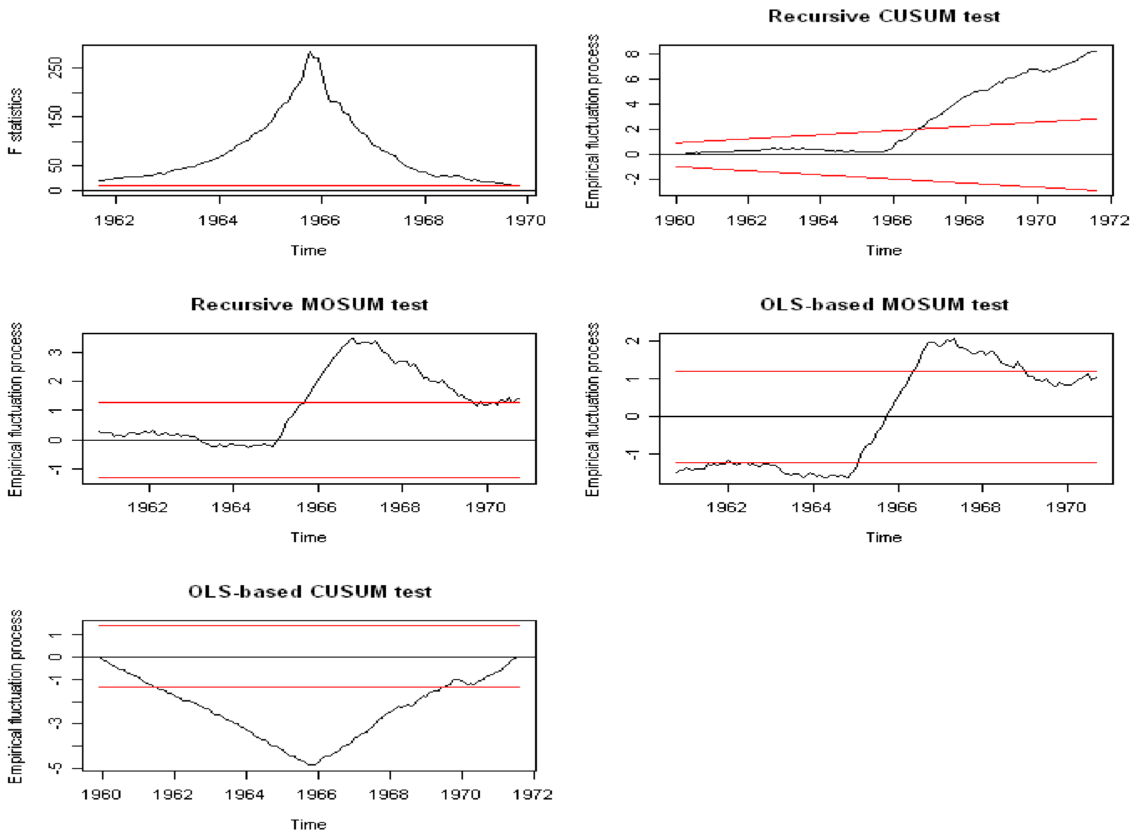


Figure 4.1 Plot of the structural change tests in the case of the simulated data analysis

Inspecting the results of the F-tests, we can conclude that the null hypothesis of parameter constancy (no breaks) is rejected in favour of the alternative hypothesis that states that there is a single break. From the plot of the F-statistics we can conclude that the estimate of the break date is at October 1965, since in this point is observed the highest F-statistic value of all the series of the F-statistics. As we have said in the point that there exists the higher F-statistic value, this point is the structural break. After the conclusions that we reached with F-test we are now going to examine the results of the Bai-Perron (BP) method. Their method finds a structural break at the same position with the F-tests which position corresponds to the 70th observation, that is to say October 1965. From the fluctuation tests we reach in quite the same conclusions we reached with the F-tests and the BP method. In particular we can conclude that in all the fluctuation tests the null hypothesis of parameter constancy (no break) is rejected. So we can conclude that based on the fluctuation tests there are breaks detected in our model. In other words the fluctuation tests are able to detect the possible breaks that exist in our time series, a thing that also the F-test and the BP method did. Looking one by one the plots of the different fluctuation tests we can see that the Recursive CUSUM test and the OLS-based MOSUM test recognize the position of the structural break around the first months of the year 1966. On the contrary the Recursive MOSUM test and the OLS-based CUSUM test identify the position of the structural break at October 1965. So someone can see that both the Recursive MOSUM and the OLS-based CUSUM test identify the same position of the structural break as the F-tests and the BP method did. Also the Recursive CUSUM test and the OLS-based MOSUM test identify the position of the structural break a few months later but around the time area of the end of the year 1965 and the beginning of the year 1966. Approximately we could say that all the fluctuation tests identify the structural break position in agreement with the F-test and the BP method. After the application of the above methods that we used in order to identify the possible structural breaks in our time series we could say that it was known when we simulated the data that the break position is in the 70th observation because after this observation the mean and the standard deviation of the data generating process changed noticeably. So we expected the methods for identifying the possible structural breaks to identify the structural break at the 70th observation or a few observations later. This happened and we can say that the methods we used above recognized correctly and accurately the structural break.

 https://doi.org/10.26219/heal.aueb.6067

4.3 Application of the on-line monitoring method to the simulated data

4.3.1 Moving Estimates Method (Bandwidth h=1)

Here we are concerning with the on-line monitoring method and its application in the time series data that we have simulated. As we have already mentioned this method is of the most useful methods in the structural break models due to the fact that with this method we are able to estimate future breaks in a data series that we already have and in which day by day new data arrive. Now, since we already know the mechanism and the usefulness of the on-line monitoring method, we are going to apply this method to the data that we have simulated before. We already know that our data consist of 70 observations that have been simulated from a normal distribution with mean $\mu=150$ and standard deviation $\sigma=2$ and 70 observations from a normal distribution too with mean $\mu=159$ and standard deviation $\sigma=4$. These 140 observations make our final sample of observations. In the case that we examine here, the on-line monitoring case, we are going to split our data in two clusters. The first cluster will be the cluster that consists of the history data and the second cluster will be the cluster that consists of the future data, the data that arrive day by day and enrich our history sample. In other words we are going to assume that we know the history data and what has happened in the past but on the contrary we don't know the future data and we want to predict them and so predict any possible structural breaks. So by having the history sample we update it with every new observation that arrives and by this we try to predict the possible structural breaks. Continuing in the application of the on-line monitoring method we have to mention the following things. We have already found that in our time series there occurs a structural break at the 70th observation, that is to say October 1965. Due to this fact we are going to cut our data before this breakpoint position and so the data after this cut-point will be the future data. By doing this we are going to be able to check and test in which point (time period) will the breakpoint position be observed. In fact we already know that the structural break position is observed in October 1965 but with the on-line monitoring

https://doi.org/10.26219/heal.aueb.6067

method we test how «fast» and how accurately will that point be detected. It is obvious and understandable that due to the fact that we have cut the data in the history and the incoming sample, the breakpoint position will be observed in later points than it was observed in the case that we used all the sample data as history sample. It is also obvious that the structural break position be observed near the point of October 1965 or a few observations later. So proceeding to the application of the on-line monitoring method we cut our data as we described above in the history and the future data. The structural break was observed in the 70th observation so we decide on our own to cut the history data in the 60th observation, that is to say December 1964. Therefore our history sample consists of the first 60 observations, that is to say the observations from January 1960 to December 1964. The rest of the observations are going to be the future sample, the observations that arrive one by one and enrich our already known historic sample. In the following table and figures we present the results of the on-line monitoring method with the Moving Estimates test. We have to mention at this point that in the procedure of testing for structural breaks and later interpreting the correspondingly results we use a significance level of 5%.

Parameter estimate on history
Intercept
150.0084
Break detected at observation # 78

Table 4.1 Results of the application of the on-line monitoring method to the simulated data via the Moving Estimates test using as bandwidth h=1

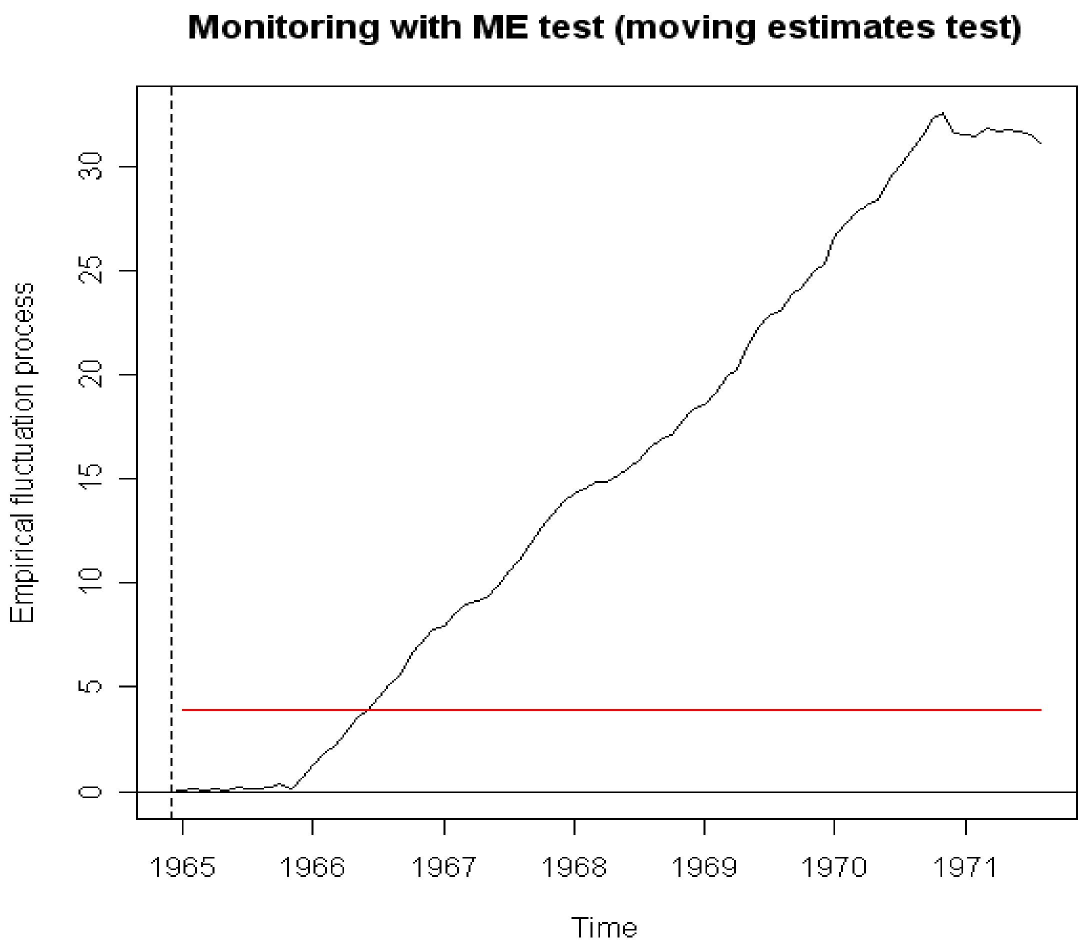


Figure 4.2 Plot of the application of the on-line monitoring method to the simulated data using the Moving Estimates test and bandwidth value h=1

The above figure depicts the monitoring method for the Moving Estimates test, along with a horizontal line which depicts the boundaries, based on formula (2.47), and a dashed vertical line for the beginning of the monitoring period. Also the above figure is based on the empirical fluctuation process, based on a modification of formula (2.46), computed by using the Moving Estimates estimating functions. So looking at the above results of the on-line monitoring method as well as the above figure we can conclude that there is predicted a structural break at the 78th observation. This point corresponds to June 1966. In other words, according to the above figure, we can say that using the sample which corresponds to 1960-1964 data as the history sample we predicted that there is going to be a structural break at the point 78 of the full sample, namely at June 1966. Also we found in the previous subchapter

 https://doi.org/10.26219/heal.aueb.6067

that there exists a structural break at position 70 which corresponds to October 1965, using the complete dataset. So, using the on-line monitoring method we predicted the existence of the structural break 8 months later. We have to mention at this point that in the same way, as the analysis that we did above, we analyze the plots of the different on-line monitoring processes used in the present thesis, depending each time in the different theory that characterizes each test.

4.3.2 Moving Estimates Method (Bandwidth h=0.5)

From the previous results we are able to see that the structural break is detected at the point 78 using as bandwidth h=1 in the statistical process of the on-line monitoring method. Here, we are going to apply the same method as before, the on-line monitoring method, using as bandwidth the value h=0.5. We expect the results to vary a little and that is the reason that we attempt the change of the bandwidth value. We do this in order to check those differences and compare the different results due to the different bandwidth values. The results of the application of the on-line monitoring method with the bandwidth value h=0.5 are illustrated in the following table and plot.

Parameter estimate on history
Intercept
149.8763
Break detected at observation # 76

Table 4.2 Results of the application of the on-line monitoring method to the simulated data via the Moving Estimates test using as bandwidth h=0.5

https://doi.org/10.26219/heal.aueb.6067

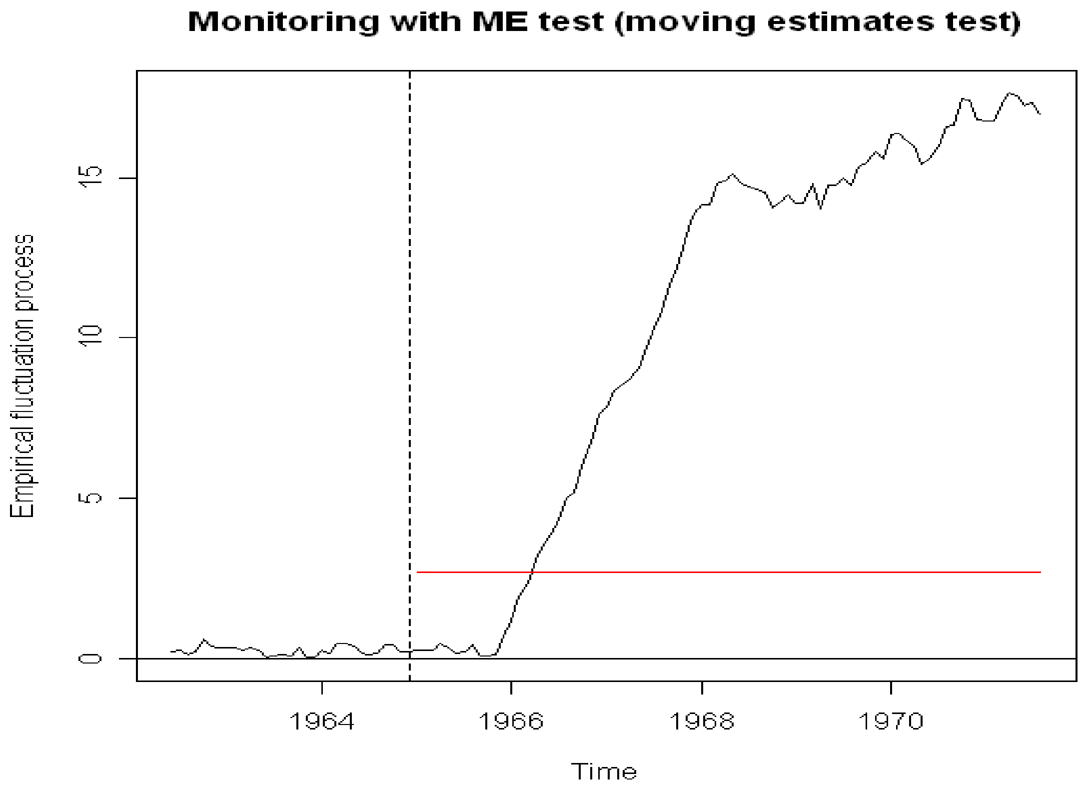


Figure 4.3 Plot of the application of the on-line monitoring method to the simulated data using the Moving Estimates test and bandwidth value h=0.5

From the above results of the on-line monitoring method with bandwidth value h=0.5 we conclude that there occurs a structural break at point 76. It is obvious that we discover the structural break 2 observations earlier in the case that we used the bandwidth value h=0.5 instead of h=1. This structural break position corresponds to April 1966. Using the bandwidth h=1 the structural break that occurred corresponded to June 1966. Consequently, the on-line monitoring process with h=0.5 returns better prediction results than the process with h=1. The process in which we use h=0.5 detects the structural break 6 months later from the already known structural break time of October 1965 that occurs in the complete data set. On the contrary the process in which we used the bandwidth h=1 detects the structural break 8 months later from the position of the structural break that corresponds to October 1965. So, the process in which the bandwidth h=0.5 is used predicted the structural break point earlier than the process in which the bandwidth h=1 is used. Lastly, commenting on the above results of the on-line monitoring method we

https://doi.org/10.26219/heal.aueb.6067

can say that both the use of different bandwidth values emphasize the results of the historic tests because the prediction of the future structural break is somewhat close with the results that the historic tests showed.

4.3.3 Recursive Estimates Method (Bandwidth h=1 and h=0.5)

In this subsection we are dealing with the application of the on-line monitoring procedure to the simulated data using as method the recursive estimates method and as bandwidth the values h=1 and h=0.5. We have already found that in our time series there occurs a structural break at the 70th observation, that is to say October 1965. At this point we call in mind that our history sample consists of the first 60 observations, that is to say the observations from January 1960 to December 1964 while the rest of the observations are going to be the future sample, the observations that arrive one by one and enrich our already known historic sample. We expect the results of the application of the recursive estimates methods to vary with the results of the application of the moving estimates method that we dealt with in the previous subsections. In the following table and plot we present the results of the recursive estimates method. It is worthy to mention that the results were the same in the case of bandwidth value h=1 and h=0.5 so the table and the plot is common between these analyses.

Parameter estimate on history
Intercept
149.8763
Break detected at observation # 75

Table 4.3 Results of the application of the on-line monitoring method to the simulated data via the Recursive Estimates test using as bandwidth h=1 and h=0.5

https://doi.org/10.26219/heal.aueb.6067

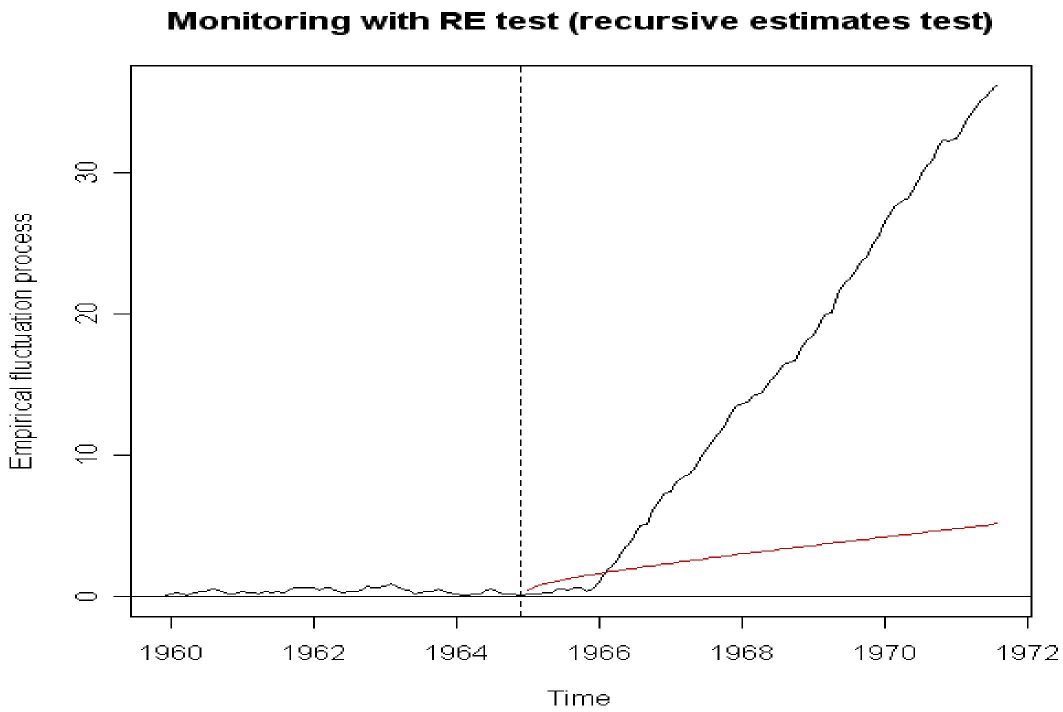


Figure 4.4 Plot of the application of the on-line monitoring method to the simulated data using the Recursive Estimates test and bandwidth values h=1 and h=0.5

Inspecting the above table and plot we conclude that there occurs a structural break in our time series and this break is detected at the 75th observation, that is to say March 1966. Comparing the results that we reached using the Moving Estimates method with the results that we reached using the Recursive Estimates method we see that using as method of the on-line monitoring procedure the recursive estimates method we predicted the existence of the structural break 2 observations earlier than the Moving estimates method using as bandwidth value the value h=1 and 1 observation earlier than the moving estimates method using as bandwidth value the value h=0.5. Although these differences in the forecasting ability of each method are small and almost inconsiderable the fact is that the recursive estimates method has better forecasting ability than the moving estimates method in the above simulated data. We should remember at this point that the recursive estimates method yielded the same results either using as bandwidth value the value h=1 or the value h=0.5.

https://doi.org/10.26219/heal.aueb.6067

https://doi.org/10.26219/heal.aueb.6067

Chapter 5

Application of the structural change tests to the hedge fund data

5.1 Introduction

In this chapter we are going to apply the methods and the tests that we described in the 2[nd] chapter in a real-time problem that has to do with hedge funds. As we have seen before, it is too difficult to evaluate the performance of the hedge funds because of their characteristics and the activities of their managers. Hedge funds strategies are too complex and as a result, the corresponding portfolios are exposed to many economic risk factors, characteristic that makes the possibility of defining a wrong model rise up since there exists no generally accepted model (Vrontos et al., 2008). Also due to some special features and characteristics of hedge funds, their return time series might be characterized by a high degree of non-normality, fat tails and skewness (Amin and Kat, 2003, Kouwenberg and Ziemba, 2007) and also might be characterized by nonlinearities (Fung and Hsieh, 2004). These nonlinearities and generally the process that generates them is a result (consequence) of the market events that have taken place over the previous years. So in order to describe and fully evaluate the hedge fund's performance we need a model that takes in charge all the previous hedge fund features and all the market events that take place over the test period.

There are too many models that have the above characteristics and can be applied to such problems but here we are going to discuss and use the break-point risk factor models. These models can capture the nonlinearities that we mentioned above and also they can detect the changes in risk exposure because of the market events that happen. Specifically the model that we are going to use here is the model that Fung and Hsieh (2004)

 https://doi.org/10.26219/heal.aueb.6067

proposed, a model for fund of hedge funds returns, linear to seven asset-based style (ABS) factors, with dynamic coefficients. Before continuing in the analysis of the model and the interpretation of the results that we got by applying the model to the data, let's give some details about the ABS factors that we mentioned above. First of all there exist two equity factors, the excess returns on the S&P 500 index and the SCMLC factor, which is the spread between small and large cap stock returns. We could say that the SCMLC factor is constructed as the difference of the Wilshire Small Cap 1750 index returns and the Wilshire Large Cap 750 index returns. There also exist two fixed income factors, the change in the 10-year Treasury yields (10Y) and the difference between the change in the Moody's Baa bonds yields and the change in the 10-year T-bonds yields (CredSpr). The last three factors are the so called primitive trend following factors, which are three portfolios of lookback straddles on bonds (PTFSBD), the currencies (PTFSFX) and the commodities (PTFSCOM). These factors have the ability to explain the returns of trend-following funds. These ABS factors have been shown to be valuable explanatory variables for fund of funds and hedge fund returns (Fung and Hsieh (2001, 2002), Fung et al. (2008), and Kosowski et al. (2007)). Having defined and explained the model that we are going to use in our application we are now going to give some brief notes of the data that we use in the specific application. We are going to analyze the Hedge Fund Research Composite Index (HFRCI hereafter) and the Credit Suisse/Tremont Hedge Fund Composite Index (CSTHFCI hereafter). By examining these indexes, someone can have enough information about investments that concern and affect the hedge funds. In other words, these indexes are of prime importance in the procedure of deciding about the distribution of the capital in different bonds that belong to the whole hedge fund strategy. In this point we have to mention the period of time in which the data were collected. Specifically our data consist of the HFRCI monthly excess returns, the CSTHFCI monthly excess returns and also the 7 factors monthly excess returns starting from January 1994 and ending at November 2005. In this period there took place some important events that attainted the global economic history and consequently affected the hedge fund's economic results. These events were the brisk boost in the US interest rates in 1994, the

https://doi.org/10.26219/heal.aueb.6067

Asian crisis in 1997, the Long-Term Capital Management crisis in 1998, the Technology and Internet bubble in 2000, the Japanese crisis in 2001 and lately the terrorism case in USA in 2001.

The tests that we are going to use in order to examine if there are structural changes in Linear Regression are the classical tests. These tests are the fluctuation tests, the F-tests and some tests which are based on least squares estimation. Especially we are going to apply the F-test, the Recursive CUSUM and Recursive MOSUM test, the OLS CUSUM and OLS MOSUM test, the Bai and Perron Method, as well as the Rolling Regression and the Recursive Regression analysis.

5.2 Analysis of Hedge Funds Composite Indices

5.2.1 HFRCI Index Analysis

According to the results of Meligkotsidou and Vrontos (2008) who applied Bayesian methods to the HFRCI index, the model that has the most probabilities to occur is the single break model which includes the S&P, SCMLC, 10Y and CredSpr factors. So we applied all the methods that we are going through in this dissertation in the model that includes the S&P, SCMLC, 10Y and CredSpr factors. In particular we applied the F-test, the Recursive CUSUM and Recursive MOSUM tests, the OLS CUSUM and OLS MOSUM tests, the Bai and Perron method, as well as the Rolling Regression and the Recursive Regression analysis methods in the model that we specified above.

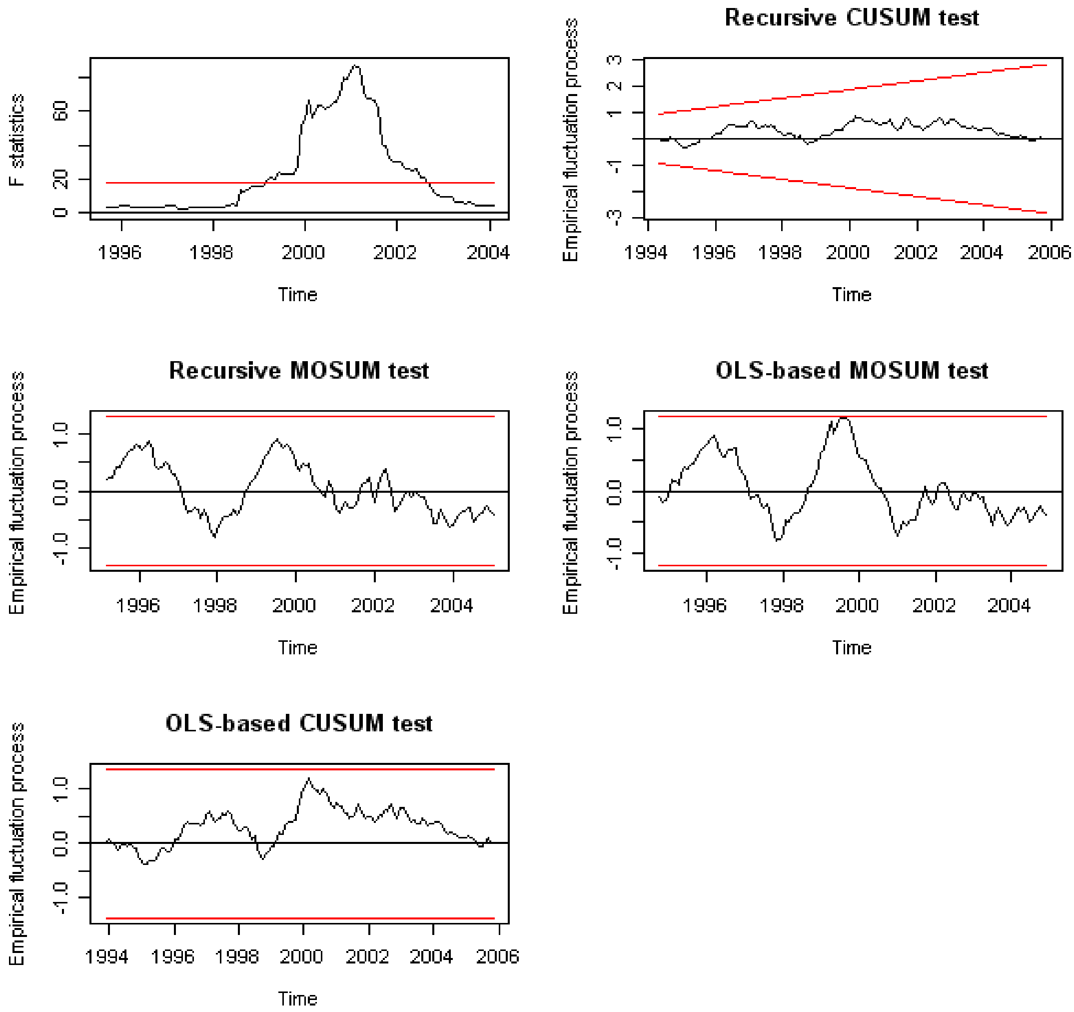


Figure 5.1 Plot of the structural change tests in the case of the HFRCI index analysis

Looking at the results of the F-tests, we conclude that the null hypothesis of parameter constancy (no breaks) is rejected in favour of the alternative hypothesis that states that there is a single break. From the chart of the F-statistics we can conclude that the estimate of the break date is at February 2001, since in this point is observed the highest F-statistic value of all the series of the F-statistics. The same results with the results of the F-test are obtained also with the Bai-Perron method (BP). Their algorithm additively with the F-tests also finds a break at February 2001 as we can see from the table of the corresponding test, based on the BIC criterion.

https://doi.org/10.26219/heal.aueb.6067

Number of Structural Breaks and the Values of the RSS and the BIC criterion					
m = 1		2001(2)			
m = 2	1996(3)	2001(2)			
m = 3	1996(1)	1998(6)	2001(2)		
m = 4	1996(1)	1998(6)	2001(2)	2003(4)	
m = 5	1995(10)	1997(7)	1999(4)	2001(2)	2003(4)
Fit:					
m	0	1	2	3	4
RSS	1.2653e-02	7.6292e-03	6.600e-03	5.9452e-03	5.6357e-03
BIC	-8.9898e+02	-9.4155e+02	-9.325e+02	-9.1766e+02	-8.9553e+02

Table 5.1 Results of the BP method applied to the model for the HFRCI index analysis

After applying the F-test and BP method we are now going to apply and interpret the results of the Recursive Regression Analysis. In the following table we can see the mean and the standard deviation of the intercept, the β's and the recursive residuals.

	Coefficients					**Recursive Residuals Summary**
	α	β_1	β_2	β_3	β_4	
Mean	0.0040	0.3825	0.3143	-0.186	-0.586	0.0001
Standard Deviation	0.0011	0.0457	0.0681	0.8859	3.2068	0.0096

Table 5.2 Results of the recursive regression analysis in the case of the model for the HFRCI index analysis

From the above results of the recursive regression analysis we can see that the average and standard deviation of the $\hat{a}$ values are 0.0040 and 0.0011

 https://doi.org/10.26219/heal.aueb.6067

respectively and the average and standard deviation of the $\hat{\beta}_1$ values are 0.3825 and 0.0457 respectively. Also the average and standard deviation of the $\hat{\beta}_2$ values are 0.3143 and 0.0681, the average and standard deviation of the $\hat{\beta}_3$ values are -0.186 and 0.8859 and lastly the average and standard deviation of the $\hat{\beta}_4$ values are -0.586 and 3.2068 respectively. Furthermore the mean of the recursive residuals is equal to 0.0001 and the standard deviation is equal to 0.0096. In the following lines we present the plots with the recursive coefficients, the CUSUM of residuals and the CUSUM of squared residuals.

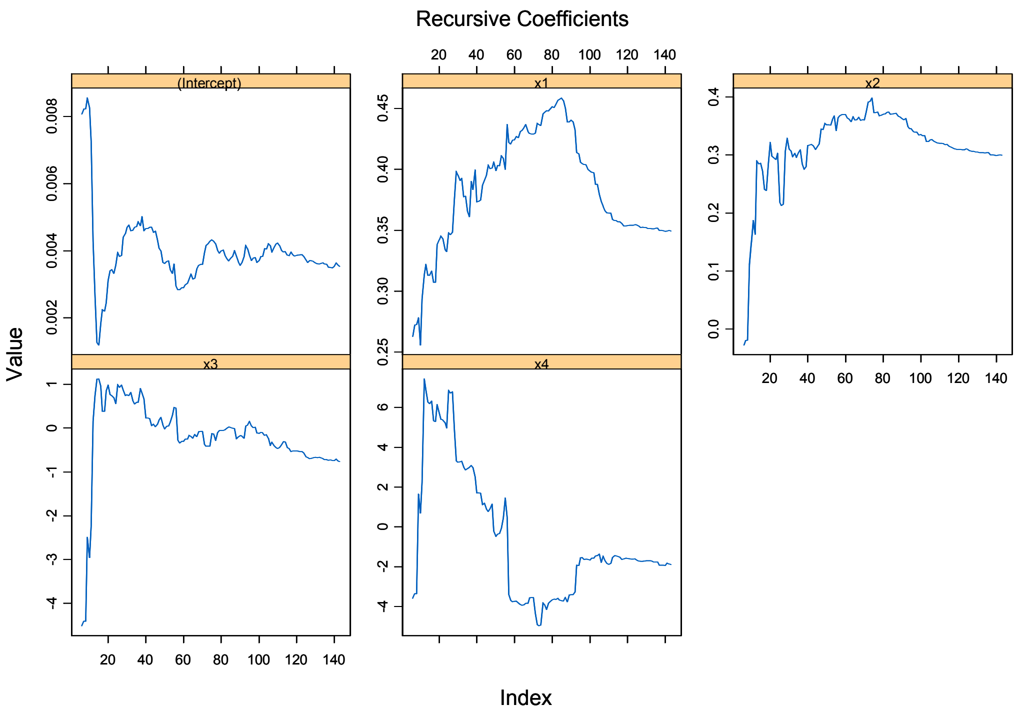


Figure 5.2 RLS coefficient estimates for the model used in the HFRCI index analysis

From the above plot of the recursive coefficients it is obvious that the recursive intercept doesn’t seem to vary but the slopes estimates seem to vary too much. The RLS estimate of α is stabilized in a specific area of values from approximately the very beginning of the sample until the end of the sample. Also the RLS estimates of β_2 and β_3 seem to remain stable in a set of

https://doi.org/10.26219/heal.aueb.6067

values while on the contrary, the RLS estimates of β_1 and β_4 vary too much. The values of the β_1 RLS estimate are increasing rapidly until almost the middle of the sample while later are decreasing until the end of the sample. Also the values of the β_4 RLS estimate are decreasing until the middle of the sample but then are increasing until they are stabilized in a specific area of values until the end of the sample. In the following figures we observe the figure of the CUSUM of Residuals and the CUSUM of Squared residuals.

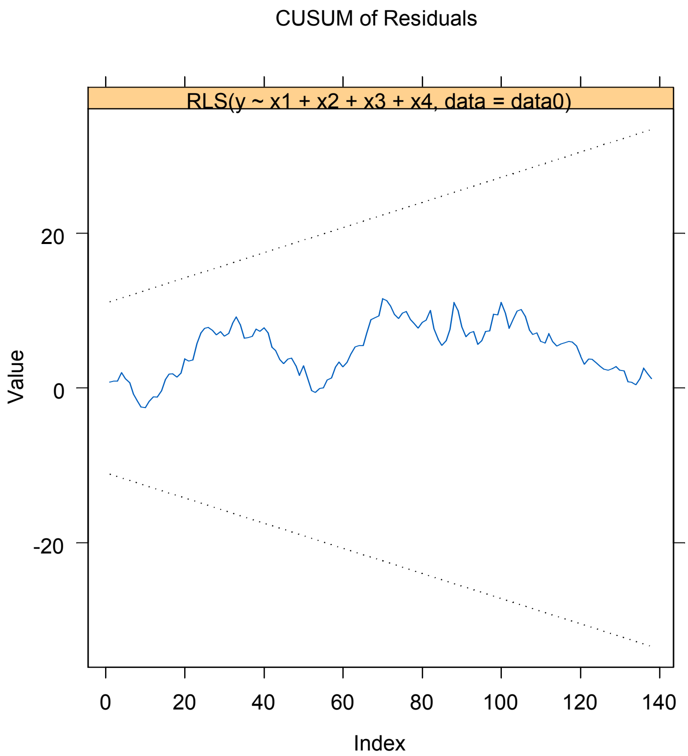


Figure 5.3 CUSUM of residuals for the model used in the HFRCI index analysis

https://doi.org/10.26219/heal.aueb.6067

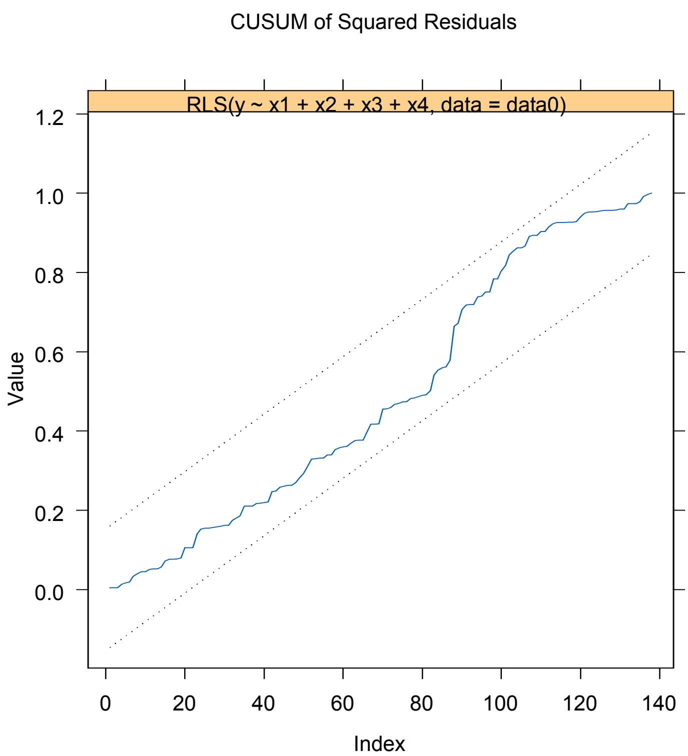


Figure 5.4 CUSUMSQ of residuals for the model used in the HFRCI index analysis

From the above figures it is obvious that the $CUSUM_t$ and the $CUSUMSQ_t$ statistics stay within the 95% confidence bands. Hence, due to the fact that there is a lot of variability in the values of the β's, this is some evidence for instability in the coefficients of the under testing linear model. After applying the Recursive Regression Analysis we are now going to apply the Rolling Regression Analysis in order to examine if there are structural breaks in our dataset using the under testing model. In the following lines we present the results of this method and especially the plots with the rolling coefficients and the rolling scale.

 https://doi.org/10.26219/heal.aueb.6067

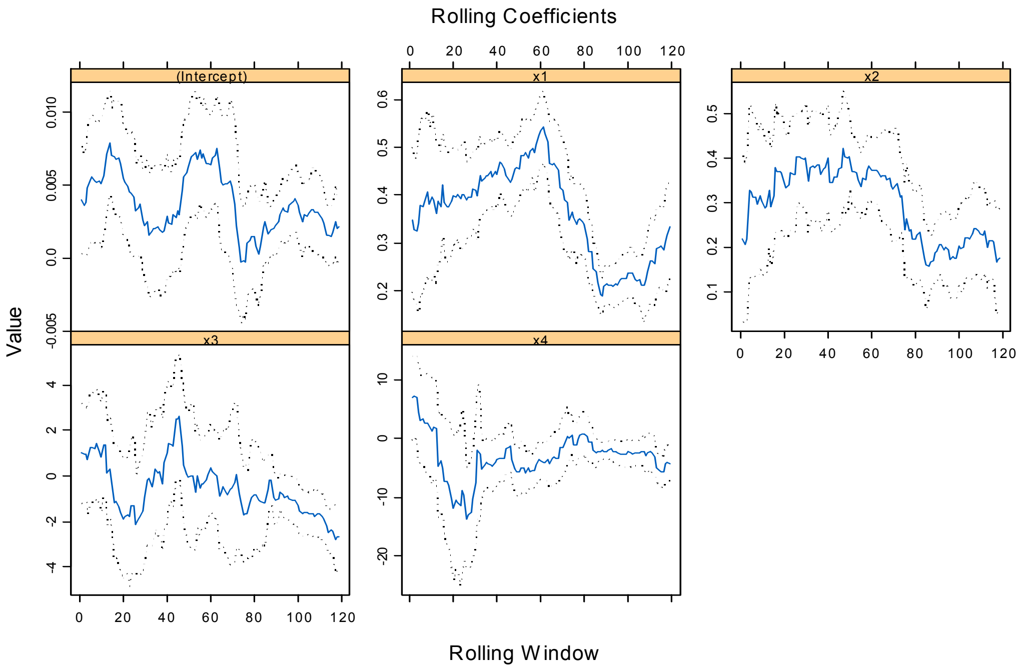


Figure 5.5 Rolling Regression coefficient estimates for the model used in the HFRCI index analysis

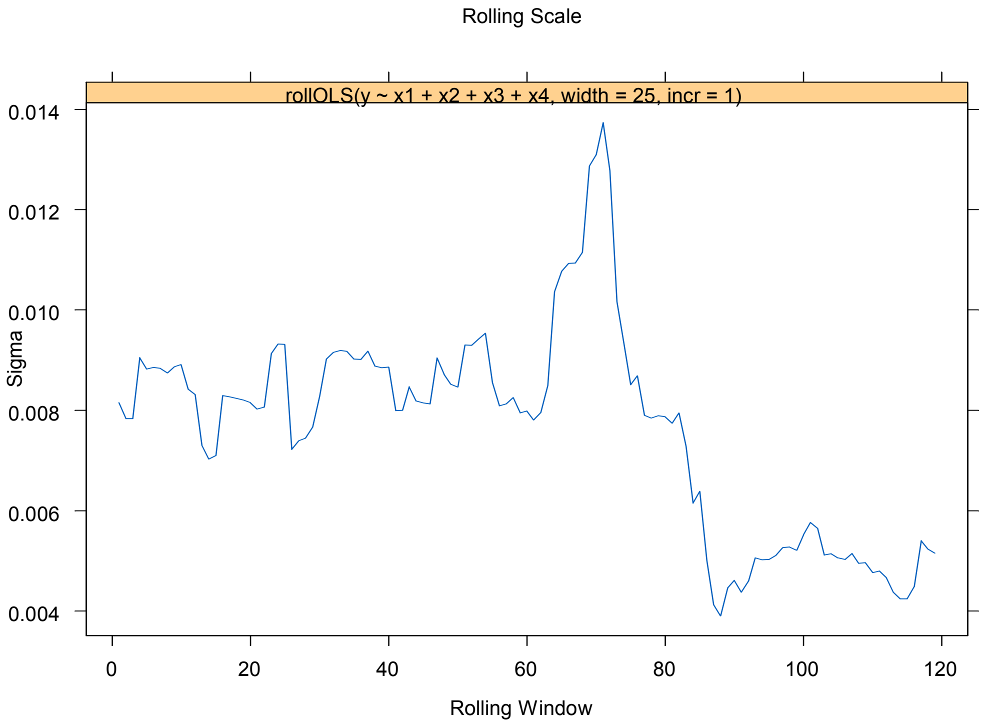


Figure 5.6 Rolling Regression estimates for the residual standard error of the model used in the HFRCI index analysis

https://doi.org/10.26219/heal.aueb.6067

From the graphs of the rolling estimates it is clear that $\hat{a}$ as well as the $\hat{\beta}$'s coefficients vary a lot across the sample. We see that in all the cases there is a lot of variability in the values of each coefficient. There are a lot of peaks in the data series of every coefficient. This is an indication that there exists a structural break in our under testing data. Also the residuals scale estimates $\hat{\sigma}$ are stabilized in a specific area of values from the beginning of the sample until the 60^{th} observation where there takes place a rapid increase until the70^{th} observation. Then the residuals scale estimates $\hat{\sigma}$ decrease abruptly until the 85^{th} observation and from the 90^{th} observation and onwards are stabilized in a specific area of values until the end of the sample. This rambling increase and decrease in the series of the residual scale estimates $\hat{\sigma}$ is also an indication that there exists a structural break in our data. After applying the Rolling Regression Analysis we are now going to apply and interpret the results of the fluctuation tests. In the fluctuation tests we can see from the corresponding chart (figure 5.1) that the null hypothesis of parameter constancy (no break) is not rejected. So we can conclude that based on the fluctuation tests there are no breaks detected in our model. In other words the fluctuation tests cannot detect any breaks in our data, something that F- tests, the BP algorithm, the Recursive and the Rolling Regression Analysis did. After obtaining the results of the above methods we are now ready to continue to the interpretation of them. Looking up the economic history and the events that happened in the time period that our data were collected (Jan 1994-Nov 2005) someone would imagine and would say that the variance of the returns in the pre-break period would be higher than the variance in the post-break period. This assumption is stable due to the fact that in the pre-break period there took place more economic events (the Asian Crisis, the Long Term Capital Management, the Technology and Internet bubble) than in the post break period. Actually, as we see from the results obtained in the following table this assumption is affirmed. The variance σ^2 in the pre-break period is higher than the variance in the post-break period. As a consequence and due to the fact that more events that took place in the pre-break period the coefficients of the risk factors in the pre-break period take higher values than the coefficients in the post-break period.

https://doi.org/10.26219/heal.aueb.6067

	First Segment	Second Segment
α	0.0037801 (0.0009922)*	0.0025435 (0.0007385)*
$\beta_{S\&P}$	0.4587172 (0.0241206)*	0.2541979 (0.0184146)*
β_{SCMLC}	0.3710481 (0.0274316)*	0.1925648 (0.0292953)*
β_{10Y}	0.0108187 (0.5566731)	-1.5838140 (0.3979893)*
$\beta_{CredSpr}$	-3.7167029 (1.1431679)*	-2.4538323 (0.7528907)*
σ^2	0,7758086e-04	0,2933306e-04

* Statistically significant

Table 5.3 Estimations and standard deviations of the most probable model in the case of the HFRCI index

Having commented on the coefficients for the pre and the post break period we are now going to comment on the results of the model specification for the two periods. For the pre-break period, based on the corresponding results we can conclude that all the coefficients are statistically significant (a=5%) except the coefficient β_{10Y}, which reflects the 10Y factor. All the other factors and the intercept are statistically significant. Also all the coefficients take positive values except the coefficient $\beta_{CredSpr}$, which reflects the CredSpr factor. For the post-break period we can conclude that all the coefficients and consequently the factors are statistically significant. Also the intercept, the $\beta_{S\&P}$ and β_{SCMLC} variables take positive values apart from the variables β_{10Y} and $\beta_{CredSpr}$ which take negative values. Also the variables $\beta_{S\&P}$, β_{SCMLC}, β_{10Y} including the intercept take larger values in the pre-break period than the post-break period apart from the variable $\beta_{CredSpr}$ which takes its larger value in the post-break period.

5.2.2 CSTHFCI Index Analysis

As before with the HFRCI index, according to the results of the paper that Meligkotsidou and Vrontos (2008) wrote, applying Bayesian methods to the data of the CSTHFCI index and the ABS factors, the model that has the highest probability to occur is the two break model which includes the S&P,

 https://doi.org/10.26219/heal.aueb.6067

SCMLC, 10Y, CredSpr and PTFSFX factors. So we applied all the methods that we did before with the case of HFRCI index dealing with the model that includes the S&P, SCMLC, 10Y, CredSpr and PTFSFX as factors. Specifically we applied the F-test, the Recursive CUSUM and Recursive MOSUM tests, the OLS CUSUM and OLS MOSUM tests, the Bai and Perron method as well as the Recursive and the Rolling Regression analysis methods in the model that we specified above for the CSTHFCI index.

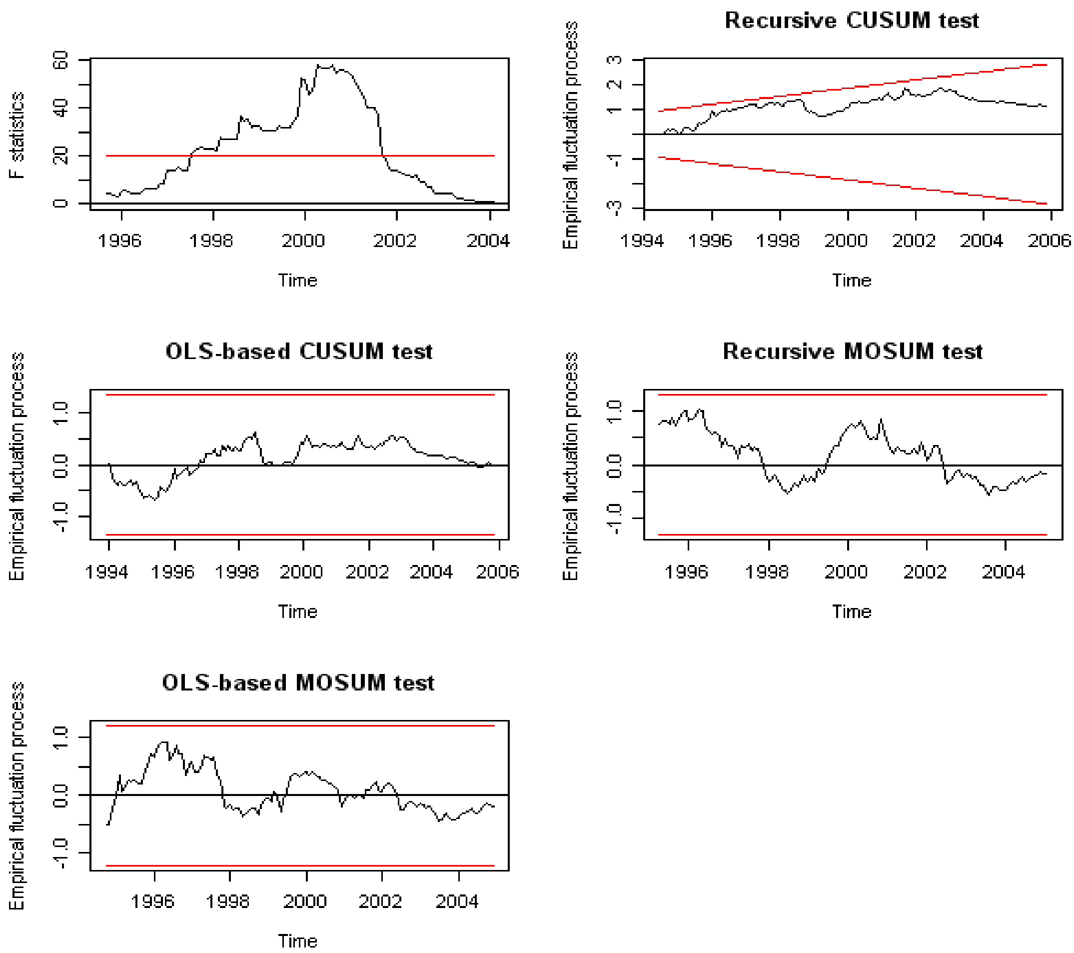


Figure 5.7 Plot of the structural change tests in the case of the CSTHFCI index analysis

Firstly we are going to describe the results of the F-test method. As we can see from the F-test plot there are a lot of values that exceed the limit. The period that these F-tests exceed the limit is around 1998 and 2001. This limit violation is an indication of a break around that period. From the results and the plot of the F-test it is concluded that the estimated break date of the F-test

https://doi.org/10.26219/heal.aueb.6067

method is April 2000, since in this point we observe the highest F-statistic value of all the consequence of the F-statistics. The next method that we applied to our data was the Bai and Perron method (BP). According to this method and based on the BIC criterion, there exists only one structural brake in our data and the estimated breakdate is located at April 2000.

Number of Structural Breaks and the Values of the RSS and the BIC criterion						
m = 1	2000(4)					
m = 2	1998(4)	2000(4)				
m = 3	1995(9)	1998(5)	2000(4)			
m = 4	1995(9)	1997(7)	1999(5)	2001(2)		
m = 5	1995(9)	1997(7)	1999(5)	2001(2)	2003(4)	
Fit:						
m	0	1	2	3	4	5
RSS	0.04309	0.029793	0.02483	0.022454	0.021217	0.0206968
BIC	-718.7974	-736.81852	-728.14205	-707.7806	-681.14004	-649.95271

Table 5.4 Results of the BP method applied to the model for the CSTHFCI index analysis

We observe that both the F-test and the BP method estimate a model with only one structural break and the estimated breakdate is at April 2000. After applying the F-test and BP method we are going now to apply and interpret the results of the Recursive Regression Analysis. In the following table we can see the mean and the standard deviation of the intercept, the β's and the recursive residuals.

 https://doi.org/10.26219/heal.aueb.6067

	Coefficients						**Recursive Residuals Summary**
	α	β_1	β_2	β_3	β_4	β_6	
Mean	0.0019	0.376	0.1311	-1.761	-5.239	0.0148	0.0008
Standard Deviation	0.0025	0.112	0.1516	1.514	7.537	0.0236	0.0178

Table 5.5 Results of the Recursive Regression analysis in the case of the model for the CSTHFCI index analysis

From the above results of the recursive regression analysis we can see that the average and standard deviation of the $\hat{a}$ values are 0.0019 and 0.0025 respectively and the average and standard deviation of the $\hat{\beta}_1$ values are 0.376 and 0.112 respectively. Also the average and standard deviation of the $\hat{\beta}_2$ values are 0.1211 and 0.1516, the average and standard deviation of the $\hat{\beta}_3$ values are -1.761 and 1.514, the average and standard deviation of the $\hat{\beta}_4$ are -5.239 and 7.537 and lastly the average and standard deviation of the $\hat{\beta}_6$ values are 0.0148 and 0.0236 respectively. Furthermore the mean of the recursive residuals is equal to 0.0008 and the standard deviation is equal to 0.0178. In the following lines we present the plots with the recursive coefficients, the CUSUM of residuals and the CUSUM of squared residuals.

https://doi.org/10.26219/heal.aueb.6067

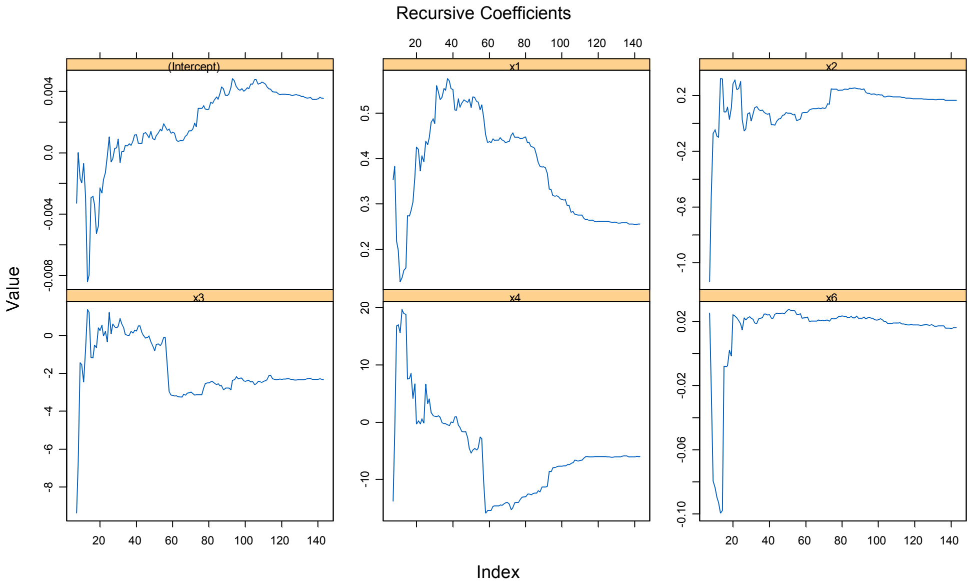


Figure 5.8 RLS coefficient estimates for the model used in the CSTHFCI index analysis

From the above plot of the recursive coefficients it is obvious that the recursive intercept and some of the slopes estimates seem to vary too much. The RLS estimate of α is stabilized in a specific area of values from approximately the very beginning of the sample until the middle of the sample where it increases until almost the end of the sample. Also the RLS estimates of β_2 and β_6 seem to remain stable in a set of values while on the contrary, the RLS estimates of β_1, β_3 and β_4 vary too much. The values of the β_1 RLS estimate are increasing rapidly until almost the middle of the sample while later are decreasing until the end of the sample. Also the values of the β_4 RLS estimate are decreasing until the middle of the sample but then are increasing until they are stabilized in a specific area of values until the end of the sample. Furthermore the values of the β_3 have a peak around the middle of the sample where the values decrease until a specific value and remain stable until the end of the sample. In the following figures we observe the figure of the CUSUM of Residuals and the CUSUM of Squared residuals.

https://doi.org/10.26219/heal.aueb.6067

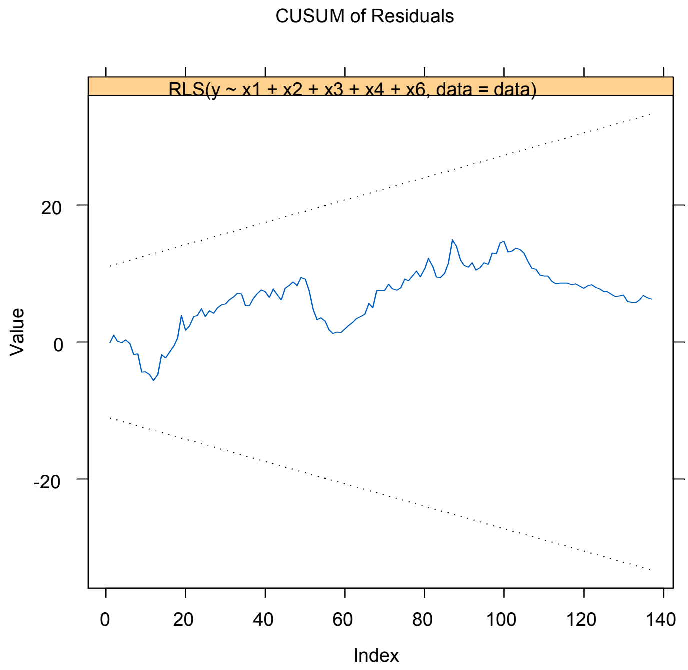


Figure 5.9 CUSUM of residuals for the model used in the CSTHFCI index analysis

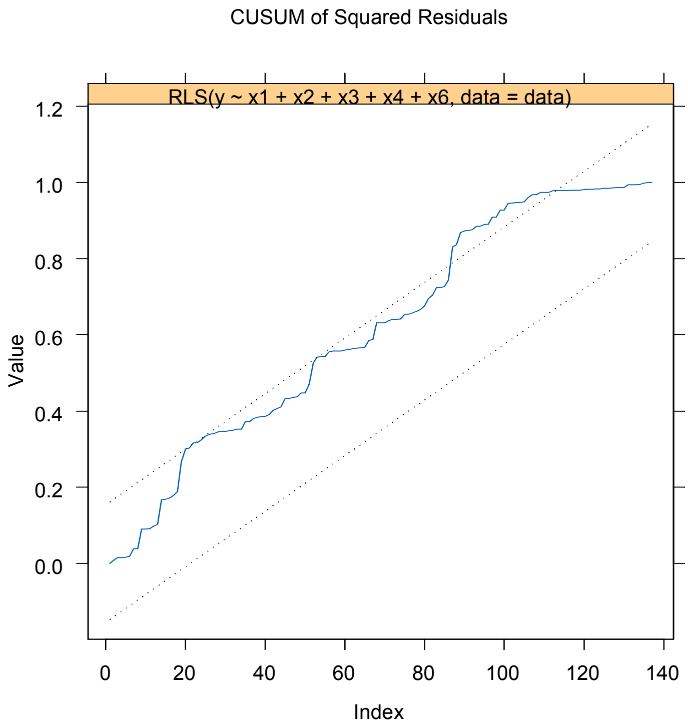


Figure 5.10 CUSUMSQ of residuals for the model used in the CSTHFCI index analysis

https://doi.org/10.26219/heal.aueb.6067

From the above figures it is obvious that the $CUSUM_t$ statistics stay within the 95% confidence bands. On the contrary the $CUSUMSQ_t$ statistics do not stay within the 95% confidence intervals. Hence, due to this fact and due to the fact that there is a lot of variability in the values of some of the β's estimates, this is some evidence for instability in the coefficients of the under testing linear model. After applying the Recursive Regression Analysis we are now going to apply the Rolling Regression Analysis in order to examine if there are structural breaks in our dataset using the under testing model. In the following lines we present the results of this method and especially the plots with the rolling coefficients and the rolling scale.

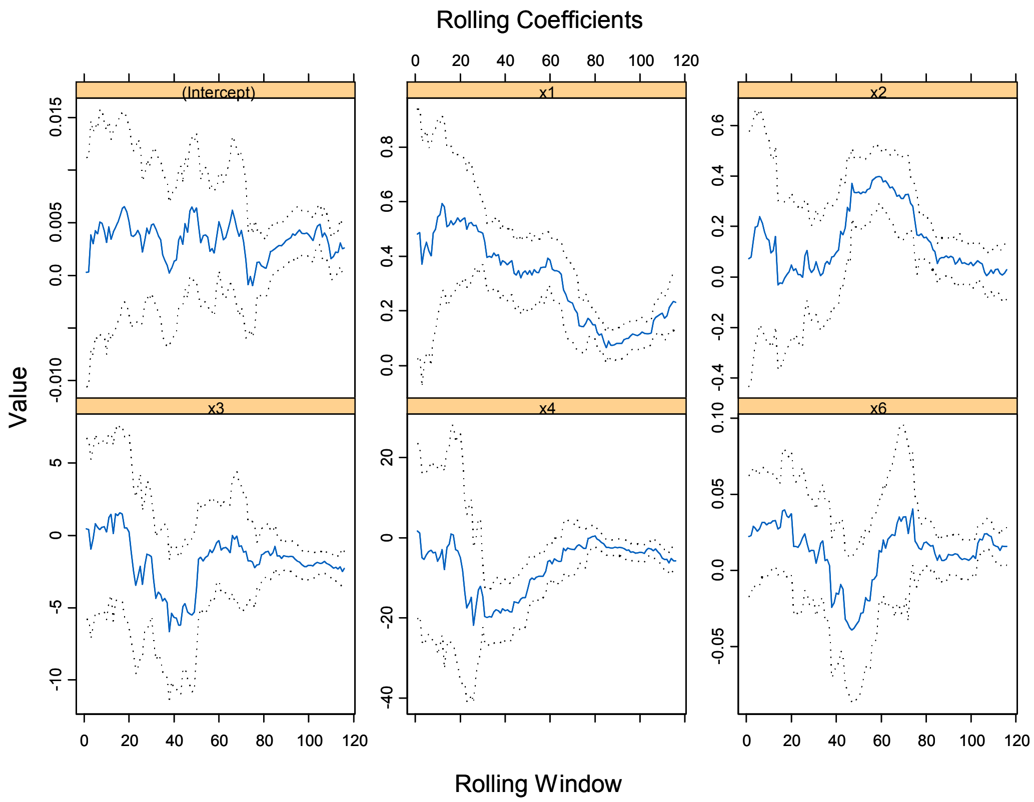


Figure 5.11 Rolling Regression coefficient estimates for the model used in the CSTHFCI index analysis

https://doi.org/10.26219/heal.aueb.6067

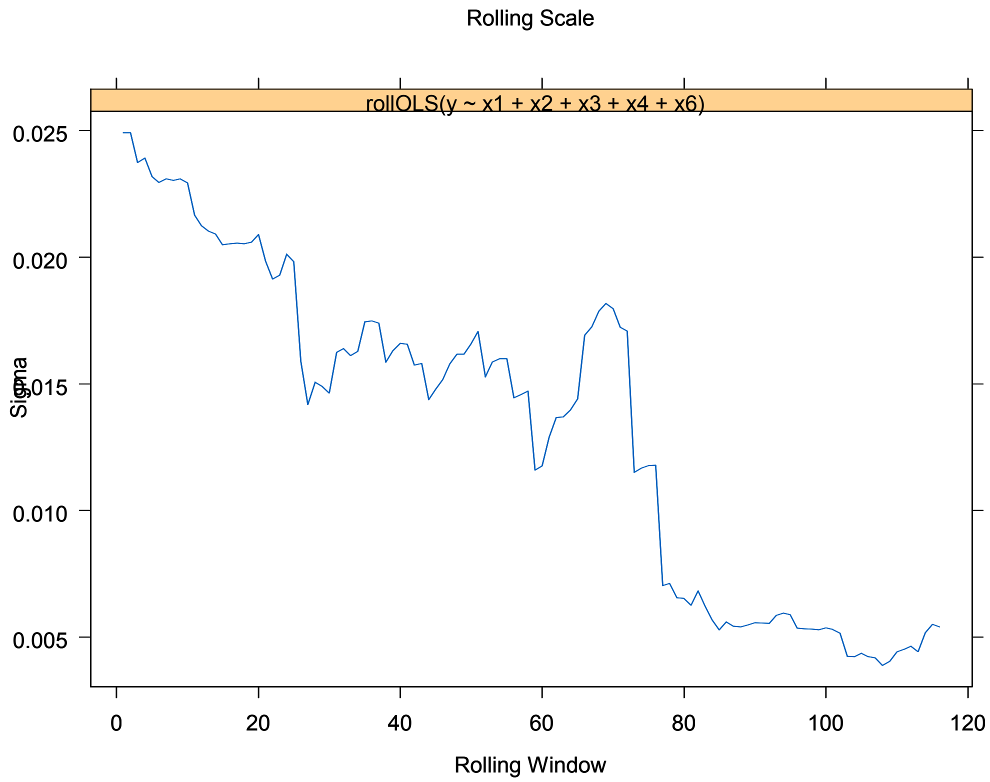


Figure 5.12 Rolling Regression estimates for the residual standard error of the model used in the CSTHFCI index analysis

From the graphs of the rolling estimates it is clear that the $\hat{a}$ coefficient seems to remain stable in a specific area of values, across the sample. On the contrary the $\hat{\beta}$'s coefficients vary a lot across the sample. We see that in all the cases there is a lot of variability in the values of each $\hat{\beta}$ coefficient. This is an indication that there exists a structural break in our under testing data. Also the residuals scale estimates $\hat{\sigma}$ are decreasing from the beginning of the sample until the 60^{th} observation where there takes place an increase in the values of the $\hat{\sigma}$ estimate until almost the 70^{th} observation. From this observation and on, there takes place a decrease in the $\hat{\sigma}$ values until the end of the sample. This decrease and increase in the series of the residual scale estimates $\hat{\sigma}$ is also an indication that there exists a structural break in our data. After the F-test, the BP method, the Recursive and the Rolling Regression Analysis we applied the fluctuation tests in our data in order to check whether there exists or not any structural break. Again, as in the HFRCI

 https://doi.org/10.26219/heal.aueb.6067

index case, as we can see from the corresponding plots (figure 5.7) that none of the fluctuation tests can distinquish any structural breaks. In other words, the fluctuation tests do not reject the null hypothesis of parameter constancy over the time. So based on the fluctuation tests there are no breaks detected in the model that we specified for the CSTHFCI index case. In the following table we present the results for the parameters of the estimated models for the two segments that were created in the case of the F-test, the BP method, the Recursive and the Rolling Regression analysis, where we estimated that there exists a structural change in the data.

	First Segment	**Second Segment**
α	0.002895 (0.002447)	0.002174 (0.001069)*
$\beta_{S\&P}$	0.447514 (0.060221)*	0.145503 (0.026370)*
β_{SCMLC}	0.247132 (0.065622)*	0.117322 (0.040985)*
β_{10Y}	-3.129831 (1.416903)*	-2.260629 (0.561458)*
$\beta_{CredSpr}$	-14.041235 (2.957212)*	-3.654631 (1.050000)*
β_{PTFSFX}	0.021614 (0.011541)	0.016016 (0.006190)*
σ^2	0.0003944196	0.00007182562

* Statistically significant

Table 5.6 Estimations and standard deviations of the most probable model in the case of the CSTHFCI index

We conclude from the above results that the variance in the first segment is larger than that in the second segment. Also it is obvious that there are differences in the risk exposures between the two segments. There is a decrease in the intercept, the $\beta_{S\&P}$, the β_{SCMLC} and the β_{PTFSFX} factors and an increase in the β_{10Y} and the $\beta_{CredSpr}$ factors. Furthermore we conclude that in the first segment all the coefficients except the intercept α and the β_{PTFSFX} factor are statistically significant and in the second segment that all the coefficients are statistically significant. These results are consistent with the conditions that existed in the market before the Technology and Internet

https://doi.org/10.26219/heal.aueb.6067

boom in 2000 and the conditions that existed in the market after the above crisis.

5.3 Application of the on-line monitoring method to the Hegde Funds Composite Indices

In this subsection we are dealing with the on-line monitoring method. This method is of prime importance in the structural break field. In fact it is the most useful method in the structural break models due to the fact that with this method we are able to estimate future breaks in a data series that we have in hand and which, day by day, is enriched by new data arriving. Especially we have some data which we treat as the history data and for which we assume a regression model and next we are testing whether new data that arrive are consistent with the specific model that is already known from the history data. We could say that it is more realistic, more important and more interesting to test for future structural breaks because in the real world new data arrive all the time and so we want to estimate or make a prediction for the future and not for the past. There is no practical sense to estimate a break in a data series that is «old» and this break might be located in a past period from the time being. Always our aim is to try to predict the future more accurately and with the lower loss.

The above statements will be clearer with the example that follows. Let's imagine in our case a hedge fund. The manager in this hedge fund has collected some data which reflect some past returns of this hedge fund. The manager in order to plan the hedge fund's strategy wants to know what about is going to happen in the future in the global economy. Having the data set of his hedge fund if he examines this series for structural breaks the only result that he might find is a break or some breaks that occurred in the past. That is he detects some structural breaks corresponding to some events that already have happened and belong to history. These results won't help him plan his future strategy. On the contrary, if he uses the data that he has as a history period and also collects the daily data that arrive for the hedge fund he can use a structural break model in order to estimate a break or breaks in the

https://doi.org/10.26219/heal.aueb.6067

future. This case is important for the structure and the future life of the hedge fund because if the manager estimates in some way the future and the possible breaks he can change his strategy and his investment tactics in order to avoid some future economic dangers for the hedge funds, or he can also change his investment tactics if he examines that the future, based on the events that might take place, reveals some lucrative economic opportunities for the hedge fund that he manages. We see that the on-line monitoring method is much more important than the attempt of estimating breaks in the past.

Now, having described the on-line monitoring method we are going to apply this method to our data. Especially we are going to use the same model and data as before with the only difference being that the data will be cut in two clusters. The first cluster will be the history data and the second cluster will be the incoming data, the so called «future data». We will act as if we know the history data and don't know the future data. Having estimated before the number of breaks and the break-point in every model (HFRCI and CSTHFCI case) we are going to cut the data before these breakpoints (history data) and the rest of the data will be the future data. By applying this method we are going to be able to check and test in which point (time period) will the breakpoints be observed. In fact we already know these points but with the on-line monitoring method we test how «fast» will these points be detectable with the model that we are dealing with. It is obvious that due to the fact that we have cut the data in history and incoming sample, the breakpoints will be observed in later points than they were observed using all the sample data as history sample. It is also obvious that it will be better if these breakpoints are observed closer to the estimated breakpoint position.

5.3.1 On-line monitoring method applied to the HFRCI index

5.3.1.1 Moving Estimates Method (Bandwidth h=1)

For the HFRCI index we found that there exists a model with a single break in this data series and the estimated breakdate is February 2001. These

https://doi.org/10.26219/heal.aueb.6067

results were obtained with the F-test, the Bai-Perron method (BP) and the Recursive and Rolling Regression Analysis. Let's now consider that we want to monitor the model specified for the HFRCI index around 2000 for structural breaks using years 1994-1999 as the history period. We already know that there exists a break somewhere around that period but we act as if we didn't know this for the aim of the on-line monitoring method. Firstly we cut the complete data in the historic sample and the incoming data (future sample). The idea is that we use the years 1994-1999 as the history sample and then when new data (future data) arrive we integrate them in our existing model and data. Thus, by this way it's like we are a manager or a researcher in 1999 and we want to find out whether the model which was calculated for the history sample is stable or unstable when new data arrive. By doing this we are trying to detect any possible future breaks. For this procedure to take place we use a significance level of 5%.

	Parameter estimate on history	**Last parameter estimate**
α	0.004162902	0.002544072
$\beta_{S\&P}$	0.437989787	0.303945903
β_{SCMLC}	0.391490318	0.274370256
β_{10Y}	-0.407735093	-1.454691469
$\beta_{CredSpr}$	-4.895352452	-1.936597746
Break detected at observation #112		

Table 5.7 Estimations of the most probable model using the on-line monitoring method via the Moving Estimates test with bandwidth h=1 in order to predict the possible breaks in the case of the HFRCI index

https://doi.org/10.26219/heal.aueb.6067

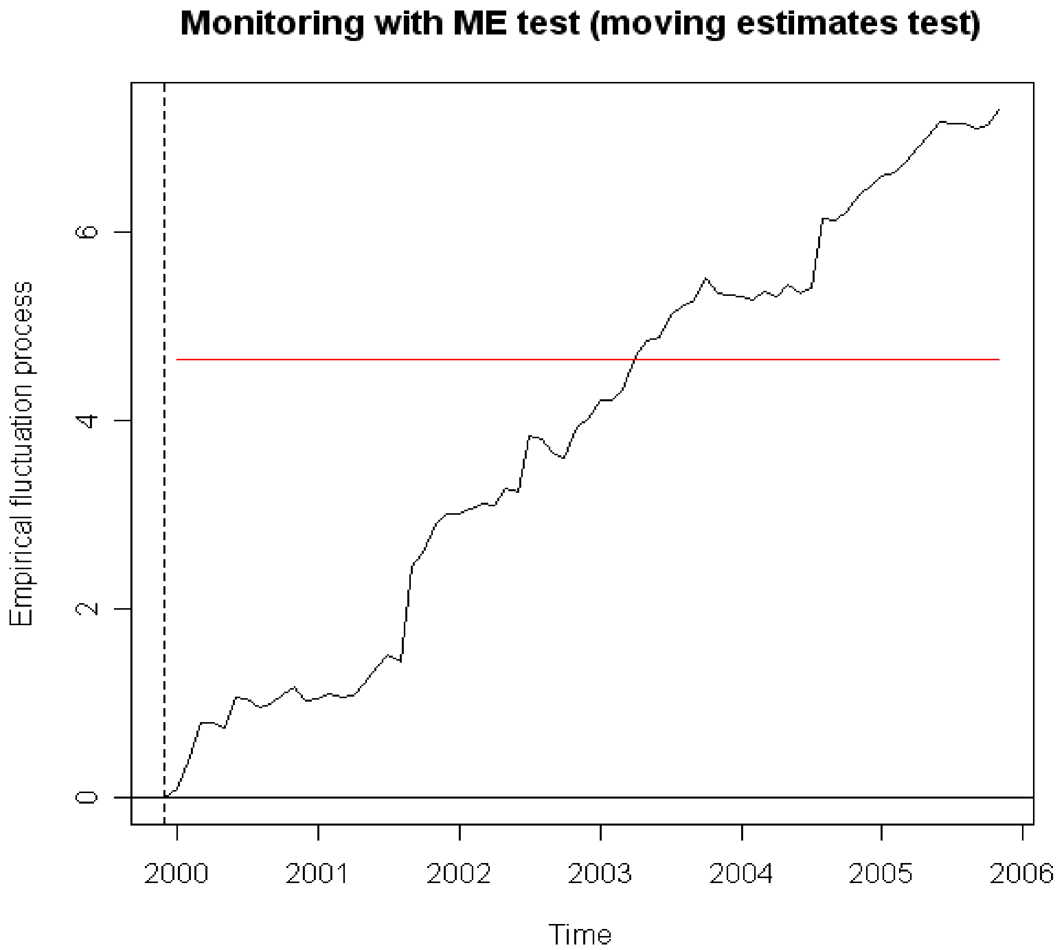


Figure 5.13 Plot of the on-line monitoring method in the case of the HFRCI index using the Moving Estimates test with bandwidth h=1

From the above results we conclude that there occurs a structural break at point 112. This point corresponds to April 2003. Also looking the above figure of the ME test process with the history sample on the left and the monitoring sample on the right we reach at the same conclusion as before as we see that around 2003 there exists a shift. That is to say that using the sample which corresponds to 1994-1999 data as the history sample and using also the model of Fung and Hsieh (2004) we predicted that there is going to be a structural break at the point 112 of the full sample, namely at April 2003. Also we know from the previous subchapter that there exists a structural break at position 86 which corresponds to February 2001, using the complete

https://doi.org/10.26219/heal.aueb.6067

dataset. So, with the on-line monitoring method we predicted the existence of the structural break 2 years and 2 months later.

5.3.1.2 Moving Estimates Method (Bandwidth h=0.5)

We saw from the previous results that the structural break is detected at the point 112 using as bandwidth h=1 in the statistical process of structural break detection. It would be useful if we did the previous process using as bandwidth h=0.5 and compare the results that appear using as bandwidth h=1. In the following lines we illustrate the results of the above idea.

	Parameter estimate on history	**Last parameter estimate**
α	0.004162902	0.002935262
$\beta_{S\&P}$	0.437989787	0.238398450
β_{SCMLC}	0.391490318	0.238715598
β_{10Y}	-0.407735093	-1.865484909
$\beta_{CredSpr}$	-4.895352452	-3.203891238
Break detected at observation #103		

Table 5.8 Estimations of the most probable model using the on-line monitoring method via the Moving Estimates test with bandwidth h=0.5 in order to predict the possible breaks in the case of the HFRCI index

https://doi.org/10.26219/heal.aueb.6067

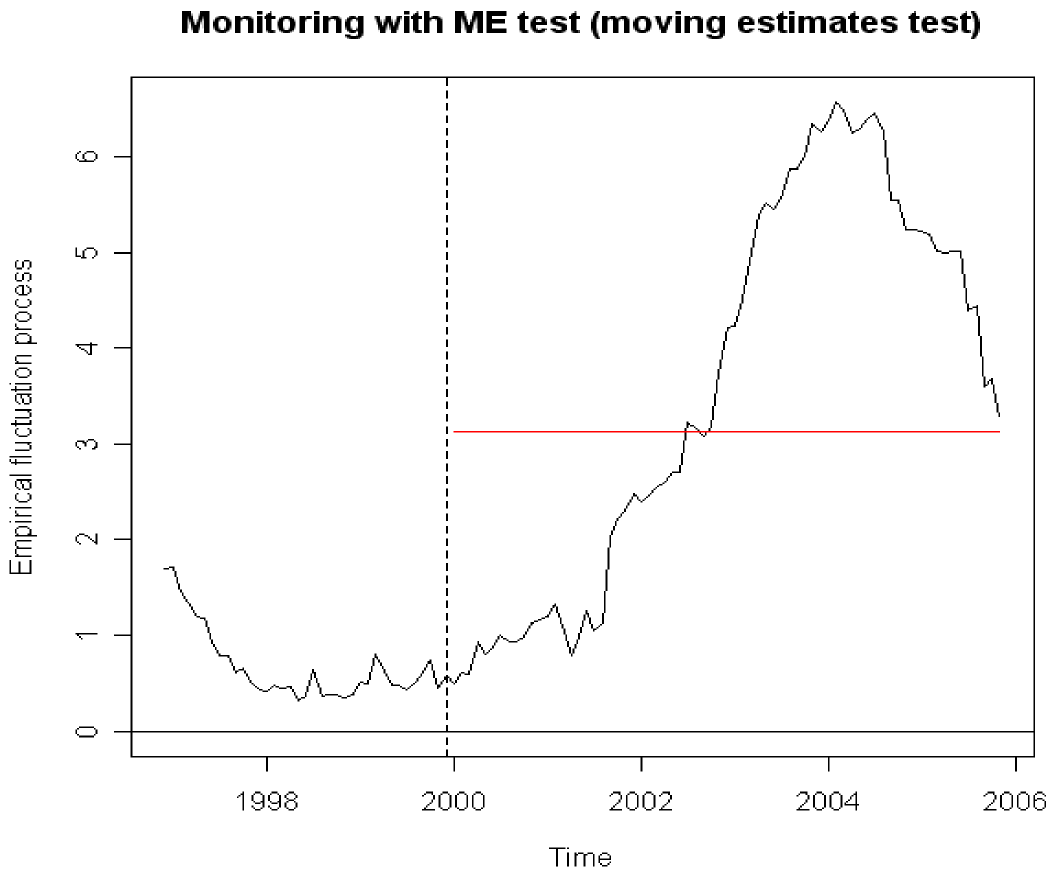


Figure 5.14 Plot of the on-line monitoring method in the case of the HFRCI index using the Moving Estimates test with bandwidth h=0.5

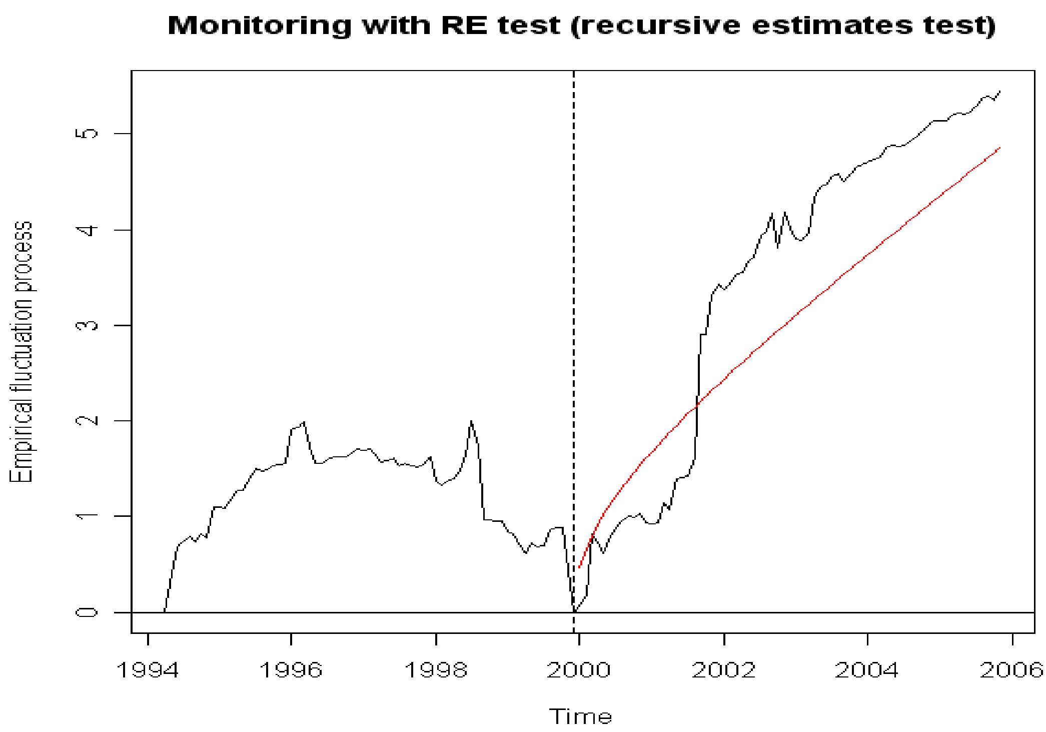
From the results we conclude that there occurs a structural break at point 103. So, it is obvious that we discover the structural break even 9 observations earlier using the bandwidth h=0.5 instead of h=1. This structural break detection corresponds to July 2002. Using the bandwidth h=1 the structural break that occurred corresponded to April 2003. Consequently, the process with h=0.5 returns better prediction results than the process with h=1. The process in which we use h=0.5 detects the structural break 1 year and 5 months later from the already known structural break time of February 2001 that occurs in the series of HFRCI index using all the sample of the data. On the contrary the process in which we used the bandwidth h=1 detects the structural break 2 years and 2 months later from the position of the structural break that corresponds to February 2001. So, the process in which the bandwidth h=0.5 is used has better prediction properties than the process in

 https://doi.org/10.26219/heal.aueb.6067

which the bandwidth h=1 is used. However, despite the difference in the prediction of the structural breaks between the two processes, fact that is reasonable to happen, both the two prediction results we can say that emphasize and complete in some way the results of the history tests.

5.3.1.3 Recursive Estimates Method (Bandwidth h=0.5 and h=1)

In this section we monitor the model specified for the HFRCI index around 2000 for structural breaks using years 1994-1999 as the history period via the Recursive Estimates test. We already know that the single break model is favoured and the estimated breakdate is February 2001. Also using the on-line monitoring method, the Moving Estimates method detected and estimated the structural break points in the position of April 2003 and July 2002 using as bandwidth h=1 and h=0.5 respectively. In the following, we present the results of the application of the Recursive Estimates test in the HFRCI index. At this point we shall mention that in the Recursive Estimates method both the bandwidth values h=1 and h=0.5 revealed the same results.



https://doi.org/10.26219/heal.aueb.6067

Figure 5.15 Plot of the on-line monitoring method in the case of the HFRCI index using the Recursive Estimates test with bandwidth values h=1 and h=0.5

	Parameter estimate on history	**Last parameter estimate**
α	0.004162902	0.003539799
$\beta_{S\&P}$	0.437989787	0.349297687
β_{SCMLC}	0.391490318	0.299386060
β_{10Y}	-0.407735093	-0.760517896
$\beta_{CredSpr}$	-4.895352452	-1.893884090
Break detected at observation #75		

Table 5.9 Estimations of the most probable model using the on-line monitoring method via the Recursive Estimates test with bandwidth h=0.5 and h=1 in order to predict the possible breaks in the case of the HFRCI index

We conclude from the above results that the Recursive Estimates test using as bandwidth h=1 and h=0.5 identifies a structural break at the point 75, which corresponds to March 2000. It is obvious that the Recursive Estimates method identifies the structural break point earlier than the Bai-Perron method and the Moving Estimates method did. This structural break point can be associated with the Technology and Internet bubble in March 2000.

5.3.2 On-line monitoring method applied to the CSTHFCI index

5.3.2.1 Moving Estimates Method (Bandwidth h=1)

For the CSTHFCI index we found that there exists a model with a single break in this data series and the estimated breakdate is April 2000. These results were obtained using the F-test, the Bai-Perron method (BP) and the Recursive and Rolling Regression Analysis. Consider that someone would like to monitor the model specified for the CSTHFCI index around 1999 for structural breaks using years 1994-1999 as the history period. We proceed as we did before with the HFRCI index case. Firstly we cut the complete data in the historic sample and the incoming data (future sample). We use the years 1994-1999 as the history sample and then when we have the incoming data (future data) we integrate them in our existing model. For this procedure to take place we use a significance level of 5% as we did with the previous analyses.

	Parameter estimate on history	**Last parameter estimate**
α	0.001935879	0.002427847
$\beta_{S\&P}$	0.452216155	0.147728606
β_{SCMLC}	0.141355200	0.247305701
β_{10Y}	-3.152007611	-1.749401947
$\beta_{CredSpr}$	-15.261813272	-2.912038741
β_{PTFSFX}	0.020718668	0.010805047
Break detected at observation #105		

Table 5.10 Estimations of the most probable model using the on-line monitoring method via the Moving Estimates test with bandwidth h=1 in order to predict the possible breaks in the case of the CSTHFCI index

https://doi.org/10.26219/heal.aueb.6067

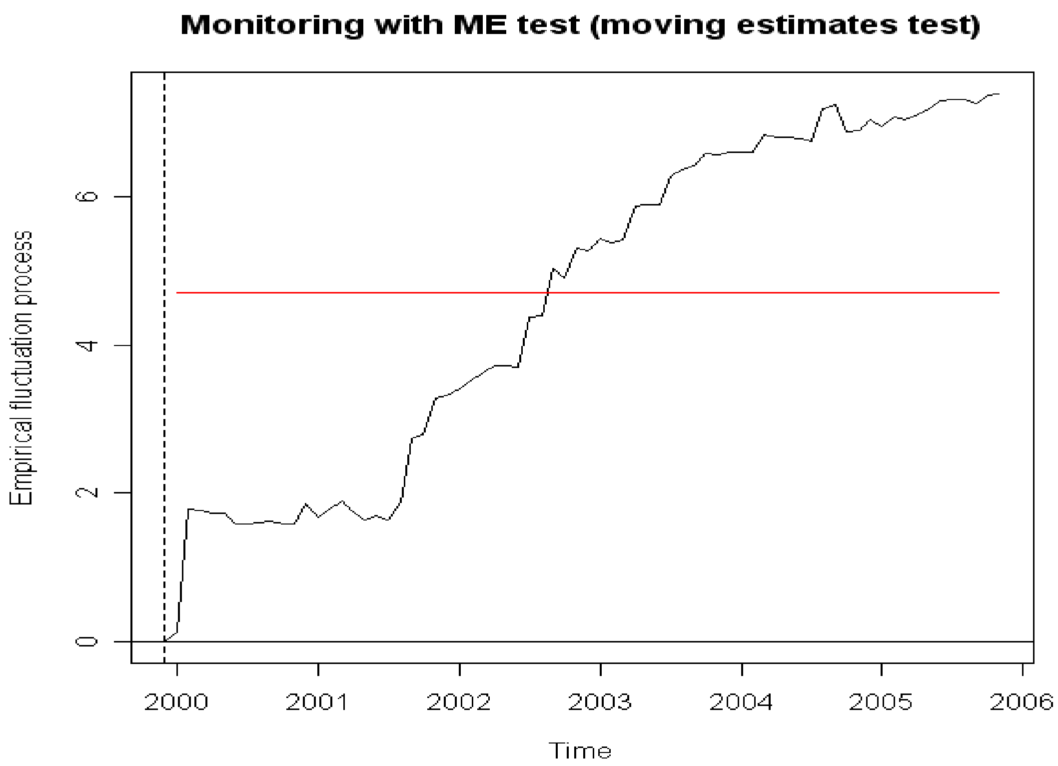


Figure 5.16 Plot of the on-line monitoring method in the case of the CSTHFCI index using the Moving Estimates test with bandwidth value h=1

Looking at the above results and the corresponding figure we can conclude that there occurs a structural break at point 105. This point corresponds to September 2002. That is, using the sample which corresponds to 1994-1999 data as the history sample and using also the model of Fung and Hsieh (2004) we predicted that there is going to be a structural break at the point 105 of the full sample of the CSTHFCI index case, namely at September 2002. From the results of the Bai-Perron method we already know that there exists a structural break at position 76 which corresponds to April 2000. So we are able to say that using the on-line monitoring method we predicted the existence of the structural break 2 years and 5 months later. These results were obtained using as bandwidth in the on-line monitoring process the value h=1.

https://doi.org/10.26219/heal.aueb.6067

5.3.2.2 Moving Estimates Method (Bandwidth h=0.5)

We saw from the previous results that the structural break is detected at the point 112 using as bandwidth h=1 in the statistical process of structural break detection. It would be useful if we did the previous process using as bandwidth h=0.5 and compare the results that appear using as bandwidth h=1. In the following lines we illustrate the results of the above idea.

	Parameter estimate on history	**Last parameter estimate**
α	0.001935879	0.002775767
$\beta_{S\&P}$	0.452216155	0.169532740
β_{SCMLC}	0.141355200	0.037687223
β_{10Y}	-3.152007611	-2.207580751
$\beta_{CredSpr}$	-15.261813272	-5.122999963
β_{PTFSFX}	0.020718668	0.017654933
Break detected at observation #96		

Table 5.11 Estimations of the most probable model using the on-line monitoring method via the Moving Estimates test with bandwidth h=0.5 in order to predict the possible breaks in the case of the CSTHFCI index

https://doi.org/10.26219/heal.aueb.6067

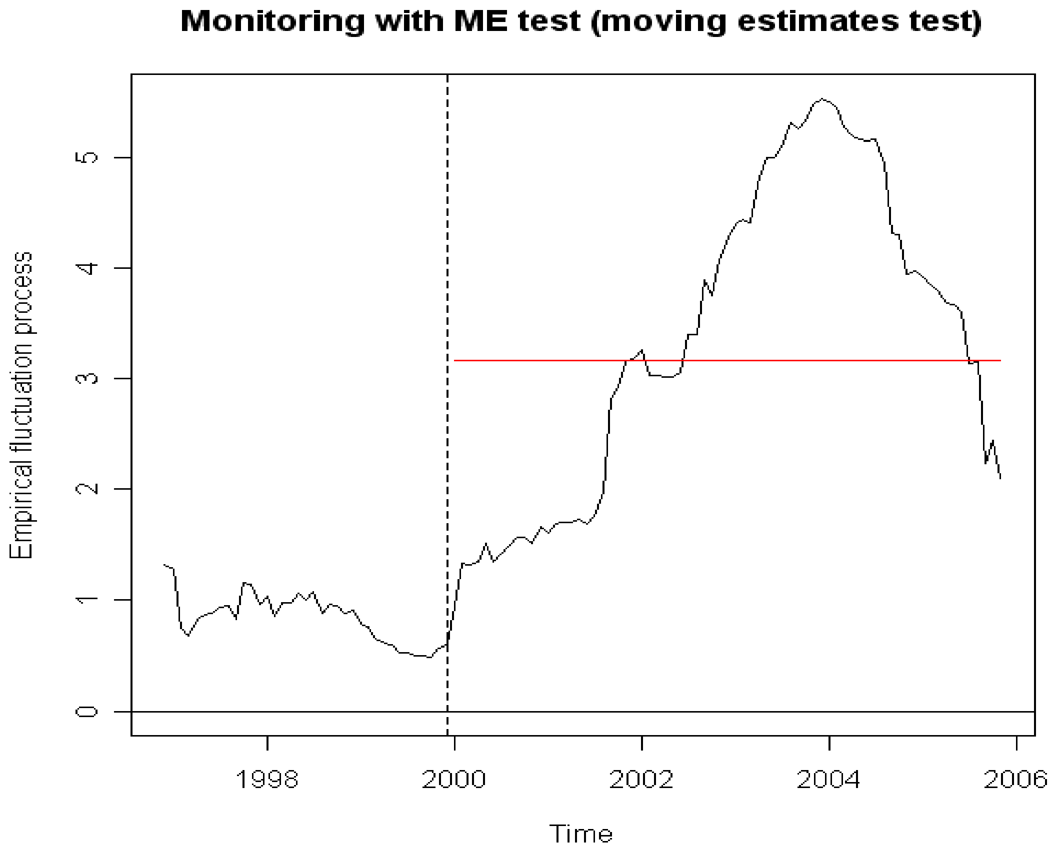


Figure 5.17 Plot of the on-line monitoring method in the case of the CSTHFCI index using the Moving Estimates test with bandwidth value h=0.5

From the above results for the h=0.5 case we are able to conclude that there occurs a structural break at point 96. So, it is obvious that we discover the structural break 9 observations earlier as we used the bandwidth h=0.5 instead of h=1. This structural break detection corresponds to December 2001. Using the bandwidth h=1 the structural break that occurred corresponded to September 2002. Consequently, the process with h=0.5 returns better prediction results than the process with h=1. The process in which we use h=0.5 detects the structural break 1 year and 8 months later from the already known structural break time of April 2000 that occurs in the series of CSTHFCI index using all the sample of the data. On the contrary the process in which we used the bandwidth h=1 detects the structural break 2 years and 5 months later from the position of the structural break that corresponds to April 2000. However, despite the differences in the prediction of the

https://doi.org/10.26219/heal.aueb.6067

structural breaks between the two processes, both the prediction results emphasize and complete in some way the results of the history tests.

5.3.2.3 Recursive Estimates Method (Bandwidth h=0.5 and h=1)

In this section we monitor the model specified for the CSTHFCI index around 2000 for structural breaks using years 1994-1999 as the history period via the Recursive Estimates test. We already know that in this data series the single break model is favoured and the estimated breakdate is April 2000. Also using the on-line monitoring method, the Moving Estimates method detected and estimated the structural break points in the positions of September 2002 and December 2001 using as bandwidth h=1 and h=0.5 respectively. In the following lines we present the results of the application of the Recursive Estimates test in the CSTHFCI index. In the Recursive Estimates test both the bandwidth values h=1 and h=0.5 revealed the same results.

	Parameter estimate on history	**Last parameter estimate**
α	0.001935879	0.003540182
$\beta_{S\&P}$	0.452216155	0.255504641
β_{SCMLC}	0.141355200	0.165776577
β_{10Y}	-3.152007611	-2.330089198
$\beta_{CredSpr}$	-15.261813272	-6.042979185
β_{PTFSFX}	0.020718668	0.015954958
Break detected at observation #74		

Table 5.12 Estimations of the most probable model using the on-line monitoring method via the Recursive Estimates test with bandwidth h=1 and h=0.5 in order to predict the possible breaks in the case of the CSTHFCI index

https://doi.org/10.26219/heal.aueb.6067

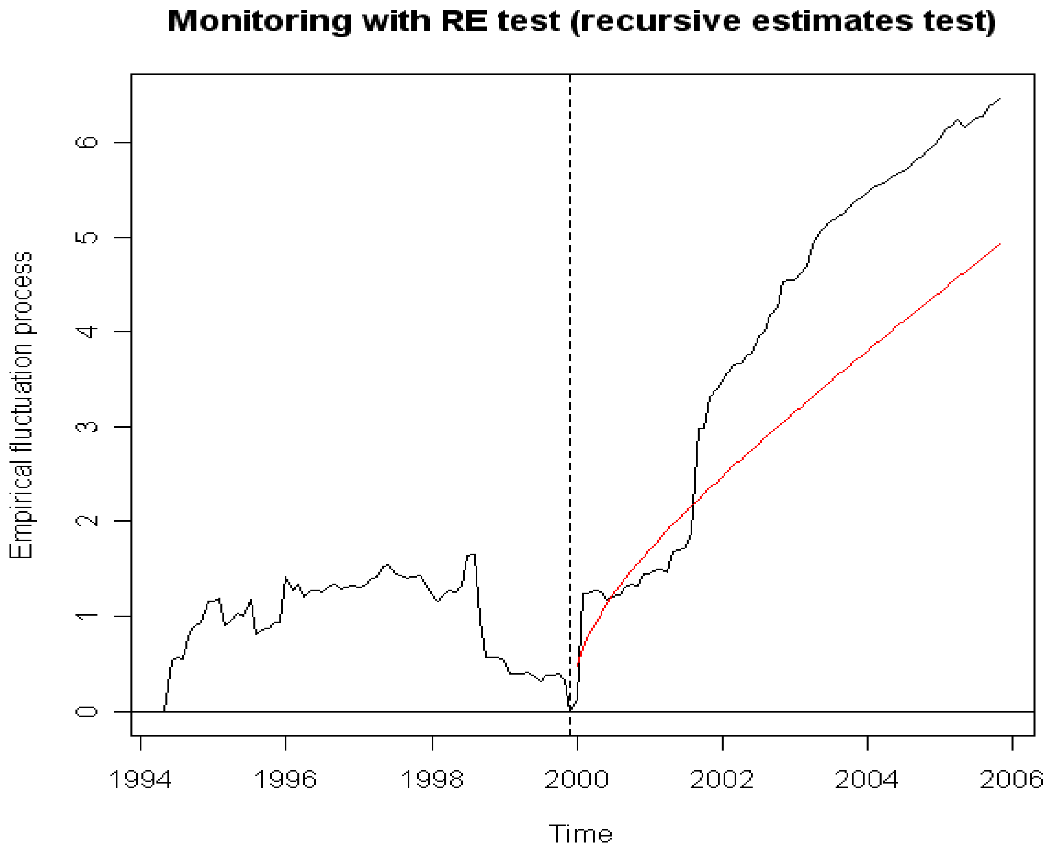


Figure 5.18 Plot of the on-line monitoring method in the case of the CSTHFCI index using the Recursive Estimates test with bandwidth values h=1 and h=0.5

We conclude from the above results that the Recursive Estimates test using as bandwidth h=1 and h=0.5 identifies a structural break at the point 74, which corresponds to February 2000. It is obvious that the Recursive Estimates method identifies the structural break point earlier than the Bai-Perron method and the Moving Estimates method did. However we could say that the Bai-Perron method identified the structural break point close to the structural break point that the Recursive Estimates test did because the difference between these two break points is 2 months. This structural break point can be associated with the Technology and Internet bubble in March 2000.

https://doi.org/10.26219/heal.aueb.6067

https://doi.org/10.26219/heal.aueb.6067

Chapter 6

Conclusions

6.1 Conclusions for the analysis of the HFRCI and the CSTHFCI indices

The results of the application of the structural break point models and the structural change tests to the data of the HFRCI and CSTHFCI indices suggest that the break point risk factor models can identify the changes that occur in the above two hedge fund composite indices. Firstly we have found that both indices have significant risk exposures to the two equity ABS factors with the exposures of the HFRCI being higher. These significant exposures decrease after 2001 for both indices. On the contrary the CSTHFCI index has higher risk exposures to the fixed income ABS factors and furthermore it is affected by the trend following factor on currencies. Another important conclusion that we reach by analyzing the above indices is that in both indices there occurs a structural break point around the years 2000 and 2001 which can be attributed to the Technology and Internet bubble. In the specific years all the hedge funds overweighted high technology stocks in their portfolios (Brunnermeier and Nagel, 2004) and this strategy is reflected to the values of the HFRCI and the CSTHFCI indices. Lastly, we can say that applying the on-line monitoring method in the above indices in order to predict the possible breaks using a subsample of the initial sample as the history data we found that the Recursive Estimates test performed better than the Moving Estimates test. The Recursive Estimates test using as bandwidth the values h=1 and h=0.5 predicted the break earlier than the Moving Estimates test with bandwidth values h=1 and h=0.5 did.

https://doi.org/10.26219/heal.aueb.6067

6.2 Conclusions for the simulated data

Comparing the results that we reached to using the Moving Estimates method with the results that we reached to using the Recursive Estimates method we see that using as method of the on-line monitoring procedure the recursive estimates method we predicted the existence of the structural break 2 observations earlier than the Moving Estimates method using as bandwidth value the value h=1 and 1 observation earlier than the Moving Estimates method using as bandwidth value the value h=0.5. We see that in the case of the simulated data, as in the case of the real hedge funds data, the Recursive Estimates test performed better than the Moving Estimates test in the procedure of estimating the possible breaks via the on-line monitoring method.

https://doi.org/10.26219/heal.aueb.6067

References

Amin, G.S., and Kat, H.M. (2003). Hedge Fund Performance 1990-2000: Do the "Money Machines" Really Add Value? , *The Journal of Financial and Quantitative Analysis*, 38, 251-274

Andrews, D.W.K. (1993). Tests for Parameter Instability and Structural Change with Unknown Change Point, *Econometrica*, 59, 817-858

Andrews, D.W.K. and Ploberger, W. (1994). Optimal Tests when a Nuisance Parameter is Present Only Under the Alternative, *Econometrica*, 62, 1383-1414

Bai, J. and Perron, P. (1998). Estimating and Testing Linear Models with Multiple Structural Changes, *Econometrica*, 66, 47-78

Bai, J. and Perron, P. (2003). Critical values for multiple structural change tests, *Econometrics Journal*, 6, 72–78

Brown, R.L., Durvin, J., Evans, J.M. (1975). Techniques for Testing the Constancy of Regression Relationships over Time, *Journal of the Royal Statistical Society*, 37, 149-192

Chow, G.C. (1960). Tests of Equality Between Sets of Coefficients in two Linear Regressions, *Econometrica*, 28, 591-605

Chu, C.S.J., Hornik, K., Kuan, C.M. (1995a). The Moving Estimates test for Parameter Stability, *Econometric Theory*, 11, 699-720

Chu, C.S.J., Hornik, K., Kuan, C.M. (1995b). MOSUM Tests for Parameter Constancy, *Biometrika*, 82, 603-617

https://doi.org/10.26219/heal.aueb.6067

Chu, C.S.J., Stinchcombe, M., White, H. (1996). Monitoring Structural change, *Econometrica*, 64, 1045-1065

Fung, W. Hsieh, D.A. (2004). Hedge Fund Benchmarks: A Risk-Based Approach, *Financial Analysts Journal,* 60, 65-80

Kouwenberg, R., Ziemba, W.T. (2007). Incentives and Risk Taking in Hedge Funds, *Journal of Banking and Finance*, 31

Kramer, W., Ploberger, W., Alt, R. (1988). Testing for structural change in dynamic models, *Econometrica*, 56, 1355-1369

Leisch, F., Hornik, K., Kuan, C.M. (2000). Monitoring Structural Changes with the generalized fluctuatiton test, *Econometric Theory*, 16, 835-854

Meligkotsidou, L. and Vrontos,I.D. (2008). Detecting structural breaks and identifying risk factors in hedge fund returns: A Bayesian approach, *Journal of Banking and Finance*

Ploberger, W., Kramer, W., Kontrus, K. (1989). A New Test for Structural Stability in the Linear Regression Model, *Journal of Econometrics*, 40, 307-318

Ploberger, W., Kramer, W. (1992). The CUSUM Test with OLS Residuals, *Econometrica,* 60(2), 271-285

Vrontos, S.D., Vrontos, I.D., Giamouridis, D. (2008). Hedge Fund Pricing and Model Uncertainty, *Journal of Banking and Finance*, 32, 741-753

Zeileis, A., Leisch, F., Kleiber, C., Hornik, K. (2005). Monitoring Structural Change in Dynamic Econometric Models, *Journal of Applied Econometrics*, 20, 99-121