# STRUCTURAL BREAK DETECTION AND MODEL SELECTION: COMPARISON OF REGULARIZATION TECHNIQUES WITH AUTOMETRICS

Pakistan Institute of Development Economics


***By***

**Sara**

**PIDE2017FPHDETS04**

**Supervisor**

**Dr. Amena Urooj**

**PIDE School of Economics**

**Pakistan Institute of Development Economics,**

**Islamabad**


**2023**

# Author's Declaration

I Ms. Sara hereby state that my PhD thesis titled **"Structural Break Detection and Model Selection: Comparison of Regularization Techniques with Autometrics"** is my own work and has not been submitted previously by me for taking any degree from **Pakistan Institute of Development Economics, Islamabad'** or anywhere else in the country/world.

At any time if my statement is found to be incorrect even after my Graduation the university has the right to withdraw my PhD degree.

Ms. Sara

PIDE2017FPHDETS04

# Plagiarism Undertaking

I solemnly declare that research work presented in the thesis titled **"Structural Break Detection and Model Selection: Comparison of Regularization Techniques with Autometrics"** is solely my research work with no significant contribution from any other person. Small contribution/help wherever taken has been duly acknowledged and that complete thesis has been written by me.

I understand the zero tolerance policy of the **HEC** and **Pakistan Institute of Development**

declare that no portion of my thesis has been plagiarized and any material used as reference is properly cited.

I undertake that if I am found guilty of any formal plagiarism in the above titled thesis even after award of PhD degree, the University reserves the rights to withdraw/revoke my PhD degree and that HEC and the University has the right to publish my name on the HEC/University Website on which names of students are placed who submitted plagiarized thesis.

Students/Author Signature:____________________

Ms. Sara

PIDE2017FPHDETS04

## Certificate of Approval

This is to certify that the research work presented in this thesis, entitled: **"Structural Break Detection and Model Selection: Comparison of Regularization Techniques with Autometrics"** was conducted by **Ms. Sara** under the supervision of **Dr. Amena Urooj** No part of this thesis has been submitted anywhere else for any other degree. This thesis is submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Econometrics from **Pakistan Institute of Development Economics, Islamabad.**

**Student Name:** Ms. Sara
PIDE2017FPHDETS04

Signature: ____________

**Examination Committee:**

a) **External Examiner:**
Dr. Zahid Asghar
Professor
Quaid-e-Azam University
Islamabad

Signature: ____________

b) **Internal Examiner:**
Dr. Saud Ahmed Khan
Assistant Professor,
PIDE, Islamabad

Signature: ____________

**Supervisor:**
Dr. Amena Urooj
Assistant Professor PIDE,
Islamabad

Signature: ____________

**Dr. Shujaat Farooq**
Head, PIDE School of Economics (PSE)
PIDE, Islamabad

Signature: ____________

**ABSTRACT**

The indicator Saturation method is a popular method for structural break and outlier detection that simultaneously detects the structural break/outliers in a model. Step Indicator Saturation (SIS) does not possess any restriction on the number or length of breaks, breaks at the start or end of observations. In contrast, IIS is a more efficient technique for handling outliers in cross-sectional modeling. The indicator saturation method uses Autometrics techniques for computation. However, the thriving model depends on the selection of significance level (with a significance level of 0.01 or 0.001 model drops the significant break, and with a nominal significance level 0.05, it retains irrelevant breaks).

Meanwhile, regularization techniques efficiently deal with the saturated model even if the regressors are far greater than the number of observations. This study uses well-known regularization techniques, Least Absolute Subset Selection Operator (LASSO), Adaptive Least Absolute Subset Selection Operator (AdaLASSO), Minimax Concave Penalty (MCP), and Smoothly Clipped Absolute Deviation (SCAD) for structure break and outlier detection and compared with Autometrics. We assess the performance of regularization techniques in terms of Gauge ('Size'), Potency ('Power'), RMSE, and MAE with different Data Generating Processes (DGP) in the simulation study. For structure break detection in simulation experimental, we consider three different scenarios single break at the end of observation, single break at the start of observations, and unknown break with two-step indicators. However, for outlier detection, we consider two different scenarios outliers with AR(1) process and different magnitudes. The second simulation experiment was with a static multivariate model with varying outlying observations of 5%, 10%, and 20% obtained by assuming $\varepsilon_i{\sim}(0, \sigma + 4)$ *and* $\varepsilon_i{\sim}(0, \sigma + 6)$.. The final simulation experiment is based on the covariate and its lag selection with varying autocorrelation coefficients and sample sizes in time series modeling.

The simulation result indicates that MCP and SCAD perform near Autometrics in average potency with fixed tuning parameter in single and multiple breaks detections. On the other hand, LASSO estimates work well for single break detection, whereas it selects more irrelevant dummy indicators for multiple breaks. The SCAD and MCP perform better in forecasting and covariate selection in simulation studies with a 4SD outlier (20% and 5% outlying observations), nonetheless, as compared to Autometrics  Meanwhile, LASSO and AdaLASSO select more covariates and possess higher RMSE than SCAD and MCP. Overall, SCAD and MCP possess least RMSE than Autometrics. Although, for covariate and its lag selection, compared to Autometrics, the WLAdaLASSO outperforms in covariate and its lag selection as well as in forecasting, especially when there is a greater linear dependency between predictors. In contrast, the efficiency of Autometrics in potency decreases with a strong linear dependence between predictors. However, under the large sample and weak linear dependency between predictors, the Autometrics potency approaches to 1 and gauge approaches to $\alpha$.

In contrast, LASSO, SCAD, and MCP, select more covariates and possess higher RMSE than WLAdaLASSO and Autometrics. The real data analysis has been performed for each simulation experiment on a popular macroeconomic variable of Pakistan. The real data analysis is aligned with simulation findings.

**Keywords:**

Step Indicator Saturation, Data Generating Process, Regularization Techniques, Autometrics, Simulation

## Dedication

To my late father, mother, my sisters, and my brothers, I dedicate my thesis with the utmost love and thanks. My accomplishments have been motivated by your unfailing love, support, and encouragement. Thank you for being my pillars of strength and inspiring me to reach new heights.

# Acknowledgements

First and foremost, I am thankful to Almighty Allah, the most gracious, the most merciful and the magnificent, who bestowed me with the abilities and skills to complete my PhD. I bow in front of him in gratitude for all His blessings. I pay my heartiest tributes to the Holy Prophet Muhammad (Peace Be Upon Him) who is forever the source of guidance and knowledge for the humanity as a whole.

I would like to express my gratitude to my Supervisor, Dr. Amena Urooj for her supervision, support, and encouragement throughout this research work Without her precious support it would not be possible to conduct this research. In fact, I am privilege to carry out my research under the supervision of such outstanding professor.

I am also grateful to Dr. Zahid Asghar for reviewing my thesis and providing very constructive comments and suggestions for improving my research work. I am benefited and steered from the constructive comments, criticism and suggestions of Dr. Saud Ahmed Khan and Dr. Zahid Asghar. I would also like to express my gratitude to all my respectable teachers: Dr. Atiq ur Rehman, Dr. Hafsa Hina, and Dr. Saud Ahmed Khan for their nice attitudes and setting themselves as examples of great teachers.

I am indebted to my colleagues and friends who helped me throughout this tenure. Especially my discussions with Dr. Faridoon Khan regarding the simulation study and shrinkage methods. Special thanks go to Dr. Faridoon Khan, Nauman Ahmed, Iqra Iqbal, and Farah Naz, who always encouraged me during my PhD.

# TABLE OF CONTENTS

## LIST OF TABLES

## LIST OF FIGURES

## LIST OF ABBREVIATIONS

DGP Data Generating Process

SD Standard Deviation

IS Indicator Saturation

SIS Step Indicator Saturation

IIS Impulse Indicator Saturation

GUM Generalized Unrestricted Model

CV Cross-Validation

MCP Minimax Concave Penalty

SCAD Smoothly Clipped Absolute Deviation

LASSO Least Absolute Subset Selection Operator

AdaLASSO Adaptive Least Absolute Subset Selection Operator

WLAdaLASSO Weighted Lag Adoptive LASSO

CV Cross Validation

# CHAPTER 1

# INTRODUCTION

## 1.1 Background of Study

The ordinary least squares (OLS) approach has been a widely chosen technique among the numerous available methods in regression analysis because it is computationally straightforward and possesses the best linear unbiased estimate. However, it possesses a strong assumption on the distribution of the error term (ε), as the error term is normally distributed with mean zero and constant variance $\varepsilon \sim N(0, \sigma^2)$. This assumption is usually violated while dealing with real data analysis due to structural breaks and outliers in macroeconomic variables. The outlier and breaks in macroeconomic variables are due to socioeconomic and political instabilities, pandemics, and other technological revolutions, which is the significant cause of these distortions.

Since the last century, the world has altered dramatically in every measurable way, including World War I 1914-1918, the Spanish flu 1918–20, the Great Depression 1929, World War II 1939-1945, the oil crises 1970s, the Asian financial crisis 1997, the 2008–2012 financial crisis or Great Recession and COVID-19 pandemic, recently (Castle & Hendry, 2019a). These socioeconomic and political crises distort the flow of macroeconomic variables, leading to parameter instability, out-of-sample poor forecast, model misspecification, and possibly affecting variables and their lags in non-linear functions (Castle & Hendry, 2014). On the other hand, outliers are the primary cause of distortion in the distribution of error terms, which goes against the residuals' normality assumption.

Risks associated with least squares regression arise from outliers in the dependent and explanatory variables that, if undetected, might have a negative effect on the estimate (Zaman et al., 2001).

However, several real data analyses were performed in which OLS residuals failed to detect any outliers despite significant outliers (Rousseeuw and Leroy, 2005). Whereas, new statistical procedures have been proposed that are less susceptible to outliers; Rousseeuw, (1984) introduced the first feasible robust regression estimators (least median squares (LMS), least trimmed squares (LTS), and variations) that perform correctly even when a high number of outliers are present. Huber M Estimation, MM Estimation, Least Absolute Value Method (LAV), and S Estimation are robust approaches (Berk, 1990; Birkes & Dodge, 2011; Wilcox, 2011).

However, time series analysis structural breaks and outliers possess different meanings for model stability and accuracy. Structural breaks in the time series model impact the parameter consistency, whereas outliers do not affect the parameter stability but distort the model's residual. In contrast to outliers, the traditional structural break tests in time series modeling based on the prespecified model assumption, starting with Quandt, (1960), Farley & Hinich, (1970), Ploberger et al., (1989), Ploberger & Krämer, (1990), Perron, (1989), Perron & Vogelsang, (1992), Andrews, (1993), Perron, (2006), Hansen, (1992 and 2012), Jansen & Teräsvirta, (1996), Bai & Ng, (2002), and Bai & Perron (2003a, 2003b, 2003c, 2006). However, the model's accuracy in time series analysis is always unknown. Only one in a million models can be accurate; ‘‘Essentially, all models are wrong, but some are useful’’(Box, 1979). Based on the strong assumption of the prespecified model, what if the model is far from ‘accurate specifications’, and in what way to ‘repair’ it is constantly unclear, such methods can be ineffective for break detection (Castle & Hendry, 2019a).

Nevertheless, the Indicator Saturation (IS) method is based on univariate and multivariate analysis for structural beak and outlier detection. However, for break detection Castle & Hendry, (2019a) emphasis to use in univariate structure to overcome the problem of prespecified model specification in advance. IS methods are a popular method for outlier detection and structural break detection in the mean (Castle et al., 2015; Hendry et al., 2006; Castle et al., 2015a, 2012; Santos et al., 2008). Step-Indicator-Saturation (SIS) and Impulse Indicator Saturation (IIS) methods are well-known techniques of IS method for multiple breaks and outlier detection. The IIS method was initially designed to detect unknown numbers of outliers with unknown magnitudes at uncertain points in the sample, including the beginning and end of observations (Hendry et al., 2006). On the other hand, Step Indicator Saturation (SIS) method is a modified version of IS techniques for multiple structural break detection. Step Indicator Saturation (SIS) takes over the Chow, (1960) and Bai & Perron, (1998) tests as it does not require prior knowledge of the break, See (Castle et al., 2015a, 2012). The SIS method does not restrict the number or lengths of breaks and breaks at the start or end of observations (Castle et al., 2015; Pretis et al., 2018). Meanwhile, Doornik, (2009) recommended that Impulse Indicator Saturation (IIS) is a robust estimator in the presence of outliers compared to all other existing techniques. However, Johansen and Nielsen, (2009) describe and demonstrate a split-sample estimator for the indicator-saturated regression model as a one-step M-estimator that is iterated twice.IS method is a popular method for outlier and structural break detection that simultaneously detects the outlier and step impulses and underlying covariate in modeling (Doornik, 2009; Hendry et al., 2006; Johansen & Nielsen, 2009).

### 1.2. Problem statement and significance of study

The SIS and IIS methods for structural break and outlier detection exhibit several dummy regressors in the model, which is equal to the number of observations. SIS and IIS are feasible to estimate because of Autometrics; it can manage N candidate variables greater than T observations during model selection by extending and contracting multiple-path searches. The SIS and IIS method uses Autometrics to simultaneously detect and estimate the model, which is also an advantage for being used in conjunction with all other aspects of model selections. IIS and SIS are built-in functions of Autometrics (part of PcGive in Oxmetrics) and get package in R (Pretis et al., 2018). It automatically detects the break and estimates the model simultaneously (Castle et al., 2012; Doornik, 2009).

However, the fundamental of Autometrics is based on the concept of PcGets which automatically selects the model proposed based on the pre-specified level of significance (Hoover et al., 1999; Krolzig & Hendry, 2001). However, the concept of general to specific (gets) modeling serves as the foundation for this approach. They use their standard testing approaches to reduce the complexity of the dataset by removing statistically insignificant variables and examining the validity of reduced modelat each stage to ensure the congruence of the chosen model. It begins with a general unrestricted model that captures all the core attributes of the underlying model (theoretical or empirical).. Campos et al., (2003) used Monte Carlo simulation to ensure the consistency of the PcGets approach, and they studied the PcGets probability of retrieving the data-generating process (DGP) and achieved reliable findings.

Autometrics is the third generation of the PcGets method that was put forth by Doornik (2009). Autometrics use a tree-path search to locate and eliminate statistically unimportant factors; if the significant covariate is eliminated by chance in that case, the algorithm continues to function and does not get stuck in a single path comprising other covariates as proxies (like in stepwise regression). Autometrics is effective even if the model's covariates are greater than the number of observations (Castle et al., 2012).

On the contrary, other than Autometrics, a considerable amount of literature exists on Machine Learning techniques that can handle saturated models efficiently, called the regularization technique. The regularization techniques govern the loss function via an additional parameter called tuning parameter. Tuning parameters regularize the loss function in terms of bias and variation. Among the literature of regularization technique, the Least Absolute Subset Selection Operator(LASSO) is one of the well-known technique introduced by Tibshirani, (1996). Primarily, among other regularization techniques, the LASSO estimate selects more irrelevant regressor and does not possess oracle property compared to other methods like Smoothly Clipped Absolute Deviation (SCAD), Adaptive LASSO (AdaLASSO), and Minimax Concave Penalty (MCP) (Fan & Li, 2001; Zhang, 2010; Zou, 2006). SCAD and MCP satisfy the oracle properties asymptotically, opt for the correct subgroup of variables with nonzero coefficients, and have an optimum estimation bias (Fan & Li, 2001; Zhang, 2010). However, AdaLASSO can possess the oracle property if the weights are data-dependent and carefully chosen.

In a recent decay, the use of sparse modeling has grown widely in time series analysis as it can efficiently handle big macroeconomic data sets and substitute the factor models (Bai & Ng, 2002; De Mol et al., 2008; Hua, 2011; J. Li, 2012; J. Li & Chen, 2014; Marsilli, 2014; Nicholson et al., 2017). For the time being, the Adaptive LASSO (AdaLASSO) consistently chooses the important covariates as the number of observations grows (model selection consistency) even when the errors are non-Gaussian and conditionally heteroscedasticity (Medeiros & Mendes, 2012). The theoretical and empirical efficiency of AdaLASSO is demonstrated by Audrino & Camponovo, (2013) as it asymptotically selects covariates with finite-sample in time series regression models. Covariates and their lag selections are challenging in time series modeling, mainly when a mixture of serial correlation exist (Song & Bickel, 2011). Konzen & Ziegelmann, (2016) purpose weighted lag adaptive LASSO(WLagAdaLASSO) for covariate and its lag selection and compare it efficiency with LASSO and AdaLASSO under different scenarios. Their result indicates that WLAdaLASSO outperforms LASSO and AdaLASSO in forecasting and covariate selection, even in a greater linear dependency between predictors with many candidate lags (Konzen &

Ziegelmann, 2016). Although, Uematsu & Tanaka, (2019) uses folded concave penalties for ultra-high-dimensional time series forecasting and covariate selection. They verify the oracle inequalities of folded concave penalties (SCAD and MCP) for macroeconomic time series under appropriate conditions with the theoretical and empirical contribution.

In the meantime, very few studies exist that utilize the classical technique (Autometrics) in macroeconomic forecasting and covariate selection (Castle et al., 2013, 2021). However, in cross-sectional modeling, Epprecht et al., (2021) compare the LASSO and AdaLASSO estimates with classical techniques (Autometrics) in forecasting and covariate selection in the static model. Their result indicates that LASSO and AdaLASSO estimates outperform Autometrics in prediction. Their study finding indicates that LASSO, AdaLASSO outperform in forecasting than Autometrics, and for covariate selection, LASSO and AdaLASSO select more irrelevant regressors than Autometrics.

Castle et al., (2015) and Pretis et al., (2017) used the LASSO estimate for break detection via SIS method and compared its efficiency in gauge and potency with Autometrics. Their analysis indicates that the LASSO estimate works well for a single-step shift due to the forward selection approach, and selection over multiple-step functions fails to detect shifts once multiple breaks occur. However, Castle et al. (2021) utilize IIS and SIS techniques as a robust model discovery that aims to simultaneously handle all types of misspecifications on the cross-sectional dataset (Boston house prices) and see its improvement on prior empirical studies. Although, Autometrics possesses several flaws as the model selection based on the prespecified "significance level", including pre-designated test statistics, thrives in different models as the level of significance changes (Doornik, 2009). Autometrics retain some irrelevant step dummies/variables even though they may be insignificant with a nominal significance level 0.05, whereas, with a 1% or 0.1%

significance level, it only includes variables/dummy indicators that are highly significant and omit relevant variable/dummy indicators in the final model (Castle et al., 2012). The choice of the significance level is imperative as it controls the trade-off between the irrelevant and relevant dummy indicators/variables (Castle et al., 2015a, 2015b; Pretis et al., 2018).

On the contrary, regularization techniques dependent on the shrinkage penalty control the tradeoff between bias and variance. It is well known that the LASSO estimate does not fulfill specific statistical properties like oracle and unbiased and performs poorly for multiple break detection (Castle et al., 2015 & Pretis et al., 2017). Meanwhile, other regularization techniques, such as SCAD and MCP estimates, meet sparsity, continuity, and oracle conditions and possess unbiased estimates. Fan & Li (2001) and Zhang (2010) illustrate the efficacy of SCAD and MCP over LASSO for covariate selection.

However the efficiency of regularization techniques such as SCAD, MCP, and AdaLASSO in structural break detection and model selection does not exist in the core of the existing literature. For this purpose, we employ SCAD, MCP, and Adaptive LASSO for structural break and outlier detection and compare their performance with the classical approach (Autometrics). However, another goal of this study is to compare WLAdaLASSO, SCAD, and MCP with classical techniques (Autometrics) in dynamic time series modeling. We haven't come across any study that compared the efficiency of regularization techniques. We assess the performance of regularization techniques in terms of Gauge ('Size'), Potency ('Power'), Root Mean Square Error (RMSE), and Mean Absolute Error (MAE) under different Data Generating Processes (DGP) in the simulation studies.

### 1.3. Objective of the study

The main goal of this study is to evaluate the efficacy of regularization techniques versus Autometrics for model selection and structural break identification. The specific goals are:

- The first objective of this study is to compare the SCAD, MCP, and AdaLASSO estimates with Autometrics and LASSO for break detection in the univariate case and breaks with different shift magnitudes. We used fixed and cross-validation tuning parameters for break detection.
- The second objective of this study is to analyze the efficiency of SCAD, MCP, AdaLASSO, and LASSO in the univariate autoregressive series with outliers and multivariate static model with different outlying observations. We consider three scenarios 5%, 10%, and 20% obtained by assuming $\varepsilon_i \sim (0, \sigma + 4)$ $and$ $\varepsilon_i \sim (0, \sigma + 6)$..
- The third objective of this study is to compare the WLAdaLASSO with SCAD, MCP, and Autometrics for the time series dynamic model. We considered different scenarios with various autocorrelation coefficients (0.1, 0.5, and 0.8) of regressors and T sample sizes (50, 100, and 500).

- Finally, we apply the SIS technique with SCAD, MCP, and AdaLASSO estimates to detect the break-in GDP growth and GDP deflator of Pakistan to evaluate our simulation experiment with real data application. For multivariate analysis of outlier detection data; we use COVID-19 cross-sectional data collected from July 2021 and 30 September 2021 in Isolation. For dynamic time series modeling, we use macroeconomic indicators of the balance of trade in the case of Pakistan.

The efficiency of these techniques is assessed with gauge, potency, and in-sample/out-of-sample Root Mean Square Error (RMSE) in the simulation experiment. Meanwhile, in DGP, we intake orthogonal cases for this purpose, we use some well-known orthogonal techniques of regularization like LASSO (Tibshirani, 1996), Adaptive LASSO (Zou, 2006), Smoothly Clipped Absolute Deviation (SCAD) (Fan & Li, 2001), and Minimax Concave Penalty (MCP) (Zhang, 2010).

**1.5. Organization of thesis**

The remaining part of the thesis is organized as follows.

Section 2 is based on the literature overview of structural break detection and model selection techniques.

Section 3 discusses the considered methods for structural break detection and model selection techniques, including orthogonal regularization techniques and the classical approach (Autometrics). We also discuss the evaluation method for assessing considered techniques and selecting tuning parameters.

Section 4 illustrates the data generating process and result of simulation experiment for univariate and multivariate breaks and outlier detections. the results of the simulation experiment are subsections according to the study's objective.

Section 5, based on the real data analysis in this section, each of the real data analyses is linked to the objective of the study and simulation experiment.

Section 6 is based on the in-depth and summarized discussion of real data analysis and simulation experiments. This section also discusses the research's limits and future directions.

# Chapter 2

# Literature Review

## 2.1. Introduction

This part of study covers the literature review for structural break detection and Model selection techniques. We divided this chapter into three sections; the first part of the literature review covers the test for structural break and outlier detection in time series. The second part of the literature review contains model selection techniques, which are further subdivided into two parts: conventional model selection techniques and machine learning techniques for high-dimensional analysis.

## 2.2. Methods for Structure Break and Outlier Detection in Time Series

The structural breaks are typical in economic time series interactions and ignoring it can be dangerous. The structural break is usually found in macroeconomic data that follow an extensive time-series dataset influenced by various economic factors (Muthuramu & Uma Maheswari, 2019). In general, structural breaks occur due to changes or sudden shifts in the socio-economic structure, political instabilities, and pandemics across the globe. Economic links can be misinterpreted, estimates can be wrong, and policy suggestions might be misleading or worse. The new tools created in recent years are beneficial aids in specifying, analysing, and evaluating econometric models. The techniques for estimating structural break are the core concerns of econometrics methods. In the last few decades, tremendous development in testing structural breaks has existed in literature. The econometrics of structural change seeks systematic ways to identify structural breaks. The following three developments are among the most significant additions to this literature over the last decades: 1) Tests for known timing in a structural break; 2) Structural Breaks Tests for Unknown and Multiple Breaks; 3) Indicator Saturation Method for Structural Break and Outlier Detection.

### 2.2.1. Test of Known Structural Breaks

The traditional credit for the structural change test usually goes to Chow (1960). His well-known testing method divides the sample into two subperiods, estimates the parameters for each subperiod, and then uses a classic F statistic to evaluate the equivalence of the two sets of parameters. This test has been widely used for years and expanded to include the most relevant econometric models. However, the Chow test has a significant limitation: the break date must be known in advance and was supposed to be a test for a single break. There are two options available to a researcher: choose an arbitrary candidate break date or a break date based on a well-known data attribute. The break date may be omitted in the first scenario, making the Chow test useless. As the candidate break date in the second scenario is endogenous—it is correlated with the data—the Chow test may be deceptive because it may falsely show the existence of a break when none exists.

Furthermore, since the outcomes can be susceptible to these arbitrary decisions, it is simple for researchers to come to different conclusions. Following the linear regression by k number of observations and vectors of n1 and n2 segments of observations were performed using the conventional F test. The Chow test has an extensive history; it is produced in various procedures, but the most commonly used procedures are those mentioned by Dufour, (1982).

The chi-square critical values are inappropriate if the break date is unknown at the beginning. What alternative crucial values are appropriate? This question remained unsolved for many years, and Quandt, (1960) had no real-world use. Although Quandt, (1960) test was used for an unknown structural break. However, modifications within the error variance, he examined the constant-coefficient versus the alternative. The issue was resolved concurrently by many groups of authors in the early 1990s, with Andrews (1993) and Andrews and Ploberger (1994) providing the problem's most elegant and comprehensive solutions. These authors offer critical value tables, and Hansen (1997) offers a formula for calculating p-values.

Though, the empirical prove of the chow test's correctness under heteroskedasticity demonstrated by Toyoda, (1974) which occurs when one of the sample sizes is very big. However, the level of significance of the test affected because of sample size. The level of significance always increases when there is heteroscedasticity. The precision and proof of chow test using a Toyoda device confined by Schmidt & Sickles, (1977). They concluded that Toyoda's conclusion was incorrect somehow and that there are two sample sizes with distinct variances in each. The analysis of variance test for simultaneous equations proposed by Lo & Newey, (1985) and Park, (1991). The analysis of variance test was used by Andrews & Fair (1988) for broad nonlinear econometric models, the introduced Lagrange Multiplier test, Wald test, and Likelihood Ratio test. Their findings revealed a poor heteroskedasticity controlling condition. The broad nonlinear dynamic simultaneous equation models to investigate structural stability prediction used by Dufour et al, (1994). The study also considered a substantial subsample of data prior to the structural break; as a result, structural alterations in the second half remain unclear. A Chow test with different regimes covering less than k subsamples conducted by Dufour, (1982). The analysis of variance test to investigate broad nonlinear models used by Lo & Newey, (1985) conducted simultaneous equations of the chow test Andrews & Fair, (1988). The Quandt, (1960) test with basic linear regression models (with intercept changes alternatively intercept and slope changes) used by Kim & Siegmund, (1989). The generic nonlinear models were used with the predictive test by Cantrell et al., (1991) and Dufour et al., (1994). The p-values reported in great detail by Hansen, (2001). Overall, these tests statistics for structural break detection used to produce the subsequent formation: examine for identified breakpoints, test for unidentified breakpoints, test for numerous breakpoints that are not known. Usually, the defined techniques assume known breaks, identifying breakpoints based on exogenous occurrences/outcomes (for example, the Great Recession, Oil Shock, Liberalization, Global Financial Crisis, and Eurozone Crisis) or arbitrary dates makes sense in general.

### 2.2.2. Structural Breaks Tests for Unknown and Multiple Breaks

In the late 1970s, the literature on structural break evolved to detect parameter uncertainty or changes that happened during an unidentified period. It focuses on parameter uncertainty in dynamic models with trending predictors, co-integrated regressors, heteroskedasticity in error, and perhaps Unit root (Bai, 1994). It received much consideration in relations of theoretical and empirical confirmation in the arena of econometrics Andrews, (1993), Farley & Hinich, (1970), Hansen, (1992 and 2012), Jansen & Teräsvirta, (1996), Perron, (1989), Perron & Vogelsang, (1992), Perron, (2006), Ploberger et al., (1989), Ploberger & Krämer., (1990), and Quandt., (1960). The synoptic assessments of the single unidentified structural break were assembled in the chapter, which was added by Maddala & Kim., (1998), Stock & Watson., (2010), and Vilares., (1986). On a single unknown structural break, there exist three significant tests categorized by Vilares., (1986). Depending on the size of the model, the number of parameters, and other variables, these asymptotic critical values are significantly bigger than the corresponding chi-square critical values. Tests are created for multiple structural changes by Bai and Perron (1998). Starting with a single structural break test, their procedure is sequential. The sample is divided in half (depending on the break date estimate supplied in the next section) and the test is repeated on each subsample if the test rejects the null hypothesis that there is no structural break. This process continues until no further evidence can be found by the subsample tests of a break.

Conversely, in the second half of the 1970s,  Brown et al., (1974) presented the CUSUM test as a method for analyzing recursive residuals. Ploberger, (1983) proposed it for the first time in a linear regression model with lagged regressand variable and local power of CUSUM, Ploberger & Krämer, (1990 and 1992) expanded the CUSUM and CUSUM squared tests. They used a dynamic linear regression model to demonstrate the structural change and introduced the instability test relatively than the recursive residual on parameter estimation. Though their sample has just undergone a structural transformation. However, a recent structural change happened in their sample. They demonstrated that the CUSUM test had a flaw in that its regression coefficient was asymptotically negative. However, rather than constant coefficients, it does not cause heteroskedasticity in its disturbances. Following that, the CUSUM squared test revealed asymptotically similar results.  Krämer et al., (1988) and Ploberger & Krämer, (1990) expanded the Quant test for dynamic linear regression model. Ploberger et al., (1989) suggested that the CUSUM test showed local power against heteroskedasticity. They presented a fluctuation test for the power problem (regarding sequential parameter estimates relative than recursive residuals). CUSUM and CUSUM squared test were used by Westlund & Törnkvist, (1989) to assess the structural stability of the test statistics and used the Monte Carlo method. The test statistic parameters were estimated differently, and the Monte Carlo approach had the smallest chance of generalization. For various parameters, CUSUM and CUSUM squared test statistics were unknown. Another type of test introduced by Andrews, (1993) as the Sup F test . He evaluated a parameter uncertainty test and a previous structural alteration with an unknown break point. The study has nontrivial asymptotic local power versus all alternatives with non-constant parameters or unknown break points (structural breaks). Andrews, (1993) methodology was used to analyze nonlinear models since the Sup F test (with asymptotic critical values of 10%, 5%, and 1% significance levels) offered superior power attributes than the fluctuation in CUSUM test. The optimum analysis was used by Andrews et al., (1994) as soon as a nuisance parameter was existing under the alternative. They strictly regarded stationary series for this purpose, and an ideal test was generated using a weighted average power criterion. On the other hand, the structural breaks were known; one could use Wald test and Lagrange Multiplier test with no stochastic or deterministic trends. The Likelihood ratio-like test using on nonlinear models built on the Generalized Method of Moments (GMM) estimators presented by Andrews. He also gave an asymptotic critical value, which he called Sup F test (Maddala & Kim, 1998).

Macroeconomic time-series variables are frequently subjected to several unknown structural breaks. The detection of multiple unknown breaks has drawn attention in recent decades. The problems are estimating the break date and getting confidence intervals for the break date when treating the date of structural change, or the "break date," as an unknown parameter. The date that produces the highest value in the Chow test sequence is a clear candidate for a break date estimate (in our labor productivity example, May 1991). The Chow test is generated with the "homoskedastic" version of the covariance matrix in linear regressions, and it turns out that this is the only case where this is known to be a reasonable estimate.

In regression models, least squares are an appropriate method for estimating the parameters, including the break date. The sample is divided at each potential break date, the other parameters are estimated using ordinary least squares, and the total squared errors are computed and saved. The date that minimizes the full-sample sum of squared errors is known as the least squares break date estimate (equivalently, minimizes the residual variance). Jushan Bai has created a theory of least squares estimation in several papers, solo and by coauthors. The asymptotic distribution of the break date estimator is derived by Bai (1994, 1997a), who also demonstrates how to create confidence intervals for the break date. These confidence intervals indicate the degree of estimating accuracy, which is simple to calculate and particularly helpful in applications.

This methodology is expanded upon by Bai, Lumsdaine, and Stock (1998) to include several time series with simultaneous structural breaks. They demonstrate that increasing estimation precision by utilizing multiple time series. The simultaneous estimation of several break dates is covered by Bai and Perron (1998). Chong (1995) and Bai (1997b) demonstrate how to estimate several break dates sequentially. The essential finding is that the sum of squared errors (as a function of the break date) may have a local minimum close to each break date when there are many structural breaks. As a result, it is possible to employ the global minimum as a break date estimate while also cautiously considering the other local minima as potential candidates. Following the sample's division at the break date estimate, further analysis is performed on the subsamples. Bai (1997b) demonstrates how iterative modifications can yield significant gains by re-estimating break dates using improved samples.

Under the Bai-Perron class of tests, there are several alternative approaches for estimating multiple breaks Bai (1994 and 1997) and Bai & Perron (1998, 2003a, 2003b, 2003c, and 2006) have all made important contributions to the field. Sequential analysis, global maximizer, and hybrid versions are mutual components among the available process. Bai, (1994) investigated inference models with a simultaneous relatively than sequential structural break (consecutive estimation to each breakpoint). The goal was to find the factors of many breaks and estimate the number of breakpoints at various breaks and develop result estimators. They defined it as a partial structural change model with constant parameters. The goal was to find the best approach for estimating each breakpoint relatively than the exact position of the breaks. They focused on no structural changes vs arbitrary modifications under the estimated null hypothesis (L) versus (L+1) changes and employed the Sup Wald test for analytical purposes. Reducing the sum of squared residuals incorporates the linear regression model with many structural breaks. The occurrence of a structural break is decided by the estimators' attributes, the number of breaks, and the predicted break dates. In a pure structural change model, all coefficients are liable to change, and for additional information, see Bai & Perron, (1998). They utilized quarterly US ex-post real interest rate data from 1961 to 1986. They permitted up to five segments for empirical verification and detected two breaking dates (1972:3 and 1980:3) estimated under global minimization. It is, nonetheless, beneficial when dealing with linear regression models that include several structural breaks. It is not allowed for a convergence rate of sequential estimators but instead calculated the rate of breakpoint convergence.

The null hypothesis tests for structural breaks assume no structural differences vs an unspecified number of breaks provided an upper bound M (Bai & Perron 1998 and 2006). The first is regarded as a pair of maximal checks. Within this, there are two subgroups: an equal-weight variant (UDMax) and a test that discovers unique weights that produce equal marginal p-values throughout m (WDMax). However different autocorrelation conditions and error distributions, and heteroskedasticity in the explanatory variables used by Bai & Perron (2003b and 2006). The HAC approach supplied numerous restrictions in the overall framework to correct the distribution of error term and regressors across segments. Generally, Bai and Perron used diverse assumptions for autocorrelation, distribution of error, and regressors heteroskedasticity.

In current econometric practice, the Chow statistic has largely been supplanted by the Quandt (1960), Andrews (1993) and Andrews and Ploberger (1994) families of statistics. Stock and Watson's (1996) systematic application of the tests to 76 monthly time series using univariate and bivariate regressions is one example of a thorough application. More than half of their models disapprove stability at the 10% level. Ben-David and Papell (1998), who search for indications of "slowdowns" (a decline in the trend function) in the Summers-Heston GDP data from 74 nations, is another intriguing application. In 46 countries, they discover statistically significant evidence of a slowdown. The post break trend function is negative in 21 of these situations. McConnell and PerezQuiros (2000) are a final illustration that has received much recent attention. The stability of the volatility of US GDP growth rates is tested, and they discover convincing evidence of a considerable decline in volatility about 1984.

#### 2.2.3. Structural Break or Random Walk

Time series are frequently stated as consisting of a trend and a cycle. Prior to Nelson and Plosser's (1982) research, it was conventional knowledge that the trend was linear. Nelson and Plosser (1982) contested that presumption by presenting evidence that the trend might be described as a random walk for many commonly used aggregate macroeconomic time series. To put it another way, the trend would be shifted by random shocks rather than being a fixed trend to which the time series would retrace throughout the business cycle, and it would then hold at the new level until disturbed by another random shock. While various counter challenges were mounted in response to this result, Perron's(1989) was the most productive. According to Perron (1989), a sparse single structural break in a linear trend that is otherwise continuous could account for the trend's movement. This theory makes sense since a trend break results in serial correlation qualities that are comparable to those of a random walk. Perron (1989) demonstrated how to compare the random walk theory and the trend-break model. The desired broken trend specification is captured by estimating a linear autoregression enhanced with dummy interactions. An easy way to evaluate this is with a t-ratio statistic because the hypothesis of a random walk trend requires that the total of the autoregressive coefficients equals one (that is, a "unit root" in the autoregressive polynomial). Perron (1989) developed a distribution theory and crucial values even though the t-distribution ratios are non-normal. This test was used on the Nelson-Plosser (1982) macroeconomic time series by Perron (1989), who chose 1929 as the break date for the annual series and 1973 for the postwar quarterly series. For most of the series, he rejected the random walk model at the 5% level of significance, indicating that the series were stationary after taking structural changes in the trend into account.

The Perron (1989) concept has significantly and rightfully influenced empirical analysis and brought attention to the time series features of the trend. As we now realize, the main difference between a random walk and a trend break relates to how frequently the trend is permanently shocked. Such shocks happen regularly in a random walk process but seldom in a trend-break process (once or twice in a sample). Future research can look for more approaches to reduce the disparity between these models. Perron's (1989) idea and its variants have been applied to multiple applications. The main concern of Fernandez (1997) work is whether changes in the money supply, even after conditioning on lagged output, aid in forecasting output. According to past research, the outcomes depend on whether or not interest rates are factored into the regression and whether or not a time trend is incorporated to linearly detrend the dataset. Fernandez (1997) uses Perron's experiments to support his claim that a stationary process regarding a temporal trend with a single trend break adequately captures output. Fernandez then detrends output using the estimated broken trend function. Fernandez (1997) discovers that when the sample period is limited to data prior to 1985, this yields very robust results; but, when data collected after 1985 are included, he is unable to yield robust results.

Several other studies, including Christiano (1992), Zivot and Andrews (1992), Banerjee, Lumsdaine, and Stock (1992), as well as Perron and Vogelsang (1992), disagreed with Perron's (1989) & (1992) findings. According to these studies, it is improper to declare the break date to be known because it is implausible to think that the decision was taken independently of the facts. The break date that produces the highest t-ratio—the break date that produces the most evidence against the random walk hypothesis—is what these writers jointly propose as an appropriate strategy. The test statistics can have the same numerical value using this break date selection process because Perron's choices of 1929 and 1973 were quite wise. On the other hand, the test is built using a different technique and hence has a distinct sample distribution. It is more difficult to reject the null hypothesis of a random walk with the modified test since the critical values are substantially more significant. With the aid of this new explanation, the facts that supported the random walk hypothesis vanished.

However, the question arises that; Will larger data samples end the debate? In a follow-up study, Perron (1997) expanded the sample to include 1991:III and employed various techniques to choose the autoregressive lag order. Although he discovered marginally more convincing evidence against the random walk model, it was still inconclusive. The main issue is that the trend functions mentioned earlier in Perron (1989) do not make good extra-sample predictions. The trend function significantly overpredicts the 1970s and 1980s for the annual series. It underpredicts the quarterly series from 1987 to 2000. This fits the trend of a random walk. Lumsdaine and Papell (1997) discover that the evidence against the random walk is more robust when they permit two break dates instead of one. However, the distinction between the trend-break and random walk models is also diminished by the requirement for two structural breaks.

Papell, Murray, and Ghiblawi (2000) provide still another example The hysteresis in unemployment rates in 16 OECD nations is of concern to these authors. It is strongly related to the idea that trend unemployment can be compared to a random walk since hysteresis holds that a temporary shift in unemployment can have long-lasting repercussions. The Perron-Vogelsang (1992) tests enable the authors to reject the random walk hypothesis in favor of a one-off break in the time trend for ten of the 16 countries. This result raises an entirely distinct economic theory of hysteresis.

### 2.2.4. Indicator Saturation Method for Structural Break and Outlier Detection

Bai-Perron tests possess the subsequent limitations: a time series macroeconomic variable possess multiple unknown breaks. However, Bai-Perron class of tests based on the per-specified number of breaks and trimming parameters based by sensitivity analyses. In conclusion, a little adjustment in the series length or the use of a real price series rather than a nominal one impacts the results. As a result, sensitivity analysis becomes critical in this situation. Other than these limitations all of the above structural break test Andrews, (1993), Farley & Hinich, (1970), Hansen, (1992 and 2012), Jansen & Teräsvirta, (1996), Perron, (1989 and 2006), Perron & Vogelsang, (1992), Ploberger et al., (1989), Ploberger & Krämer, (1990), and Quandt, (1960) based on the assumption of pre-specified models. However, if the model is far from 'accurate specifications', such methods can be ineffective at break detection, and in what way to 'repair' it is constantly unclear (Castle & Hendry, 2019a).

In contrast, Indicator Saturation (IS) methods were primarily designed to detect unknown numbers of outliers with unknown magnitudes at uncertain points in the sample, including the beginning and end of observations (Castle et al., 2015; Hendry et al., 2006, Castle et al., 2015a, 2012; Santos et al., 2008). The Chow test is a well-known technique for a single and known structural break, and in other techniques, information of break timing, kind, or known shift magnitude is usually required. However, other various form of structural break test in a least-squares linear model investigated by Bai & Perron, (1998). They suggest specific structural break tests without trending regressors, and a selection technique established on a series of tests to estimate the number of breaks consistently. Bai-Perron tests for multiple breaks is subject to certain restrictions: the technique is not valid for trending series and is limited to a pre-defined fixed number of breaks (Bai & Perron., 2006). However, the Step Indicator Saturation (SIS) method is a modified version of IS techniques for multiple break detection. SIS method already takes over the Chow, (1960) and Bai & Perron, (1998 and 2006) tests as it does not require prior knowledge of the break, See (Castle et al., 2015a, 2012). The SIS method does not exhibit any restriction on the number or lengths of breaks and breaks at the start or end of observations (Castle et al., 2015; Pretis et al., 2018).

IIS is a popular method for outlier detection that simultaneously detects the outlier indicator and underlying covariate modeling (Doornik, 2009; Hendry et al., 2006; Johansen & Nielsen, 2009). However, the IIS method can be used as a robust estimator in a presences of outliers, whereas Johansen and Nielsen, (2009) describe and demonstrate a split-sample estimator for the indicator-saturated regression model as a one-step M-estimator that is iterated twice (Doornik, 2009). The robust least squares and IIS are more efficient than least trimmed squares in the presences of outliers (Doornik, 2009). When the regressors are fixed, and only outliers occur in the dependent variable's data occur, M estimation works effectively. However, Robust regression techniques are used significantly in the literature of outlier presences. Langford and Lewis, (1998) defined an outlier as an observation that appears inconsistent with the rest of the data. Such influential points are frequently concealed from the user since they do not always appear in the standard least-squares residual graph. the OLS residuals are ineffective in finding outliers in small and big sample sizes (Zaman et al., 2001). At the same time, a well-known technique is based on Huber's M-estimators, which offer robustness in location parameters. Regrettably, generalizations of regression models fail to achieve robustness Rousseeuw, (1984) illustrates that regression M-estimators likewise have a 0% breakdown value. The generalization of MM-estimators likewise fails to attain large breakdown values. A direct method to robust regression is to use Least Trimmed Squares (LTS) analysis in huge residuals. The LTS analysis discards outlying observations and then can run a standard OLS regression, proposed in Rousseeuw, (1984). However, removing too many data points in the case of too many outlier observations runs the risk of the final regression not reflecting the relationship that the econometrician wants to assess (Zaman et al., 2001).

Meanwhile, the SIS and IIS method exhibit several dummy regressors in the model, equal to the number of observations. IS method is being feasible to estimate because of Autometrics, as it can manage more N candidate variables than T observations during model selection by extending and contracting multiple-path searches. The choice of the significance level is the trade-off between the irrelevant and relevant dummy indicators (Castle et al., 2015b; Pretis et al., 2018). As the SIS and IIS methods possess huge dummy regressors, the conventional OLS method fails to estimate the thriving model; hence, we discuss different model selection techniques in detail in the below section.

### 2.3. Model selection techniques

Since the primal in time series analysis, modeling and forecasting have been the center of attraction. The accuracy of the model in time series analysis is always unknown. Only one in a million models can be accurate; ‘‘Essentially, all models are wrong, but some are useful’’(Box, 1979). However, the massive availability of data in the current era leads us to a new phase of time series analysis for model selection and forecasting. Including many financial and economic covariates in the time series model for superior prediction may yield considerable benefits. However, parsimonious models in time series analysis perform superior in forecasting. Failure to decrease dimensionality may lead to poor performance due to cumulative estimation losses from redundant or insignificant variables. There are two primary schools of thought: high-dimensional analysis in machine learning and the classical model selection approach. Since the last two decades, many model selection techniques existed in high-dimension machine learning techniques. However, slight enhancement exists in the literature of classical statistical techniques. This section elaborates on each of these techniques in further detail.

### 2.3.1. Classical Approach (General-to-Specific)

Autometrics is the third generation modified version of general to specific modeling (Hendry & Krolzig, 2004; Santos et al., 2008). The set of all variables is divided into two: those that are currently selected (the candidate set) and the rest (the excluded set) (Doornik, 2009). The set of variables that are not currently selected is divided into blocks. Then two steps are alternated in Expansion and Reduction Steps. An alternative to split-sample and cross-blocks algorithms where a more advanced search has been conducted when the variables are in huge amount is the block-search algorithm (Doornik, 2009).

Autometrics with impulse saturation outperforms location-scale-trend, location-scale, and stationary autoregression model selection with occurrence of multiple breaks (Castle et al., 2012). Whereas the performance of stepwise regression is poor in all cases: frequently, it does not possess power to detect change points. Though, it is harder to detect change points when the series possess trend in the Data Generating Process (DGP) at a single change point in impulse indicator (Autometrics), so far this is not the case where there are multiple change points. Thus, compared to a single change point, multiple change points are more challenging to detect in a DGP without a trend but easier in the DGP with a trend (Castle et al., 2012). When there is a change point in a stationary autoregressive with an intercept, Autometrics does not perform well, whereas with a unit-root model and a limited outlier, a stationary autoregressive with a shift does well in detecting breaks. However, estimating a Generalized Unrestricted Model (GUM) with the dependent nonstationary variable and no constant, Autometrics does not perform well (Hendry et al., 2013). Autometrics model selection with somewhat tight significance thresholds and bias correction is an effective strategy that permits many breaks to be addressed (Doornik, 2009). Despite the fact that the technique involves both expanding and contracting searches (due to the fact that there are more regressors than observations), impulse and level saturation allow dummies to identify and 'model' multiple breaks (Castle et al., 2012). Setting the nominal significance level at $\alpha \leq 1/N$ approaches to 0 as T approaches $\infty$, out of N candidates variables/dummies, on average, one extraneous variable/dummy will be preserved as potentially significant (Doornik, 2009; Doornik & Hendry, 2015). As a result, starting with GUM, it is not easy to eliminate practically all irrelevant variables. The alternative methodology for the locations, magnitudes, durations and signs of location shifts when they are unknown is introduced by Hendry et al., (2013) known as Step Indicator Saturation. Step indicators are the cumulation of impulse indicators up to each next observation. The Impulse Indicator Saturation (IIS) test possesses low power for long breaks, whereas Step Indicator Saturation (SIS) intakes important application for multiple location shifts at the forecast origin to test supper heterogeneity pernicious. While the shift's location and magnitude are unknown, the step-indicator saturation technique seems reasonable because the technique has the correct null retention frequency in persistent conditional models with a nominal level of significance size of $\alpha$ (Castle et al., 2015b). Although the derivations and Monte Carlo simulations were done just for basic static equations and specified location shifts, the ideas appear to be generic and should apply to dynamic equations (albeit with estimated null-rejection frequencies) and conditional systems.

To detect the structural break in the model selection framework (Castle et al., 2015b) proposed choosing significant step indicators among the set of saturated model with the union of all candidate step dummy variables. Split half and multipath block search algorithms are used to extract the null retention frequency, and approximate non-centrality of a selection test is derived. The study validated the accuracy of nominal significance levels under the null and demonstrated retentions when location changes occurred, enhancing the non-null retention frequency when compared to the related Impulse-Indicator Saturation (IIS)-based approach and the LASSO (Castle et al., 2015b).

When the variables are non-orthogonal, it affects the speed of the algorithm in Autometrics (tree search algorithm) however, a tree search is infeasible for large N, whereas 1-cut path search algorithm is not for non-orthogonal variable (Doornik, 2009; Doornik & Hendry, 2015). When numerous variables must be combined for them to be meaningful, 1-step forward searches fail. 1-step forward searches across N variables, on the other hand, just involves computing N correlations and then adding variables until the next highest correlated variable becomes negligible when added, making them relatively quick even for large N. Even though 1-cut is not suitable for non-orthogonal data, there is little loss from utilizing the path-search technique in Autometrics (Castle et al., 2012).

Outlier detection is a rapidly developing procedure in the healthcare and medical data industries, and it is a significant source of concern. Hauskrecht et al., (2016) study data-driven outlier-based surveillance and alerting system that uses data from former patient cases Wilson et al., (2017) used the outliers identification method for hypoglycemia safety in patients, calculating a facility outlier value within a year, comparator group, and AIC threshold while considering at-risk population proportions. Jyothi et al., (2020) used Outlier detection in healthcare data, a key source of concern for health insurers. The development of a Supervised Outlier Detection Approach in Healthcare Claims (SODAC) and carried out in two parts. Noma et al., (2020) offer optimal influence measures for network meta-analysis models with missing outcomes and appropriate degree of freedom adjustments. The real data application of the IIS method in health care and medicine with outliers for cross-sectional analysis does not exist in the current literature (Hauskrecht et al., 2016; Jenkinson et al., 2020; Jyothi et al., 2020; Noma et al., 2020; Sakurai et al., 2019; Verbanck et al., 2018; Wilson et al., 2017).

### 2.3.2. Regularization Techniques

A considerable amount of work exists in the literature regarding dimensionality reduction of variables, and different reduction techniques have been proposed in the literature. For example, Ridge regression is like an ordinary linear regression. However, it shrinks the estimated coefficient toward zero (Hoerl & Kennard, 1970), Least Absolute Shrinkage and Selecting Operator (LASSO) based on the $l_1$ norm, it shrinks some of the coefficients exactly equal to zero and introduces substantial bias, but does not possess oracle property (Tibshirani, 1996). In contrast, Smoothly Clipped Absolute Deviation (SCAD) (Fan & Li, 2001), and Minimax Concave Penalty (MCP) (Zhang, 2010) possess oracle property and reduce substantial bias. All of the techniques assumed that the variables are orthogonal, whereas, for the non-orthogonal situation, different techniques exist in the literature.

Typical microarray data consists of a thousand predictor variables and fewer observations (usually less than 100). In genomics, it is considered that genes are performing as a group. The LASSO does not perform well because it selects a group of a variable in a final model with neglecting pairwise correlation (Zou & Hastie, 2003). The LASSO is not a perfect method when p >> n in the grouped variable situation because it can only choose n variables out of p candidates (Efron et al., 2004). Zou & Hastie, (2003)presented the Elastic Net regularization technique to handle such type of problem. The Elastic Net (EN) is Similar to the LASSO while simultaneously selecting the variable and continuously shrinking the coefficients of the correlated group variable equal zero. The EN can select all p variables if required and inclines to take the correlated variables as a group. This grouping selection now makes a less parsimonious model, as more coefficients are required to represent the additional variables (Bondell & Reich, 2008).

Octagonal Shrinkage and Clustering Algorithm for Regression (OSCAR) purposed by Bondell & Reich, (2008) for non-orthogonal high-dimension regressors. The methodology refers to non-complexity to final data and subgrouping the correlated predictors. This is precise equality of the variable coefficients allowing for a sparse representation in terms of the resulting complexity of the model. The number of unique non-zero coefficients in OSCAR formulation encourages a sparse solution. Consequently, the variable collected via shrinking coefficient equal to zero; the OSCAR instantaneously accomplishes a supervised clustering assignment by yielding a particular coefficient to regulate a cluster of variables that are pooled to have a particular effect on the response (Bondell & Reich, 2008). Before OSCAR (Bondell & Reich, 2008) and Elastic Net (Zou & Hastie, 2003) all of these studies never explicitly addressed the problem of correlation structure among the variables into account when the group of the variable is linearly dependent (Bühlmann et al., 2013).

Another study of the correlated variable in regression and clustering and sparse estimation has been introduced by (Bühlmann et al., 2013). The author primarily has proposed a canonical correlation for clustering variables and agglomerative hierarchical clustering, as the variables are linearly dependent. In the first step, an agglomerative bottom-up clustering algorithm is used, based on canonical correlation, and this ends with an optimal, statistically significant solution. The canonical correlation-based clustering improves the compatibility constant for the cluster group LASSO and also address bias and deduction issues: one satisfactory situation is for (nearly) uncorrelated clusters with possibly many vigorous variables in a cluster; the bias due to working with cluster representatives is small if the inside a group correlation is high, and recognition is moral if the regression coefficients inside a group do not withdraw (Bühlmann et al., 2013).

On the other hand, the traditional time series modeling for covariates and lag selection in Autoregressive Distributed Lag (ARDL) modeling uses Akaike Information Criteria (AIC) and Bayesian Information Criteria (BIC) (Pesaran et al., 2001; Pesaran & Smith, 1995). This technique is limited to the number of covariates, and their lag must not be greater than the number of observations. The traditional Ordinary Least Square method fails to estimate the forthcoming models with huge regressors and limited observations due to inadequate degrees of freedom. Several statistical techniques exist in the literature for model selection and forecasting when covariates and their lags are more than the number of observations. Meanwhile, classical approach (Autometrics, general-to-specific) and regularization techniques (Machine Learning) are frequently used in time series modeling when covariates exceed the number of observations. Besides these techniques, complex network theories provide an efficient and reliable solution for handling time-series issues. In recent years the complex network has been extensively used in socio-economic phenomena (Cui et al., 2021; Hu et al., 2020; Qiao et al., 2021). However, this study aimed to identify the true covariate and evaluate the model's forecasting performance, and we only concentrated on regularization techniques and the classical approach.

### 2.4. Significance of study and limitations

The prevailing test for structural break possesses the limitations of a pre-specified model, initially based on single and known/unknown break Quandt, (1960), Farley & Hinich, (1970), Ploberger et al., (1989), Ploberger & Krämer, (1990), Perron, (1989), Perron & Vogelsang, (1992), Andrews, (1993), Perron, (2006) ,Hansen, (1992 and 2012) and Jansen & Teräsvirta, (1996). However, Castle & Hendry, (2019) argues that if the model is far from 'accurate specifications', such methods can be ineffective at break detection, and in what way to 'repair' it is constantly unclear. Aside from these barriers, The number and position of structural breaks are indeed very subjective to the type of Bai Perron test employed, as well as assumptions about the amount of breaks in tests based on known breaks and trimming parameters as recommended by sensitivity analysis (Bai & Perron, 1998, 2003a, 2003b, 2003c, 2006). However, the SIS technique does not exhibit any restriction on length, magnitude, number, and timing of break (Castle et al., 2015b). SIS takes over the Chow, (1960) and Bai & Perron, (1998) tests as it does not require prior knowledge of the break, See (Castle et al., 2015a, 2012). The SIS method does not exhibit any restriction on the number or lengths of breaks and breaks at the start or end of observations (Castle et al., 2015; Pretis et al., 2018).

However, SIS techniques estimated via Autometrics possess certain limitations as the choice of significance level determines the final selected models. With a tight significance level, the final model omits relevant variable/dummy indicator, whereas, with a significance level equal to 5%, the final model retains irrelevant variable/dummy indicator. On the other hand, high-dimension (Regularization technique) machine learning techniques provide promising results in the case of a saturated model. However, the power of SCAD, MCP, and AdaLASSO is unrevealed for a structural break and outlier detection. Meanwhile, Castle et al., (2015b) compares the Autometrics with LASSO for multiple and single break detection; the result indicates that LASSO provides poor results in multiple shifts due to the forward selection method. It is evident that among high-dimensional techniques, LASSO lacks oracle property and produces biased estimates; however, SCAD and MCP are unbiased estimates and possess oracle properties (Fan & Li, 2001; Zhang, 2010). The power of the regularization technique, particularly SCAD and MCP, is not identified when applied on IIS and SIS for the structural break for single/multiple shift and outlier detection in the existing literature. This study examines the power of regularization techniques specifically SCAD and MCP for structural break and outlier detection and compares it with Autometrics in terms of gauge, potency, RMSE, MAE. In the meantime, we also compare Autometrics with the WLAdaLASSO estimate, which efficiently handles the time series dynamic modeling in forecasting and covariates/lags selection even with higher linear dependence in the predictor variable. However, the empirical comparison of Autometrics, SCAD, and MCP doesn't exist in the current literature. We assess the efficiency of these techniques in gauge, potency, RMSE, and MAE.

The study has a few limitations, considering only linear models for dynamic time series analysis. However, the study is limited to orthogonal covariates selection techniques such as LASSO, AdaLASSO, SCAD, and MCP for outlier detection via IIS method. The future study can be developed to examine the performance of modern statistical and machine learning methods combined with the IIS approach in panel data.

# Chapter 3

# Methods of Structural Break Detection and Model Selection Techniques

## 3.1. Introduction

The two broad spectrums of model selection techniques exist in literature; Regularization techniques and classical approach (Autometrics or general-to-specific modeling), whenever P regressors are greater than N number of observations. The classical approach (Autometrics, general-to-specific) starts with a fully saturated model and uses a backward elimination with the multi-path search process, and the selection of the model mainly depends on the predefined significance level. However, the regularization technique applies the sparsity on the p-dimensional parameter vector, which forces many of its components to be zero. This technique combats the issues posed by high dimensionality. We describe each of these techniques in more detail, but we only consider orthogonal regularization techniques in this study. The first section based on introduction of the indicator saturation method for break and outlier detection, and then we profoundly illustrate model selection techniques in the second section.

## 3.2. Indicator Saturation Method for Structural Break and Outlier Detection

In this study, we only considered IIS and SIS methods for outlier and step break detection. SIS is the sum of impulse indicators up to each following observation. Step indicators take whole-sample vectors, the system of $l_1' = (1,1,1, \dots \dots ., 1)$, $l_2' = (0,1,1, \dots \dots .., 1)$, and $l_n' = (0,0,0,0, \dots \dots, 1)$, $l_1'$ is dummy intercept (Pretis et al., 2018). However, IIS method I is a diagonal identity matrix of each corresponding observation in the model, which is illustrated as $I_1' = (1,0,0, \dots \dots .., 0)$, $I_2' = (0,1,0,0, \dots \dots, 0)$, and $I_i' = (0,0,0, \dots \dots ., 1)$. In this study, we consider univariate time series analysis by including a set of SIS dummy indicators, whereas as for IIS, we use multivariate analysis. This model is identified as the Dummy Saturation model. Attempting to estimate the

Generalize Unrestricted Model (GUM) is not feasible because there are more variables than a sample size. Autometrics (based on general-to-specific modeling) is used to detect these breaks and estimate the model simultaneously. In the general-to-specific methodology, each observation would have one dummy variable, and additional exogenous variables can be considered that could affect the dependent variable. Let us assume a univariate time series model, which includes m shift in data series, illustrated in below equation.

$$y_t = \sum_{m=1}^{T} \gamma_m I_m + \varepsilon_t \tag{3.1}$$

The IIS and SIS method assumes a generalized model, as a dummy indicator is introduced in the model that correspondence of each observation. The number of dummy indicators equal to the number of t observation in equation 3.2.

$$Y_{it} = \alpha + \sum_{i=1}^{T} \gamma_{it} I_{it} + \varepsilon_{it} \tag{3.2}$$

$$\varepsilon_t \sim IIN(0, \sigma^2), \qquad t = 1, 2, \dots \dots \dots., T$$

Whereas $I_{it}$ for SIS in the above equation can be represented as

$$I_{it} = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 1 & \ddots & 0 \\ \vdots & \vdots & \vdots & 1 & \vdots \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

Whereas IIS can be represented as

$$I_{it} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & 0 \\ \vdots & \vdots & \vdots & 1 & \vdots \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

For SIS method, the $I_{it}$ is a matrix of step dummies introduced to observe $y_t$ equal to ones and zero for all other observations. The first column in the matrix represents the dummy for intercept; hence we do not include the intercept dummy while estimating the procedure. Where $\varepsilon_t$ is independently and identically distributed (IID) with mean zero and variance $\sigma^2$ here m is the change point subscript in the above model. $y_t$ is regressed on a full set of saturated dummies under the null hypothesis of no shifts with nominal or 1-cut selection (Castle et al., 2012).

### 3.3. Model Selection Methods

#### 3.3.1. Autometrics

Autometrics, a third-generation algorithm created on similar concepts of PcGets. Hoover et al., (1999) proposed the general-to-specific model selection technique that aggregates many elements of the "Hendry" methodologies and "London School of Economics (LSE)". Doornik, (2009) proposed PcGets is a second-generation method extended by Krolzig & Hendry, (2001), prolonging and enlightening Hoover and Perez's algorithm; (Hendry & Krolzig, 2004; Krolzig & Hendry, 2001). The concept of general to specific (gets) modeling is the cornerstone of the Autometrics approach.

- Initially, the GUM includes the overall covariates and estimates it by the OLS method by expelling statistically irrelevant covariates; the reliability of the reduced model is confirmed at each stage to prove the congruence with diagnostic tests.
- Autometrics uses a tree path search with multi-step simplifications along numerous paths. Final models are calculated using a tree-path search and confirmed using diagnostic tests; if the coefficient estimates are statistically insignificant, the model is discarded. When many terminal models are identified, Autometrics re-tests their union. A new GUM is created when the 'surviving' terminal models are combined, allowing for one more tree-

path search repeat. The entire exploration process is repeated, with the terminal models and their combination' being examined once again. If many models pass the encompassing tests, the final choice is based on pre-determined information criteria.

Diagnostic tests are used to double-check the simplified models, while comprehensive tests resolve numerous terminal models. For diagnostic tests, Autometrics uses Jarque & Bera, (1980) residual normality test, Breusch & Pagan, (1980), and Godfrey, (1978) second-order residual autocorrelation, autocorrelated conditional heteroscedasticity (ARCH) to second-order (Engle, 1982), and in-sample stability (Chow, 1960). Autometrics is a partially black box (Epprecht et al., 2021). However, it allows the user to choose between "nominal significance level" and "1-cut and tight significance level" when establishing modeling approaches. The multi-path approach avoids path dependency by using a tree structure and a similar stepwise backward elimination, a built-in function of the gets package in R environments (Pretis et al., 2018).

### 3.3.2. Regularization Techniques

Regularization techniques handle saturated models with irrelevant regressors even if regressors are more than the number of observations and shrink the irrelevant coefficient equal to zero with some bias, as like LASSO (Tibshirani, 1996). In this work, we opt LASSO, Adaptive LASSO, SCAD, and MCP for structural break detection and compare it with Autometrics. Based on the $L_1$ norm, the LASSO estimates shrink some coefficients precisely equal to zero and introduce substantial bias, but unfortunately, it does not possess oracle property. In contrast, Adaptive LASSO (Zou, 2006), SCAD (Fan & Li, 2001), MCP possess oracle property and reduce substantial bias compared to LASSO (Zhang, 2010). This study possesses three main objectives, and the third objective of this study is covariate and lag selection; the WLAdaLASSO would be used only in this case, whereas for the other two objectives, the methodology will remain the same.

Consider a linear regression model where regressors are a set of indicator matrix of SIS. Assume y = $(y_1, y_2, \dots, y_n)$` continuous response regressors, and $I_j$ = $(I_1, I_2, \dots, I_i)$ are a dummy indicator matrix, and $\gamma_j$ is the estimate break coefficient. The regularization techniques for break detection are defined as.

$$\hat{\gamma}_j = argmin_{\hat{\gamma}} \left\| y_{it} - \sum_{i=1}^{j} \gamma_j \, I \right\|_2^2 + \sum_{i=1}^{j} p_{\lambda_j} \left( |\gamma_j| \right) \tag{3.3}$$

However, for covariate and its lag selection we consider a linear regression model where y = $(y_{1t}, y_{2t}, \dots, y_{nt})$`continuous response regressors, and $x_{it}$ = $(x_{1t}, x_{1t-1}, \dots, x_{pt-1})$ covariates with its lag, and $\gamma_j$ estimated coefficients. The equation can be defined as;

$$\hat{\gamma}_j = argmin_{\hat{\gamma}} \left\| y_{it} - \sum_{i=1}^{p} \gamma_{it} \, x_{it} \right\|_2^2 + \sum_{i=1}^{p} p_{\lambda_j} \left( |\gamma_j| \right) \tag{3.4}$$

Where $p_{\lambda_j}(.)$ is a penalty function, and $\lambda_j$ is a penalty parameter. We consider four different forms of $p_{\lambda_j}(.)$. For the estimation of above equation 3.3 and 3.4 we use LASSO, AdaLASSO, SCAD, MCP, and WLAdaLASSO.

### 3.3.4. LASSO and AdaLASSO Estimate

The Least Absolute Shrinkage and Selection Operator (LASSO) is a popular estimation method in a linear regression framework because of lower computation cost of introduced by Tibshirani (1996). The LASSO method is like ridge regression; however, it sets some coefficients precisely equal to zero with a substantial bias. The resulting model is easy to interpret and possesses the most negligible forecast error.

$$\text{(Lasso) } p_{\lambda_j}\left(|\gamma_j|\right) = \lambda_j |\gamma| \tag{3.5}$$

The second term in the above equation is defined as "$L_1$ penalty," and $\lambda$ leads to a sparse solution with a shrinking specific set the coefficients precisely equal to zero with a certain amount of bias. The amount of shrinkage depends upon the selection of $\lambda$, whereas it ranges $0<\lambda<\infty$.

Zou (2006) demonstrated that the LASSO estimator lacks the oracle characteristic and introduced the adaptive LASSO, a simple and effective solution. In contrast, the coefficients in LASSO are all penalized equally in the '$L_1$ penalty. However, in AdaLASSO, each coefficient is given a distinct weight. Zou (2006) illustrates that AdaLASSO can possess the oracle property if the weights are data-dependent and carefully chosen.

$$\text{(Adaptive Lasso) } p_{\lambda_j}(|\gamma_j|) = \lambda_j w_j |\gamma|, \text{ where } w_j = |\hat{\gamma}_j|^{-\tau} \tag{3.6}$$

$\widehat{w}_j = 1/|\hat{\gamma}_j^*|^{\tau}$, $\tau > 0$, and $\hat{\gamma}_j^*$ is an initial parameter estimate. The weights for zero coefficients diverge (to infinity) as the sample size expands, nonzero coefficients converge to a finite constant. To estimate the $\hat{\gamma}_j^*$, Zou (2006) recommended the OLS method. However, when the number of candidate variables exceeds the number of observations, the OLS method does not work. A ridge estimate can be employed as an initial estimator in this case.

### 3.4.2. Weighted Lag Adaptive LASSO (WLAdaLASSO)

The Weighted Lag Adaptive LASSO (WLAdaLASSO) was introduced by Konzen and Ziegelmann, (2016) and established on the concept of Park and Sakaori, (2013) work. It is defined as another type of LASSO estimate specifically for time series modeling with lag structure. The idea is like AdaLASSO and built for the time-series ARDL framework, as the more distant lags have a more negligible effect in predicting the dependent variable, imposing more enormous penalties on them.

$$\text{(Weighted Lag Adaptive Lasso) } p_{\lambda_j}(|\gamma_j|) = \lambda_j w_j |\gamma| \tag{3.7}$$

Here $\widehat{w}_j = (|\ \hat{\gamma}_j^{ridge}\ |\ e^{-\alpha l})^{-\tau}$, l is the lag length, τ > 0, and α ≥ *0* are tuning parameters. Moreover, $\hat{\gamma}_j^*$ is an initial parameter estimate. τ = 1 like in AdaLASSO. To pick α, Konzen & Ziegelmann, (2016) suggest estimating the model for a given λ using a grid (0; 0:5; 1; : : : ; 10) and choose the one with the lowest BIC and the λ parameter selected on the same criteria of the lowest BIC.

### 3.4.3. SCAD and MCP Estimate

Smoothly Clipped Absolute Deviation is unbiased, sparse (i.e. small estimated coefficients automatically set to zero) and fulfills the condition of continuity proposed by Fan & Li, (2001). The smoothly clipped absolute deviation (SCAD) for covariate selection and it lags/dummy indicators it defined as:

$$\text{(SCAD)}\ \ p_{\lambda_j}(|\gamma_j|) = \lambda \left\{ \begin{array}{cc} |\gamma| & if\ |\gamma| \le \lambda, \\ -\frac{(\gamma^2 - 2a\lambda|\gamma| + \lambda^2)}{2(a+1)\lambda} & if\ \lambda < |\gamma| \le a\lambda\ and \\ \frac{1}{2}(a+1)\lambda & if\ |\gamma| \ge a\lambda \end{array} \right\} \qquad (3.8)$$

Where x is the matrix of covariates and its lag, the second term in the above equation is $\sum_{j=1}^{d} p_j(|\gamma_j|; \lambda; \alpha)$ is a penalized term designed to meet all three requirements (unbiasedness, sparsity, and continuity). The SCAD has proven effective in many statistical circumstances, such as cross-sectional regression and time series modeling (Uematsu & Tanaka, 2019). P(γ|λ,α) is a folded concave penalty unlike LASSO it depends on two tuning parameters, penalties depend on λ in a non-multiplicative way, so that P(α|λ) = λP(α). Additionally, the tuning parameter α controls the concavity of the penalty. The maximization of the objective function depends on α and λ, whereas α equals 3.7 and λ is selected via cross-validation (Fan & Li, 2001).

The Minimax Concave Penalty (MCP) introduced by Zhang (2010) is a non-convex penalization strategy that employs spares area up to a particular variable selection threshold, resulting in an unbiased estimate.

$$\text{(MCP) } p_{\lambda_j}(|\gamma_j|) = \lambda \begin{cases} (\lambda - \frac{|\gamma|}{\alpha}\, sign(\gamma) & if\ |\gamma| \leq \alpha\lambda \\ 0 & if\ |\gamma| > \alpha\lambda \end{cases} \tag{3.9}$$

MCP uses $\sum_{j=1}^{d} p_j(|\gamma_j|; \lambda; \alpha)$ regularization path based on the family of non-convex penalty function with two tuning parameters α and λ, where α is fixed, and λ is selected via cross-validation. The tuning parameter λ controls the amount of shrinkage and α concavity of penalty. MCP prevents the spares convexity to a greater extent due to minimizing the maximum concavity (Zhang, 2010). The regularization parameter tends to have a larger α coefficient affords less unbiased and more convexity (Zhang, 2010). SCAD and MCP estimates belong to a family of folded concave penalties, as the P(·) penalty function is neither convex nor concave.

**3.4. Selection of Tuning Parameters for Regularization Techniques**

The selection of the λ tuning parameter is crucial as it governs the complexity of the selected model. The choice of the optimal tuning parameter provides a parsimonious model with a precise prediction performance. Enormous literature exists on the selection of tuning parameters, among them cross-validation and generalized cross-validation are well-known techniques of tuning parameter selection(Craven & Wahba, 1978; Stone, 1974). The tuning parameter is frequently selected using a cross-validation approach to achieve prediction optimality. Such prediction optimality is frequently at odds with covariates selection; however, the objective is to recover the underlying set of sparse variables: frequently, a more prominent penalty parameter is required for covariate selection than the optimal prediction (Bühlmann & Van De Geer, 2011). Also, finding the tuning parameter that will generate the consistent estimator is still unclear, and the cross-validation tuning parameter does not consistently estimate the covariates. However, Pretis et al., (2016 and 2018) used cross-validation and fixed tuning parameters for break detection with LASSO estimate, whereas selecting fixed tuning parameters has not been illustrated in their studies. Although the LASSO estimates have nice properties for covariate selection, they also rely mainly on the Gaussian assumption and a known variance, which may not hold in practice, and standard deviation estimation is not a simple operation (Wang, 2013).

Wang, (2013) investigates the influence of different penalty levels on the $L_1$ PLAD estimator using a universal penalty with numerous upper limits and asymptotic options. $\lambda_1 = \sqrt{1.5\, n\, log\, p}$ $\lambda_2 = \sqrt{2\, n\, log\, p,}$ $\lambda_3 = \sqrt{4\, n\, log\, p,}$ $\lambda_4 = \sqrt{10\, n\, log\, p,}$ it is worth noting that they are all fixed options that do not rely on any assumptions or factors. The results reveal that $\lambda_1$, $\lambda_2$, and $\lambda_3$ perform pretty well in terms of prediction and covariate selection, demonstrating that the $L_1$ PLAD approach can handle a wide variety of penalty levels. Furthermore, a greater $\lambda_4$ causes the estimator to be more biased. A $L_1$ penalized LAD estimate with some linear restrictions is proposed by (Wu et al., 2021). They show that when the dimension of the estimated coefficients p is fixed, the suggested estimation has the Oracle property with adjusted normal variance. When p is substantially more significant than the sample size n, the suggested estimation's error bound is sharper than $\sqrt{k\, log(p)/n}$. In this study, we use fixed tuning parameter for break detection suggested by (Wang, 2013; Wu et al., 2021) as $\sqrt{k\, log(p)/n}$ possess oracle property with adjusted normal distribution.

Information criteria like Akaike Information Criteria (AIC) or Bayesian Information Criteria (BIC) are used as another approach for penalizing the likelihood through the degrees of freedom of the fitted model. Degrees of freedom are frequently used to measure the complexity of a model fit, and we can use them to decide how much regularization to utilize. Meanwhile, in terms of covariate selection and out-of-sample forecast, WLAdaLASSO with a BIC-based tuning parameter possesses optimal results (Konzen and Ziegelmann, 2016).

$$BIC = nlog(\hat{\sigma}^2) + \log(n) + df(\widehat{y}) \tag{3.10}$$

Whereas $\hat{\sigma}^2 = n^{-1} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$ and $df(\widehat{y})$ denotes the degrees of freedom of the fitted model. Hence, we use BIC-based tuning parameter in this study. The BIC- based tuning parameter, on the other hand, is superior to cross-validation for covariate selection, although there is no theoretical justification for this (Bühlmann & Van De Geer, 2011).

### 3.6. Theoretical Comparison

To compare these techniques, we use Gauge, Potency, and out-of-sample RMSE. Gauge is the empirical null retention frequency of how irrelevant covariates are retained, whereas potency is known as correct covariate identifications. The comparison of regularization techniques and Autometrics assessed via a correct zero identification interpreted as potency, and incorrect zero identification referred to as Gauge (Doornik & Hendry, 2015). We use RMSE for in-sample/out-of-sample forecasting to evaluate the performance of concerned techniques in a simulation study and real data analysis. If the approaches correctly identify the accurate model, the estimations of the following parameters should be expected:

1. Gauge approaches to nominal significance level α or tight significance level (0.01 or 0.001).

$$E\left(\frac{\widehat{k_{irel}}}{k_{irel}}\right) \rightarrow \alpha$$

2. Potency approaches 1 if considered estimation techniques efficiently estimate the accurate model.

$$E\left(\frac{\widehat{k_{rel}}}{k_{rel}}\right) \rightarrow 1$$

3. The efficiency of the model is further evaluated via Root Mean Square Error (RMSE) with in sample and out-of-sample forecast.

$$RMSE = \sqrt{\frac{1}{N}\sum_{K=1}^{N}(y_k - \widehat{y_k})^2}$$

4. As the RMSE is not enough parameter to assess the model accuracy, we also use the Mean Absolute Error (MAE) with in sample and out-of-sample forecast.

$$MAE = \frac{1}{N}\sum_{k=1}^{N} |y_k - \widehat{y_k}|$$

# Chapter 4

# Data Generating Process and Simulation Results

## 4.1. Introduction

The study is divided into three sub-objectives in the below section, we elaborate on the DGP for each the objective separately. The DGP for the SIS method has been followed by Castle et al., (2015) for multiple break detection in univariate setup. While for the IIS method, we follow the GDP of Castle et al., (2012) for univariate autoregressive, and for multivariate static DGP, we follow Doornik, (2009). However, as our third objective precisely focused on covariate and its lag selection, we follow the DGP of Konzen & Ziegelmann, (2016).

### 4.1.1. Data Generating Process for Step Indicator Saturation

The empirical and theoretical comparison of regularization techniques, particularly SCAD & MCP for break detection, does not exist in the core of existing literature. For this reason, we opt the Data Generating Process (DGP) of Castle et al., (2015), which provides a convenient base for comparison of regularization techniques with Autometrics for a single break and multiple break detections. The study considers different DGPs of the SIS technique with an unknown break with step indicator, unknown single break with different lengths, single shift at the end of observations, and unknown break with two-step indicators.

**Unknown break with a single indicator**

$$y_t = \delta \times 1\{t \leq T_1\} + \epsilon_t, \qquad \epsilon_t \sim (0,1) \tag{4.1}$$

**Single shift at the end of observations**

$$y_t = 10 - 10 \times 1\{t \geq 76\} + \epsilon_t, \quad \epsilon_t \sim (0,1) \tag{4.2}$$

**Unknown break with two indicators**

$$y_t = \delta\left(1_{\{t<T_2\}} - 1_{\{t<T_1\}}\right) + \epsilon_t, \qquad \epsilon_t \sim (0,1) \tag{4.3}$$

Here t equals 100, which denotes the number of observations the shift coefficient equals 2 and 4, while for single indicator shift $T_1$ equals 35, and for an unknown break with two indicators, $T_1$ equals 25, and $T_2$ equals 35. The study evaluates the computational efficiency of Autometrics and regularization techniques in terms of gauge and potency. The magnitude shift for a single and unknown two-step shift equals 2 and 4: the errors are identically and independently distributed with mean zero and variance 1.

### 4.1.2. Data Generating Process for Impulse Indicator Saturation

The Data Generating Process in this section opted from (Castle et al., 2012; Doornik, 2009), here we consider two different DGP univariate with AR(1) and multivariate static DGP. In multivariate DGP, the models consist of irrelevant regressors and outliers, whereas we only assume with and without intercept for a univariate autoregressive case. We assumed well scatter outlier in multivariate DGP with 5%, 10%, and 20% observations, which is different from Doornik (2009), as it has been illustrated 20% outlier at the end of observations with magnitude coefficients equal to 6 in the static DGP, where the DGP can be defined as:

$$y_t = \delta(I_{81} + \cdots + I_{100}) + 0.5y_{t-1} + \epsilon_t \quad \varepsilon \sim IIN(0,1) \tag{4.4}$$

Multivariate and statics DGP

$$y = 0.1 + \sum_{j=1}^{k*} \beta_j x_{ij} + 6(\tau) + \varepsilon \qquad \varepsilon \sim IIN(0,1) \tag{4.5}$$

In the above data generating process $\delta_i$=1 up to 5 magnitude coefficients and shifts at last 20 observations and $t = 1,2, \dots \dots \dots, 100$, whereas $\epsilon_t \sim IIN(0, \sigma^2)$. Whereas for Multivariate and statics DGP $\beta_1 = \cdots \dots = \beta_{k*} = 1$ for static DGP where relevant regressors k= 20 and k*relevant regressors equal 10 and i = 1, 2, ……, 100. As τ equals to 5%, 10% and 20% of scattered outliers in obtained by assuming $\varepsilon_i \sim (0, \sigma + 4)\ and\ \varepsilon_i \sim (0, \sigma + 6)$.. To estimate the above DGP, we use the Generalized Unrestricted Model (GUM). We introduce an impulse dummy indicator for each observation in the model. The GUM can be illustrated as:

$$y = \alpha + \sum_{j=1}^{k} \beta_j x_{ij} + \sum_{i=1}^{100} \gamma_i I_i + \epsilon \tag{4.6}$$

Where y is a continuous variable that exhibits an unknown outlier and x is a matrix of regressors. γ is a set of outlier coefficients, and i observation equals 100. I is a diagonal identity matrix of each corresponding observation above. $I'_1 = (1{,}0{,}0, \dots \dots.., 0)$, $I'_2 = (0{,}1{,}0{,}0, \dots \dots, 0)$, and $I'_i = (0{,}0{,}0, \dots \dots.., 1)$. Attempting to estimate the above GUM by OLS estimate is not possible as P>N. By default, Autometrics (based on general-to-specific modeling) is used to detect these outliers and estimate the model simultaneously.

**4.1.3. Data Generating Process for covariate and its lag selection.**

We use Konzen & Ziegelmann, (2016) DGP for statistical comparison as the DGP provides a connivance base for comparison in high-dimensional time series analysis, with a varying number of linear dependencies and sample size. Regarding covariate and lag selection the performance of considered techniques are assessed with gauge, potency, RMSE, and MAE. To illustrate our purpose we chose Konzen & Ziegelmann (2016) DGP with 10 independent time series covariates that follow AR(1) as $x_{i,t} = \phi x_{i,t-1} + \mu_{i,t}$ where $\mu_{i,t} \sim N(0,1)$ and i= 1,2,……..,10. We assess the performance of considered techniques under different scenarios based on the same linear model with varying autocorrelation coefficients AR(1) ϕ equals 0.1, 0.5, and 0.8 and T number of observations equal to 50, 100, and 500.

The considered DGP is as fellows.

$$y_t = 0.8y_{t-1} + 0.6x_{1,t-1} + 0.3x_{1,t-2} - 0.5x_{2,t-1} - 0.2x_{2,t-2} + 0.4x_{3,t-1} + 0.3x_{3,t-2} + 0.4x_{4,t-1} - 0.3x_{5,t-1} + 0.2x_{6,t-1} + \varepsilon_t, \qquad \varepsilon_t \sim \mathrm{N}(0,1) \;\; t = 1,2, \dots\dots, T \tag{4.7}$$

We employ WLAdaLASSO, Autometrics, and other regularization techniques to estimate the model. We consider two cases for the lag length of dependent and independent regressors is equal to 2 and 5 throughout the simulation study with varying T observations and ϕ parameter of independent regressors. We eliminate the last ten observations of the simulated series to implement the out-of-sample RMSE. The RMSE of the out-of-sample forecast is reported in the below figures, and the simulation is repeated 1000 times.

### 4.2. Simulation Result and Discussion

In this section we provide simulation result of the above DGP's separately. The results are organized according to the objective of the study. The simulation results in the tables below are reported in terms of average Gauge ('Size'), Potency ('Power'), Root Mean Square Error and Mean Absolute Error. The simulation study has been performed in R-free statistical software; for Autometrics, we used the gets package of R, which is freely available, and for regularization techniques, we use the glmnet for LASSO, AdaLASSO, and WLAdaLASSO as for SCAD and MCP, the ncvreg package. The performance of Autometrics for covariate selection and forecasting is assessed with two levels of significance 0.05 and 0.01. The simulation experiment has been repeated 1000 times. In the below section we elaborate each simulation result according to their DGP individually. In the below tables and figures, we use Auto as a short form of Autometrics.

#### 4.2.1. Simulation result of SIS method

The simulation result of equations 4.1, 4.2, and 4.3 illustrated in below table 4.1, 4.2, and 4.3. The result of regularization technique is present under two different tuning parameter criteria cross-validation and fixed tuning. For fixed tuning parameter, we use $\lambda = \sqrt{2\, log(p)/n}$ which is used most in Wang (2013), as it possess oracle property with adjusted normal distribution (Wang, 2013; Wu et al., 2021).

The results of regularization techniques with Autometrics for a single shift at the end of observations are demonstrated in table 4.1 below. The result illustrates that both considered methods easily detected the breaks with an average potency of close to 1. The empirical gauge for Autometrics is significantly less than the 1% level and retains the least gauge among all existing techniques. However, LASSO and Adaptive LASSO retain a higher gauge than SCAD and MCP with cross-validation tuning parameters. SCAD and MCP with the fixed tuning parameter perform close to Autometrics in terms of gauge. Under the fixed and Cross-Validation tuning parameters, SCAD and MCP perform more efficiently than LASSO and AdaLASSO with least gauge. LASSO and AdaLASSO retain the highest potency at the cost of higher gauge compared to all other methods. However, it also retains the least average RMSE and MAE with 0.976 and 0.776, respectively, among all other candidate estimates.

**Table 4. 1:** Single break at the end of observations

| λ=-10 | | | | |
|---|---|---|---|---|
| | **Gauge** | **Potency** | **RMSE** | **MAE** |
| **SCAD fixed** | 0.006 | 0.958 | 0.9841 | 0.7864 |
| **MCP fixed** | 0.007 | 0.998 | 0.9846 | 0.7846 |
| **LASSO fixed** | 0.0291 | 1 | 0.9905 | 0.7889 |
| **AdaLASSO fixed** | 0.0289 | 1 | 0.9905 | 0.7889 |
| **SCAD CV** | 0.012 | 1 | 0.9836 | 0.7850 |
| **MCP CV** | 0.008 | 0.983 | 0.9879 | 0.7870 |
| **LASSO CV** | 0.0461 | 1 | 0.9755 | 0.7759 |
| **AdaLASSO CV** | 0.0435 | 1 | 0.9755 | 0.7759 |
| **Auto** | 0.0007 | 1 | 0.9845 | 0.7867 |

The simulated outcome of a single shift at the beginning of the observation with various magnitudes is shown in table 4.2. The break at the first observation is more difficult to identify using the considered methods than the break at the last observation. However, regularization methods like Autometrics perform better when the magnitude shift rises. The average potency of LASSO and AdaLASSO is higher than all other techniques; meanwhile, the average gauge is also higher. It indicates that LASSO and AdaLASSO possess higher potency equals 0.855 at the cost of a higher gauge equal 0.0587. SCAD and MCP perform like Autometrics in gauge, potency RMSE and MAE. For break detection LASSO and AdaLASSO perform similarly both in terms of gauge, potency, RMSE and MAE.

**Table 4. 2:** Single Break at First half of observation

| | Gauge | Potency | RMSE | MAE |
|---|---|---|---|---|
| **λ=2** | | | | |
| **SCAD Fixed** | 0.0098 | 0.632 | 0.9795 | 0.7827 |
| **MCP Fixed** | 0.0101 | 0.538 | 0.9785 | 0.7821 |
| **LASSO Fixed** | 0.0314 | 0.853 | 0.9891 | 0.7877 |
| **AdaLASSO Fixed** | 0.0314 | 0.853 | 0.9891 | 0.7877 |
| **SCAD CV** | 0.0134 | 0.633 | 0.9725 | 0.7765 |
| **MCP CV** | 0.0101 | 0.538 | 0.9749 | 0.7785 |
| **LASSO CV** | 0.0587 | 0.855 | 0.9561 | 0.7596 |
| **AdaLASSO CV** | 0.0587 | 0.856 | 0.9561 | 0.7596 |
| **Auto** | 0.005 | 0.619 | 0.9792 | 0.7830 |
| **λ=4** | | | | |
| **SCAD Fixed** | 0.0057 | 0.951 | 0.9835 | 0.7863 |
| **MCP Fixed** | 0.0077 | 0.784 | 0.9875 | 0.7873 |
| **LASSO Fixed** | 0.0319 | 0.992 | 0.9888 | 0.7870 |
| **AdaLASSO Fixed** | 0.0319 | 0.992 | 0.9888 | 0.7870 |
| **SCAD CV** | 0.0094 | 0.952 | 0.9754 | 0.7790 |
| **MCP CV** | 0.0084 | 0.784 | 0.9816 | 0.7819 |
| **LASSO CV** | 0.0598 | 0.992 | 0.9551 | 0.7584 |
| **AdaLASSO CV** | 0.0598 | 0.992 | 0.9551 | 0.7584 |
| **Auto** | 0.001 | 0.93 | 0.9845 | 0.7873 |

Unknown breaks with a two-step shift LASSO and AdaLASSO fail to omit irrelevant breaks due to the forward selection method and possesses the highest gauge among regularization techniques, see table 4.3. However, with cross validation tuning parameters, all the regularization techniques retain higher potency at the cost of higher gauge. With fixed tuning parameters, SCAD possesses higher potency among regularization techniques and as well to Autometrics. The average potency of SCAD is 73%, whereas Autometrics possess 70%, which is slightly lower than SCAD.

The simulation experiment has been conducted under the null hypothesis of no structural break.

The empirical gauge of Autometrics in the simulation study is even less than the theoretical significance level (1%), which is the least among all considered techniques. However, among the regularization techniques, SCAD possesses the least average gauge compared to other techniques. LASSO and AdaLASSO have the highest average potency, with the cost higher average gauge due to the forward selection method. SCAD with multiple unknown shifts retains a higher average potency than Autometrics.

### 4.2.2. Simulation result of IIS method

The comparison of the IIS method estimated via regularization technique is assessed under two different scenarios: univariate AR series and multivariate static model with 5%, 10%, and 20% scattered outlying observations with 6 SD and 4 SD outlying magnitude. We use the glmnet package for R to estimate AdaLASSO and LASSO. For MCP and SCAD estimation, we use the ncvreg package of R, and the ncvreg package uses a coordinate descent algorithm. While for Autometrics, we use the gets package of R. To achieve our study objective, we use a static DGP with orthogonal covariates and dummy indicator saturation opts from Castle et al., (2012) and Doornik, (2009). It provides a convenient base for comparing regularization techniques with Autometrics in the presence of outliers. Results of the simulated scenarios are presented in Tables

4.4, 4.5, and 4.6. The table illustrates the average gauge, potency, and RMSE of out-of-sample Autometrics and regularization techniques. The experiment is repeated 1000 times. As in the above univariate analysis, the AdaLASSO performs identical to LASSO; hence, we do not consider it in the case of univariate AR series.

**Table 4. 3:** Multiple breaks with different shift magnitude

| | | λ=2 | | |
|---|---|---|---|---|
| | **Gauge** | **Potency** | **RMSE** | **MAE** |
| **SCAD Fixed** | 0.015 | 0.734 | 0.9903 | 0.7901 |
| **MCP Fixed** | 0.014 | 0.56 | 0.9900 | 0.7883 |
| **LASSO Fixed** | 0.027 | 0.629 | 1.1292 | 0.8869 |
| **AdaLASSO Fixed** | 0.027 | 0.629 | 1.1292 | 0.8869 |
| **SCAD CV** | 0.021 | 0.727 | 0.9828 | 0.7832 |
| **MCP CV** | 0.018 | 0.563 | 0.9930 | 0.7893 |
| **LASSO CV** | 0.103 | 0.889 | 0.9502 | 0.7512 |
| **AdaLASSO CV** | 0.097 | 0.881 | 0.9502 | 0.7512 |
| **Auto** | 0.007 | 0.709 | 0.9764 | 0.78019 |
| | | **λ=4** | | |
| **SCAD Fixed** | 0.0061 | 0.949 | 0.9820 | 0.7839 |
| **MCP Fixed** | 0.0085 | 0.846 | 1.0060 | 0.7913 |
| **LASSO Fixed** | 0.0296 | 0.984 | 1.1574 | 0.9020 |
| **AdaLASSO Fixed** | 0.0303 | 0.988 | 1.1574 | 0.9020 |
| **SCAD CV** | 0.013 | 0.945 | 0.9868 | 0.7851 |
| **MCP CV** | 0.0107 | 0.842 | 1.0116 | 0.7953 |
| **LASSO CV** | 0.089 | 0.996 | 0.9654 | 0.7624 |
| **AdaLASSO CV** | 0.087 | 0.996 | 0.9654 | 0.7624 |
| **Auto** | 0.001 | 0.964 | 0.9821 | 0.7846 |

**Table 4. 4:** Univariate AR(1) series with single break

| | Gauge | Potency | RMSE | MAE |
|---|---|---|---|---|
| | | **γ=5** | | |
| **SCAD BIC** | 0.129 | 0.999 | 2.282 | 2.110 |
| **MCP BIC** | 0.066 | 0.997 | 2.249 | 2.073 |
| **LASSO BIC** | 0.198 | 0.995 | 2.648 | 2.426 |
| **SCAD fix** | 0.022 | 0.873 | 3.220 | 2.719 |
| **MCP fix** | 0.013 | 0.908 | 2.922 | 2.540 |
| **LASSO fix** | 0.027 | 0.889 | 3.224 | 2.720 |
| **Auto(0.05)** | 0.027 | 0.982 | 3.699 | 2.988 |
| | | **γ=4** | | |
| **SCAD BIC** | 0.177 | 0.984 | 2.358 | 2.188 |
| **MCP BIC** | 0.100 | 0.976 | 2.319 | 2.146 |
| **LASSO BIC** | 0.173 | 0.948 | 2.672 | 2.434 |
| **SCAD fix** | 0.017 | 0.631 | 3.139 | 2.669 |
| **MCP fix** | 0.014 | 0.662 | 3.007 | 2.593 |
| **LASSO fix** | 0.022 | 0.666 | 3.140 | 2.667 |
| **Auto(0.05)** | 0.059 | 0.985 | 3.326 | 2.788 |
| | | **γ=3** | | |
| **SCAD BIC** | 0.104 | 0.626 | 2.661 | 2.392 |
| **MCP BIC** | 0.094 | 0.733 | 2.518 | 2.301 |
| **LASSO BIC** | 0.064 | 0.518 | 2.807 | 2.479 |
| **SCAD fix** | 0.010 | 0.330 | 2.959 | 2.554 |
| **MCP fix** | 0.008 | 0.328 | 2.914 | 2.529 |
| **LASSO fix** | 0.011 | 0.342 | 2.961 | 2.553 |
| **Auto(0.05)** | 0.018 | 0.716 | 2.897 | 2.588 |

The simulation result of univariate Autoregression series with a single break and outliers' magnitude equal to 5, 4, and 3 illustrate in above table 4.4. The result indicates that with γ equals 5 the MCP with BIC-based tuning parameter outperforms the lowest average gauge equal 0.06. Meanwhile, it possesses the highest average potency equal to 0.99, which is even higher than Autometrics. It also possesses the least average MAE, equal to 2.073, which is less than all other techniques. Regularization techniques with fixed tuning parameters perform close to Autometrics in average gauge compared to BIC-based tuning parameters. However, the average potency of regularization techniques with fixed tuning parameters is less than the BIC-based tuning parameter. The simulation result indicates that the BIC tuning parameter performs better in average potency than the fixed tuning parameter. Meanwhile, as the outlier's magnitude decreases to $\gamma$ equal 3 the average potency of overall techniques decreases compared to $\gamma$ equals 5. Among regularization techniques, MCP with BIC tuning parameter retains the highest average potency of 0.733 with the least RMSE (2.518) and MAE (2.301).

**Table 4. 5***:* Simulated results with different percentages of outliers with 6 SD

| **20% outliers** | | |
|---|---|---|
| | **Gauge** | **Potency** |
| **SCAD** | 0.222 | 0.367 |
| **MCP** | 0.222 | 0.367 |
| **LASSO** | 0.611 | 0.767 |
| **AdaLASSO** | 0.333 | 0.433 |
| **Auto(0.05)** | 0.011 | 0.100 |
| **Auto(0.01)** | 0.011 | 0.100 |
| **10% outliers** | | |
| **SCAD** | 0.100 | 0.500 |
| **MCP** | 0.140 | 0.550 |
| **LASSO** | 0.650 | 0.850 |
| **AdaLASSO** | 0.220 | 0.600 |
| **Auto(0.05)** | 0.010 | 0.200 |
| **Auto(0.01)** | 0.000 | 0.200 |
| **5% outliers** | | |
| **SCAD** | 0.048 | 0.600 |
| **MCP** | 0.048 | 0.600 |
| **LASSO** | 0.591 | 0.933 |
| **AdaLASSO** | 0.124 | 0.667 |
| **Auto(0.05)** | 0.000 | 0.534 |
| **Auto(0.01)** | 0.000 | 0.534 |

The results of regularization techniques with Autometrics for covariate selection and outlier detection in terms of gauge and potency are demonstrated in table 4.5. The result indicates that with a 20% outlier in data, Autometrics in average potency perform worse among all existing techniques. On the contrary, LASSO possesses the highest gauge and potency among regularization techniques. Meanwhile, SCAD and MCP perform similarly in both average gauge and potency. The simulation result indicates that as the outlier percentage decreases to 10%, the performance of considered techniques increases. The performance of SCAD and MCP improved with both gauge and potency. With 5% outlying observation, the considered techniques improved further. The SCAD and MCP estimate retains 60% average potency with an average gauge equal 5%.

The results indicate that with 20% and 4 SD outliers, Autometrics perform worse among all existing techniques in average potency, see table 4.6. However, the average potency of SCAD and MCP drastically increased compared to outliers with 6 SD demonstrated in table 4.5. Meanwhile, significant improvement in the average potency of the regularization technique with 4 SD outlier has been observed over 6 SD. However, 4 SD regularization techniques retain a higher average gauge, as Autometrics maintain a theoretical average gauge but at the cost of the least potency. On the contrary, LASSO possesses the highest gauge and potency among regularization techniques, similar to outliers with 6 SD. Compared to LASSO and SCAD, MCP performs significantly in gauge equal to 0.095 and 0.114 of SCAD with 5% outlying observations. The simulation result indicates that as the outlier percentage decreases to 10%, the performance of considered regularization techniques decreases in average potency, whereas the average gauge remains like 20% of outlying observations.

Overall, the simulation result indicates that outliers with 4 SD and 5% outlying observation regularization techniques perform better than 6 SD outliers in average potency. In contrast, the average gauge of regularization techniques with 6 SD is lower than 4 SD outliers. Autometrics possess the least average gauge in all scenarios (5%, 10%, and 20% outlying observations with

6SD and 4SD magnitude) at the cost of the least average potency among all considered techniques. In contrast, LASSO possesses the highest potency and gauge of all other techniques.

**Table 4. 6:** Simulated results with different percentages of outliers with 4 SD

| | **Gauge** | **Potency** |
|---|---|---|
| **20% outliers** | | |
| **SCAD** | 0.222 | 1.000 |
| **MCP** | 0.144 | 1.000 |
| **LASSO** | 0.611 | 0.967 |
| **AdaLASSO** | 0.189 | 0.933 |
| **Auto(0.05)** | 0.000 | 0.367 |
| **Auto(0.01)** | 0.011 | 0.367 |
| **10% outliers** | | |
| **SCAD** | 0.230 | 0.600 |
| **MCP** | 0.150 | 0.550 |
| **LASSO** | 0.650 | 0.850 |
| **AdaLASSO** | 0.360 | 0.700 |
| **Auto(0.05)** | 0.000 | 0.500 |
| **Auto(0.01)** | 0.000 | 0.500 |
| **5% outliers** | | |
| **SCAD** | 0.114 | 0.667 |
| **MCP** | 0.095 | 0.667 |
| **LASSO** | 0.657 | 0.867 |
| **AdaLASSO** | 0.352 | 0.667 |
| **Auto(0.05)** | 0.000 | 0.667 |
| **Auto(0.01)** | 0.000 | 0.667 |

The out-of-sample forecasting performance of the considered techniques is represented in Figure 4.1-4.6. The graph illustrates that the average RMSE error of LASSO with 20% and 10% outlier observations is the least among all considered techniques. The result aligns with existing literature as LASSO possesses the least forecasting error and selects a more irrelevant regressor (which can be observed from table 1) (Lee, 2015). However, with 5% outlier observations, Autometrics with 0.01 retain the highest RMSE and MAE than all other techniques, see figure 4.3. Conversely, Autometrics with 5% outlying observations possess the least gauge but retain the least RMSE and MAE, figure 4.3. Autometrics with 0.05 level of significance possesses less RMSE and MAE than

0.01 level of significance, the fact that Autometrics with 0.01 level of significance omit relevant regressors increases the average RMSE.

There is a significant improvement in average RMSE with 4 SD with 5% and 20% outlying observations compared to 6 SD magnitude with 5% and 20% outlying observations. These differences can be justified as with 5% and 4 SD outliers; the average potency is higher (means that method correctly identified the correct variables/dummy indicator) compared to 6 SD. However, it ultimately impacts the out-of-sample RMSE, and the same pattern can be observed with 20% outlying observations and 6 SD. The average potency is the least; consequently, the out-of-sample RMSE increases. However, the average potency of 20% of outlying observations with 4 SD is close to 1 for regularization techniques. Due to this, the out-of-sample RMSE of regularization techniques is least compared to 6 SD, as shown in figure 4.4. Figure 4.5 shows that Autometrics with 0.01 level retain a high RMSE equal 2.657 and the least MAE retained by SCAD and MCP equals 1.28 and 1.24, respectively. Whereas figure 4.6 illustrate that regularization techniques and Autometrics found it easier to select the model with the least average RMSE among all other experiments, even the RMSE of Autometrics with 0.01 level is least compared to 10% and 20% level of significance.

Additionally, the overall performance of regularization techniques and Autometrics improved with 5% outlying observations. Among regularization techniques, SCAD and MCP perform robustly in gauge and potency even with 20% outlying observations. Autometrics possess the least average gauge and potency simultaneously among all considered techniques. In contrast, LASSO possesses the highest potency and gauge of all other techniques.

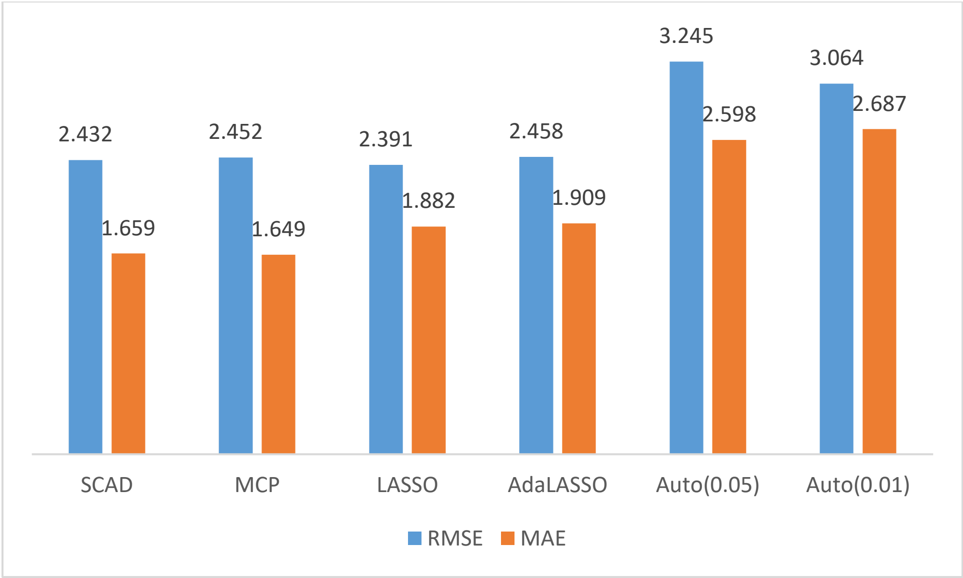



***Figure 4. 1:*** 20% Outliers with 6 SD

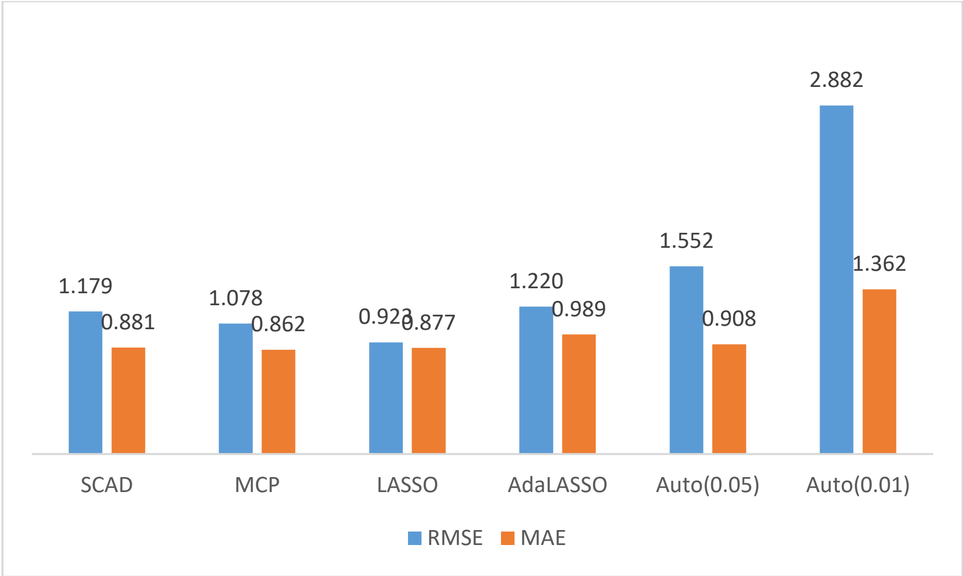



***Figure 4. 2:***10% Outliers with 6 SD

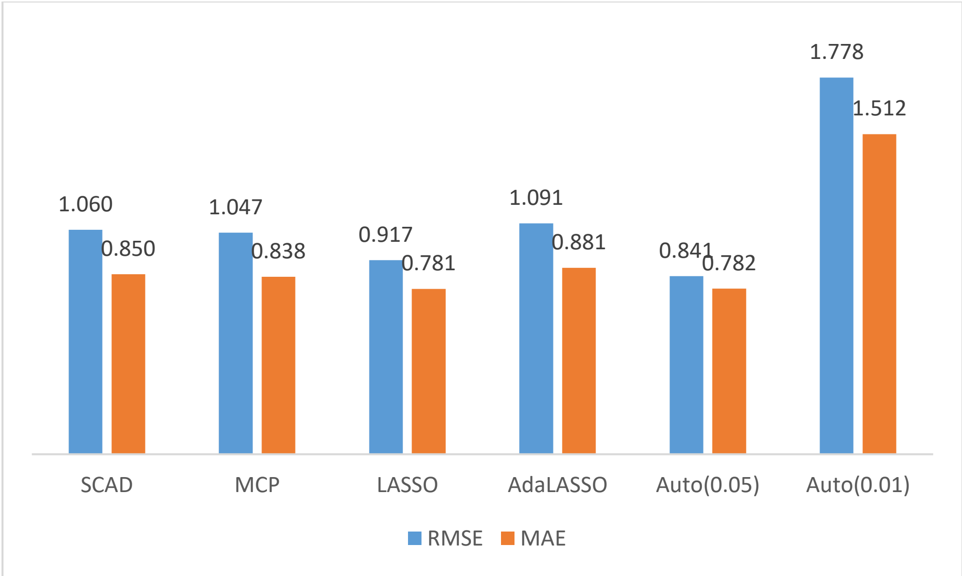



***Figure 4. 3:*** *5*% Outliers with 6 SD

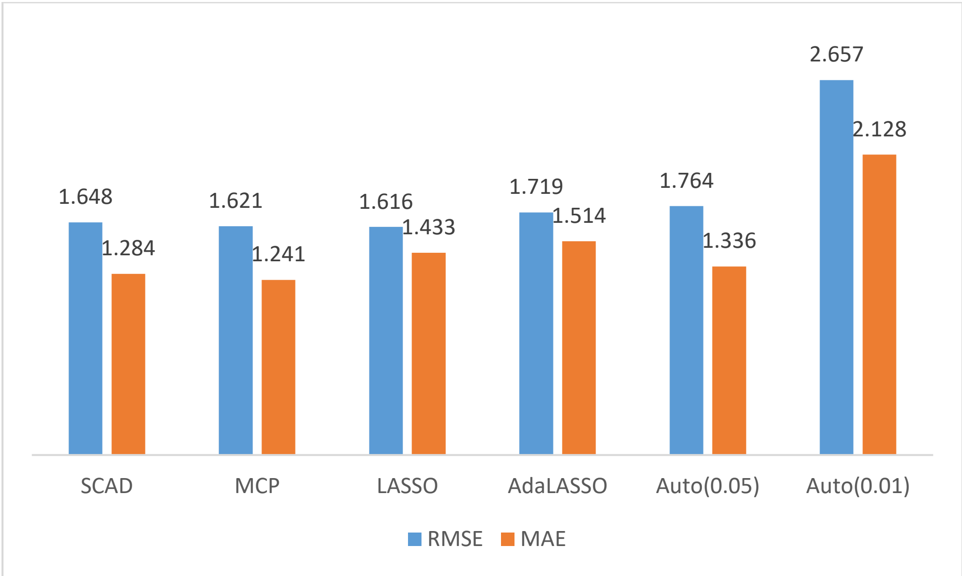



***Figure 4. 4:*** *20*% Outliers with 4 SD

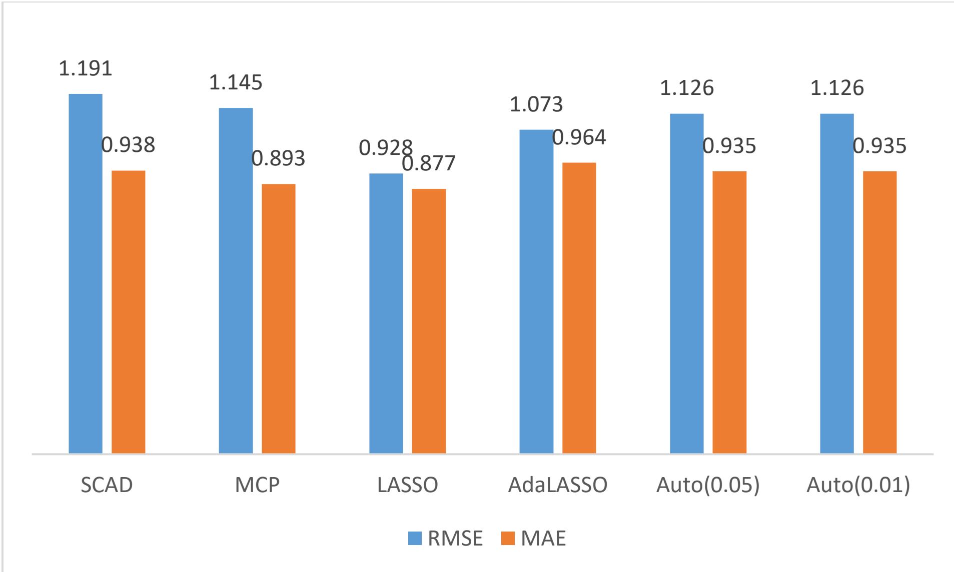



***Figure 4. 5:*** *10*% Outliers with 4 SD

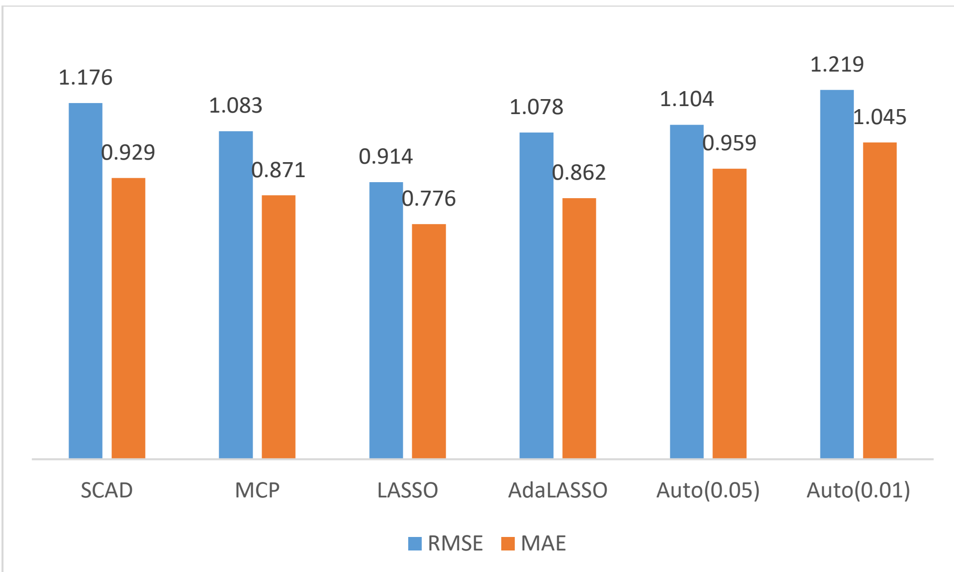



***Figure 4. 6:*** *5*% Outliers with 4 SD

The out-of-sample forecasting performance of the considered techniques is represented in figures 4.1-4.6. Figure 4.6 illustrates that the average LASSO RMSE error with 5% outlying observations and 4Sd possess the least RMSE among all considered techniques. The result aligns with existing literature as LASSO possesses the least forecasting error and selects a more irrelevant regressor (which can be observed from table 4.1) (Lee, 2015). We observed that Autometrics, with 5%

outlying observations, possess the least gauge but retain higher RMSE than SCAD and MCP. Autometrics with a 0.05 level of significance possess least MAE than 0.01 level of significance, the fact that Autometrics with 0.01 level of significance omit relevant regressors increases the average RMSE.

### 4.2.3. Simulation result of covariate and its lag selection

We consider two cases for the lag length of dependent and independent regressors equal to 2 and 5. However, in the below tables, we only report the simulation of the result of leg length equal to 5. The simulation results of lag length two are presented in table A.5 for the Appendix. we observed that the simulation result of l equal to 2 does not significantly differ from l equal to 5 with T observation equal to 50 gauge, and potency for $\phi$ equal 0.1, 0.5, and 0.8 are identical.

The simulation findings of considered techniques in terms of average gauge and potency are in table 4.7-4.9. The simulated result of out-of-sample RMSE is presented below figures. Table 4.7 indicates that among all concerned techniques, WLAdaLASSO outperforms in potency at 63.6%, with T at 50. In time series dynamic modeling, the empirical average gauge of Autometrics is 0.069, slightly higher than the nominal 5% level with $\phi$ equal to 0.1. The same outcome is observed in the case of Autometrics with 1% level; the average gauge retained in simulation equals 0.028.

Meanwhile, Autometrics retains the least potency of 16.1% among all existing techniques. As the sample size increases, the technique's performance improves in average potency (increases) and average gauge (decreases). However, with an increase in sample T equal to 500, the Autometrics with 0.05 significance perform near WLAdaLASSO in potency and gauge. Among regularization techniques, LASSO, AdaLASSO, SCAD, and MCP perform inferior to WLAdaLASSO in gauge and potency.

**Table 4. 7*:*** Simulated result of with ϕ equal 0.1

| | | T=50 | T = 100 | T=500 |
|---|---|---|---|---|
| **WLAdalasso** | **Gauge** | 0.268 | 0.055 | 0.014 |
| | **Potency** | 0.636 | 0.708 | 0.954 |
| **Autometrics(0.05)** | **Gauge** | 0.069 | 0.037 | 0.009 |
| | **Potency** | 0.256 | 0.375 | 0.960 |
| **Autometrics(0.01)** | **Gauge** | 0.028 | 0.016 | 0.001 |
| | **Potency** | 0.161 | 0.257 | 0.892 |
| **LASSO** | **Gauge** | 0.465 | 0.186 | 0.134 |
| | **Potency** | 0.619 | 0.612 | 0.713 |
| **AdaLASSO** | **Gauge** | 0.197 | 0.092 | 0.080 |
| | **Potency** | 0.389 | 0.523 | 0.681 |
| **SCAD** | **Gauge** | 0.176 | 0.174 | 0.098 |
| | **Potency** | 0.380 | 0.584 | 0.700 |
| **MCP** | **Gauge** | 0.178 | 0.155 | 0.083 |
| | **Potency** | 0.360 | 0.553 | 0.694 |

The simulated result in Tables 4.8 and 4.9 illustrate ϕ (Autocorrelation coefficients) of regressors equal 0.5 and 0.8. The WLAdaLASSO estimate outperforms an average potency of 64.5%, ϕ equal 0.5, and T equals 50. As the T sample of WLAdaLASSO increases, the average gauge approaches nominal significance level, and average potency approaches 1. The simulation result indicates that the WLAdaLASSO estimate is not sensitive to Autocorrelation coefficients as with ϕ equal 0.1 and T equal 50, the average retain potency equal 63.6%, and 64.5% with ϕ equal 0.5. However, Autometrics performs poorly as the Autocorrelation coefficient increases from 0.1 to 0.5.

Meanwhile, Autometrics with ϕ equal to 0.8 and T equal to 50 possess 11.5% gauge, which is higher than the 5% significance level, illustrated in table 4.9. The performance of Autometrics does not enhance (gauge $\rightarrow\alpha$, and potency $\rightarrow 1$) as the sample size increases with ϕ equal to 0.8. However, WLadaLASSO performs better in average potency and gauge than all other techniques. The simulation experiment indicates that WLadaLASSO performs robustly even with a stronger linear dependence between predictors. With increasing samples, the performance of Autometrics, LASSO, AdaLASSO, SCAD, and MCP does not enhance as the WLAdaLASSO.

**Table 4. 8:** Simulated result of with ϕ equal 0.5

| | | T=50 | T = 100 | T = 500 |
|---|---|---|---|---|
| **WLAdalasso** | **Gauge** | 0.210 | 0.055 | 0.013 |
| | **Potency** | 0.649 | 0.830 | 0.997 |
| **Autometrics(0.05)** | **Gauge** | 0.065 | 0.042 | 0.033 |
| | **Potency** | 0.280 | 0.407 | 0.807 |
| **Autometrics(0.01)** | **Gauge** | 0.034 | 0.018 | 0.025 |
| | **Potency** | 0.243 | 0.306 | 0.704 |
| **LASSO** | **Gauge** | 0.510 | 0.199 | 0.125 |
| | **Potency** | 0.692 | 0.660 | 0.717 |
| **AdaLASSO** | **Gauge** | 0.251 | 0.097 | 0.076 |
| | **Potency** | 0.460 | 0.520 | 0.691 |
| **SCAD** | **Gauge** | 0.210 | 0.175 | 0.093 |
| | **Potency** | 0.351 | 0.581 | 0.699 |
| **MCP** | **Gauge** | 0.221 | 0.148 | 0.079 |
| | **Potency** | 0.350 | 0.556 | 0.691 |

**Table 4. 9:** Simulated result of with ϕ equal 0.8

| | | T=50 | T=100 | T=500 |
|---|---|---|---|---|
| **WLAdalasso** | **Gauge** | 0.220 | 0.056 | 0.011 |
| | **Potency** | 0.696 | 0.707 | 0.992 |
| **Autometrics(0.05)** | **Gauge** | 0.115 | 0.050 | 0.075 |
| | **Potency** | 0.282 | 0.471 | 0.601 |
| **Autometrics(0.01)** | **Gauge** | 0.083 | 0.027 | 0.066 |
| | **Potency** | 0.206 | 0.437 | 0.522 |
| **LASSO** | **Gauge** | 0.298 | 0.231 | 0.157 |
| | **Potency** | 0.587 | 0.683 | 0.719 |
| **AdaLASSO** | **Gauge** | 0.112 | 0.108 | 0.078 |
| | **Potency** | 0.327 | 0.516 | 0.683 |
| **SCAD** | **Gauge** | 0.078 | 0.177 | 0.091 |
| | **Potency** | 0.339 | 0.537 | 0.675 |
| **MCP** | **Gauge** | 0.071 | 0.150 | 0.083 |
| | **Potency** | 0.291 | 0.528 | 0.680 |

The WLAdaLASSO performs superior to other considered regularization techniques and as well as to Autometrics in average gauge and potency even with higher and weak linear dependency between predictors and small sample size.

The result shows that the WLAdaLASSO outperforms in out-of-sample forecasting compared to other techniques with the least RMSE equals 1.33, MAE equals 1.085 with $\phi$ equal 0.8, and T equals 50. Additional metrics and factors should also be considered when assessing forecasting models since the RMSE alone does not give a comprehensive view of the model's predictive accuracy. We use Diebold-Mariano test to determine whether one method is significantly more accurate than the other to test the significance of forecasting accuracy. The Diebold-Mariano test result can be found in table A.6 of the appendix. The Diebold test statistic is equal to -5.27, which indicates that, at a particular level of significance, the difference in mean squared errors (MSE) between the two models is statistically significant. The negative sign implies that WLAdaLASSO has a much lower MSE than Autometrics (0.05).

The WLAdaLASSO has an RMSE of 1.332, and Autometrics (0.05) has an RMSE of 2.39; WLAdaLASSO has a lower root mean squared error (RMSE) than Autometrics (0.05). The RMSE measures the average difference between the expected and actual values in the data set; therefore, a lower RMSE implies higher prediction accuracy. Based on the Diebold test statistics and the RMSE values, it can be inferred that WLAdaLASSO is statistically substantially more accurate than Autometrics (0.05). However, with $\phi$ equals 0.8 and increasing sample size T equals 100 and 500, the Autometrics (0.01 and 0.05) forecasting accuracy does not enhance. WLAdaLASSO outperforms in forecasting accuracy; see table A.6 of the appendix.

***Table 4. 10:*** Simulated result of RMSE and MAE

| | T= 50 RMSE | MAE | T=100 RMSE | MAE | T=500 RMSE | MAE |
|---|---|---|---|---|---|---|
| **ϕ equal 0.1** | | | | | | |
| **SCAD** | 1.653 | 1.423 | 1.245 | 1.006 | 1.003 | 0.989 |
| **MCP** | 1.838 | 1.539 | 1.23 | 1.01 | 0.988 | 0.785 |
| **LASSO** | 1.723 | 1.496 | 1.496 | 1.236 | 1.046 | 0.964 |
| **AdaLASSO** | 1.603 | 1.394 | 1.184 | 1.019 | 0.998 | 0.745 |
| **WLAdaLASSO** | 1.427 | 1.238 | 1.229 | 0.995 | 1.053 | 0.854 |
| **Autometrics (0.05)** | 1.81 | 1.501 | 1.367 | 1.139 | 0.998 | 0.745 |
| **Autometrics (0.01)** | 1.71 | 1.435 | 1.393 | 1.194 | 1.016 | 0.987 |
| **ϕ equal 0.5** | | | | | | |
| **SCAD** | 1.8 | 1.574 | 1.155 | 0.994 | 0.99 | 0.726 |
| **MCP** | 1.981 | 1.632 | 1.152 | 0.984 | 0.987 | 0.709 |
| **LASSO** | 1.892 | 1.596 | 1.107 | 0.963 | 1.007 | 0.827 |
| **AdaLASSO** | 1.808 | 1.523 | 1.087 | 0.945 | 0.988 | 0.712 |
| **WLAdaLASSO** | 1.404 | 1.204 | 1.241 | 1.006 | 1.022 | 0.808 |
| **Autometrics (0.05)** | 1.685 | 1.401 | 1.328 | 1.041 | 1.017 | 0.834 |
| **Autometrics (0.01)** | 1.5 | 1.356 | 1.376 | 1.095 | 1.072 | 0.865 |
| **ϕ equal 0.8** | | | | | | |
| **SCAD** | 1.829 | 1.509 | 1.474 | 1.256 | 0.99 | 0.756 |
| **MCP** | 1.85 | 1.592 | 1.463 | 1.249 | 0.988 | 0.699 |
| **LASSO** | 2.16 | 1.845 | 1.251 | 1.009 | 1.007 | 0.801 |
| **AdaLASSO** | 1.963 | 1.691 | 1.374 | 1.15 | 0.992 | 0.706 |
| **WLAdaLASSO** | 1.33 | 1.085 | 1.116 | 0.981 | 1.016 | 0.81 |
| **Autometrics (0.05)** | 2.39 | 2.035 | 1.715 | 1.391 | 1.234 | 0.845 |
| **Autometrics (0.01)** | 2.252 | 1.956 | 1.612 | 1.286 | 1.233 | 0.868 |

However, the performance of other regularization techniques and Autometrics decreases as ϕ equal 0.8 with T equals 50. The WLAdaLASSO estimate is insensitive to autocorrelation coefficients, as the forecast performance and average potency have not decreased even ϕ equals 0.8 with a small sample T equals 50. However, with ϕ equals 0.8, all other techniques perform poorly in out-of-sample forecasting as the RMSE and MAE of all estimates increase. While WLAdaLASSO possesses the least RMSE. Autometrics with autocorrelation coefficients equal to 0.1 and T equal to 50 perform poorly in RMSE compared to WLAdaLASSO; with sample size increment, the RMSE decreases because the average potency increases. However, SCAD and AdaLASSO possess the least RMSE equals 1.653 and 1.603, with T equal to 50 $\phi$ equals 0.1; with increasing sample size, it decreases further. Autometrics with a $\phi$ equals 0.1 and T equals 500 outperform with the least RMSE equal 0.998 among all other techniques. The overall simulation result indicates that WLAdaLASSO outperforms Autometrics and other regularization techniques in potency and out-of-sample forecasting in terms of higher linear dependency and small samples.

The above simulation result indicates that regularization techniques with MCP and SCAD with fix tuning parameter $\lambda = \sqrt{2\ log(p)/n}$ provide a promising result for break detection compared to cross-validation tuning parameters. Overall, the performance of SCAD and MCP is close to Autometrics in average potency. The overall empirical analysis indicates that regularization techniques with fixed tuning parameters outperform cross validation tuning parameters in terms of correct break detection and RMSE. MCP and SCAD outperform in with least gauge and higher potency with fixed tuning parameters among regularization techniques compared to LASSO and AdaLASSO. The performance of SCAD and MCP is close to Autometrics in real data analysis and as well as in simulation experiment. However, the LASSO and AdaLASSO perform worse in the average gauge. LASSO and AdaLASSO for multivariate break detection are poor; the finding is aligned with (Castle et al., 2015b).

Meanwhile, for multivariate static modeling with outlier, the SCAD and MCP outperform RMSE and potency compared to LASSO and AdaLASSO. Additionally, for covariate and its lag selection, WLAdaLASSO outperforms all considered techniques even with a small size and strong linear dependence between predictor variables. However, Autometrics with $\varphi$ equals 0.1 as the sample size increases the average potency → 1 and average gauge → α, and average RMSE decreases compared to a small sample. On the other hand, in Autometrics with φ equal to 0.8, even with increasing sample size, the average potency does not → 1, and the average gauge does not → α.

# Chapter 5

# Real Data Analysis

## 5.1. Introduction

This section thoroughly elaborates on our underlying data set as an introduction to the data set before aligning our simulation findings with real data analysis. Each subheading of this chapter is related to our study objective to avoid ambiguity. The real data analysis of the concerned variables and their results are illustrated in tables and figures. The real data analysis has been performed in R-free statistical software; for Autometrics, we used the gets package of R, which is freely available, and for regularization techniques, we use the glmnet for LASSO, AdaLASSO, and WLAdaLASSO as for SCAD and MCP, the ncvreg package.

## 5.2. Real data analysis for SIS method

It might be useful to review the significant historical events in Pakistan from 1947 to 2019 before starting the empirical investigation. Since independence, Pakistan has been overwhelmed with political and socio-economic chaos, which has taken a peal on its economy. The political uncertainty and absence of democracy have underprivileged the country of an uneven record of a long-term vision, direction, and continuity of economic policies. In 1960, a huge influx of American aid and political permanence enabled Pakistan to endure high growth rates (Khan, 2002; Zaidi, 2005). Due to increasing interregional economic discrepancy, East Pakistan was dismayed alongside West Pakistan and became an autonomous Bangladesh in 1971. Extremely severe socio-economic conditions caused by the Pakistan-India war of 1971, the East Pakistani territorial issue, and the elected government's empowerment of socialism (Hasan et al., 1997; Husain, 2000; Zaidi, 2005). Due to the oil price shock, there was an upsurge in Pakistan's import bill in October 1973, 1974-77 global depression, letdowns of cotton crops in 1974-75, pest attacks on crops, and vast floods in 1973, 1974, and 1976-77 and 1972-77 inflation has been experienced, with 15% prices increment per annum (Hasan et al., 1997). Pakistan accomplished an average growth rate of above 5% over four decades ending 1988-89. In the 1990s, the second-worst inflation occurred in the wake of decreasing growth rates of GDP. The diminished growth rate prevailed until 2001 as the growth rate declined to less than 4% per year due to the "endorsed Debt Reduction and Management Committee judged the high public debt", an era of macroeconomic crises (Anjum & Sgro, 2017). Despite improvement in the growth rate 2004-05, as the growth rate was 8.6%, the following years were considered by growth slowdown, inflation upsurge, energy crisis, and decline in fiscal and balance of payments positions (Anjum & Sgro, 2017). We use Pakistan's GDP growth and GDP deflator variables for breakpoint detection in the mean from 1960 to 2019. The data is fetched from the World Data Indicator.

Before moving towards the simulation and real data analysis, we first analyze the SIS techniques for de-trend macroeconomic variables of Pakistan and assess the performance of Autometrics with Autoregressive AR(1) and without AR(1). The result of this analysis can be found in the Appendix section. The variable of this analysis has been taken GDP growth, Interest rate, Inflation rate, and unemployment rate. Tables A 3 and 4 indicate that the unemployment and interest rates are stationary with breaks and AR(1) series has less than the unity coefficient with multiple breaks. It indicates that the series depends on its past values and shifts but doesn't possess a unit root. Autometrics uses multi-path search algorithms for break detection with a 5% or 1% significance level. The result indicates that Autometrics without AR(1) series and 1% significance level omits the relevant break at the end of observations, see Table A 1. However, with AR(1) series, it estimates such breaks easily. The study indicates multiple structural breaks in unemployment, GDP growth, interest, and inflation rate. This per-analysis indicates that Autometrics omit significant breaks at the end of observation; for this reason, we use regularization techniques to assess the performance of regularization techniques for break detection.

### 5.2.1. GDP Growth

The graphical overview of break detection of each considered technique in GDP growth variable is presented in Figures 5.1 and 5.2. The graph indicates that Autometrics with 0.05 significances level omit the breaks that occurred at the last observations in 2018. However, regularization techniques with fixed tuning parameters efficiently estimate the breaks that occur at the last of observations, see figure 5.1. The graph shows that in 1970 the break coefficient of LASSO and AdaLASSO downward bias is compared to SCAD and MCP. Among regularization techniques, SCAD and MCP performed better than LASSO and AdaLASSO, as they possess the lowest bias estimate and select less irrelevant breaks.

The performance of regularization and Autometrics in Root Mean Square Error (RMSE) as like breaks are unknown, illustrated in table 5.1. LASSO and AdaLASSO possess higher RMSE than SCAD and MCP. Meanwhile, the LASSO and AdaLASSO select more irrelevant breaks for break detection; the results align with simulation experiments. In 1970 the GDP growth decreased to 10%; however, LASSO and AdaLASSO estimated coefficient is 1.815, while SCAD and MCP coefficients are 6.847 and 6.992, respectively. MCP performed close to automatics with RMSE of 1.371, the least RMSE among regularization techniques, see figure 5.3.

The regularization techniques with the Cross-Validation tuning parameter are illustrated in figure 5.2. The graphical visualization indicates that the cross-validation tuning parameter fails to select relevant breaks. It also shows that the estimated coefficients of all considered regularization techniques are downward biased, which underestimated the break coefficient and enhanced RMSE.

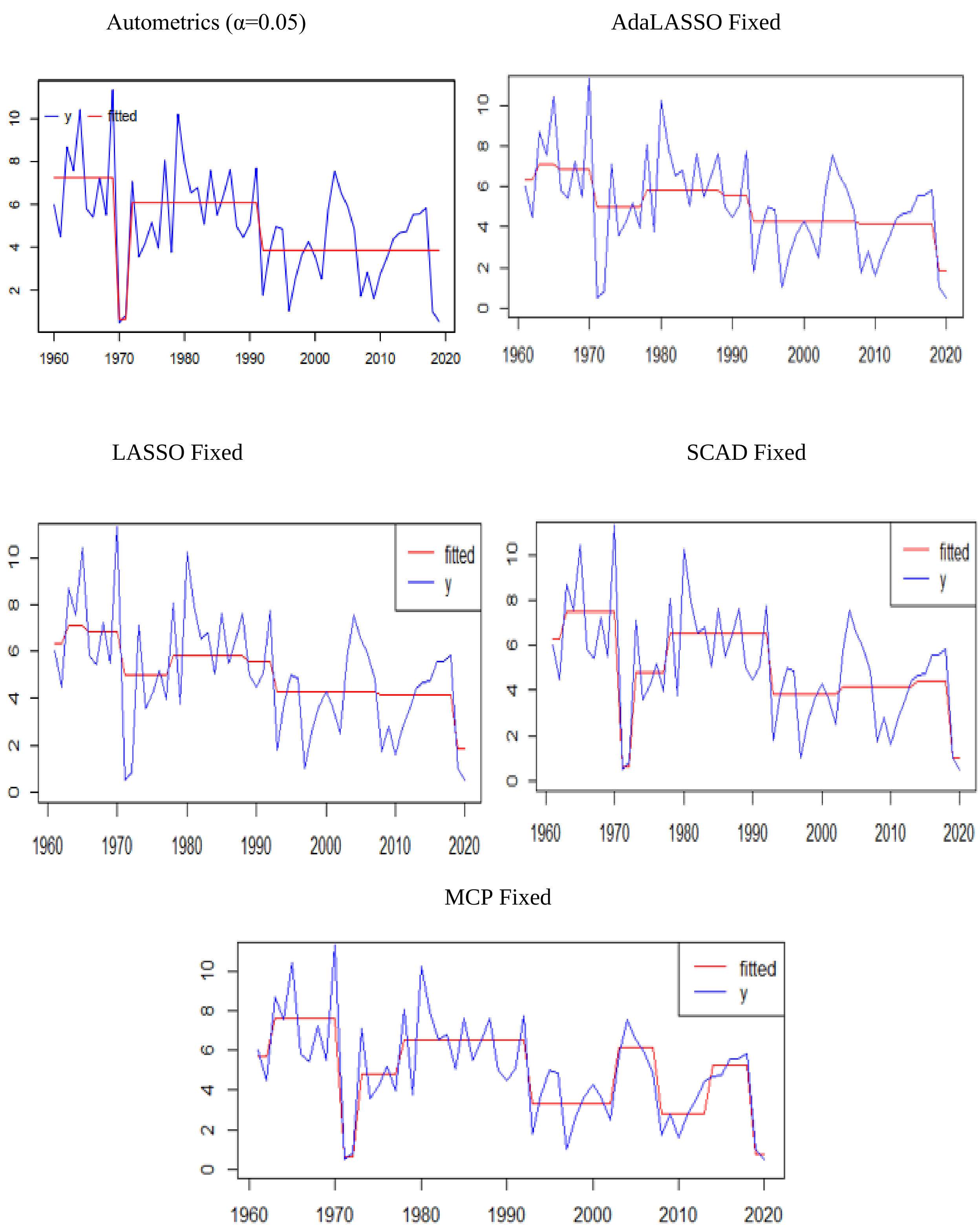



***Figure 5. 1:*** Fitted GDP Growth plot under various regularization techniques with fixed tuning parameter.

**Table 5. 1:** GDP Growth break detection via regularization techniques with fixed tuning parameter

| | | | | | |
|---|---|---|---|---|---|
| | | **Autometrics with 0.05 level of significance** | | | |
| | **Intercept** | **sis1969** | **sis1970** | **sis1972** | **sis1979** |
| **Coef.** | 6.787 | 4.567 | -10.712 | 4.463 | 1.502 |
| | **sis1992** | **sis2002** | **sis2007** | **sis2012** | **sis2018** |
| **Coef.** | -3.320 | 2.829 | -3.636 | 2.641 | -4.363 |
| | | **LASSO with Fixed tuning parameter** | | | |
| | **Intercept** | **sis1962** | **sis196** | **sis1970** | **sis1977** |
| **Coef.** | 6.311 | 0.758 | -0.244 | -1.815 | 0.818 |
| | **sis1988** | **sis1992** | **sis2007** | **sis2018** | |
| **Coef.** | -0.281 | -1.249 | -0.124 | -2.339 | |
| | | **AdaLASSO with Fixed tuning parameter** | | | |
| | **Intercept** | **sis1962** | **sis1965** | **sis1970** | **sis1977** |
| **Coef.** | 6.311 | 0.758 | -0.244 | -1.815 | 0.818 |
| | **sis1988** | **sis1992** | **sis2007** | **sis2018** | |
| **Coef.** | -0.281 | -1.249 | -0.124 | -2.339 | |
| | | **SCAD with Fixed tuning parameter** | | | |
| | **Intercept** | **sis1962** | **sis1970** | **sis1972** | **sis1977** |
| **Coef.** | 6.266 | 1.222 | -6.847 | 4.143 | 1.728 |
| | **sis1992** | **sis2002** | **sis2012** | **sis2013** | **sis2014** |
| **Coef.** | -2.677 | 0.319 | 0.017 | 0.211 | 0.0009 |
| | **sis2018** | | | | |
| **Coef.** | -3.350 | | | | |
| | | **MCP with Fixed tuning parameter** | | | |
| | **Intercept** | **sis1962** | **sis1970** | **sis1972** | **sis1977** |
| **Coef.** | 5.684 | 1.949 | -6.992 | 4.143 | 1.728 |
| | **sis1992** | **sis2002** | **sis2007** | **sis2013** | **sis2018** |
| **Coef.** | -3.227 | 2.829 | -3.316 | 2.466 | -4.507 |

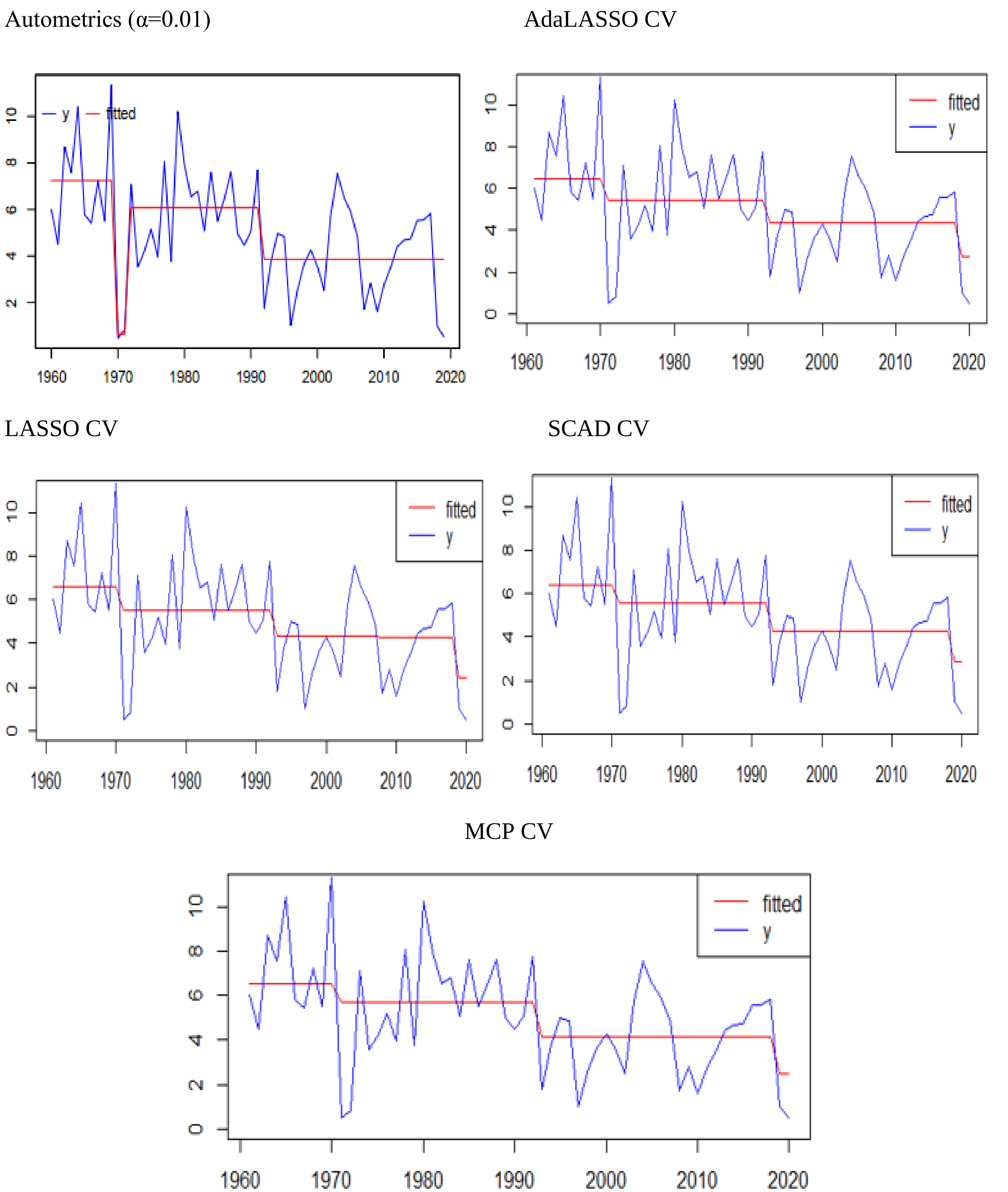



***Figure 5. 2:*** Fitted GDP Growth plot under various regularization techniques with cross-validation.

**Table 5. 2:** GDP Growth Break Detection via Regularization Techniques with Cross-Validation

| | | | | |
|---|---|---|---|---|
| **Autometrics with 0.01 level of significance** | | | | |
| | **Intercept** | **sis1970** | **sis1972** | **sis1992** |
| **Coef.** | 7.24333 | -6.60244 | 5.43918 | -2.22094 |
| **LASSO with Cross-Validation tuning parameter** | | | | |
| | **Intercept** | **sis1970** | **sis1992** | **sis2007** |
| **Coef.** | 6.563 | -1.082 | -1.147 | -0.028 |
| | **sis2018** | | | |
| **Coef.** | -1.909 | | | |
| **AdaLASSO with cross-validation tuning parameter** | | | | |
| | **Intercept** | **sis1970** | **sis1992** | **sis2007** |
| **Coef.** | 6.496 | -1.026 | -1.127 | -0.0008 |
| | **sis2018** | | | |
| **Coef.** | -1.787 | | | |
| **SCAD with cross-Validation tuning parameter** | | | | |
| | **Intercept** | **sis1970** | **sis1992** | **sis2018** |
| **Coef.** | 6.479 | -0.808 | -1.493 | -1.579 |
| **MCP with Cross-Validation tuning parameter** | | | | |
| | **Intercept** | **sis1970** | **sis1992** | **sis2018** |
| **Coef.** | 6.542 | -0.868 | -1.513 | -1.671 |

We estimate Autometrics with a 0.01 level of significance. Regularization techniques with Cross-validation possess higher RMSE than Autometrics and fixed tuning parameters in table 5.2 above. In 1970 the GDP growth decreased to 6%; however, LASSO and AdaLASSO estimated coefficient is 1.082, while SCAD and MCP coefficients are 0.808 and 0.868, respectively. Autometrics does not detect the break at the end of observation; all considered regularization techniques estimate it efficiently. For break detection in GDP growth, Autometrics is estimated under two levels of significance, 0.05 and 0.01. Autometrics with a 0.05 level retains the least RMSE of 1.305, then 0.01 level of significance, and RMSE equals 1.81, see figure 5.3.

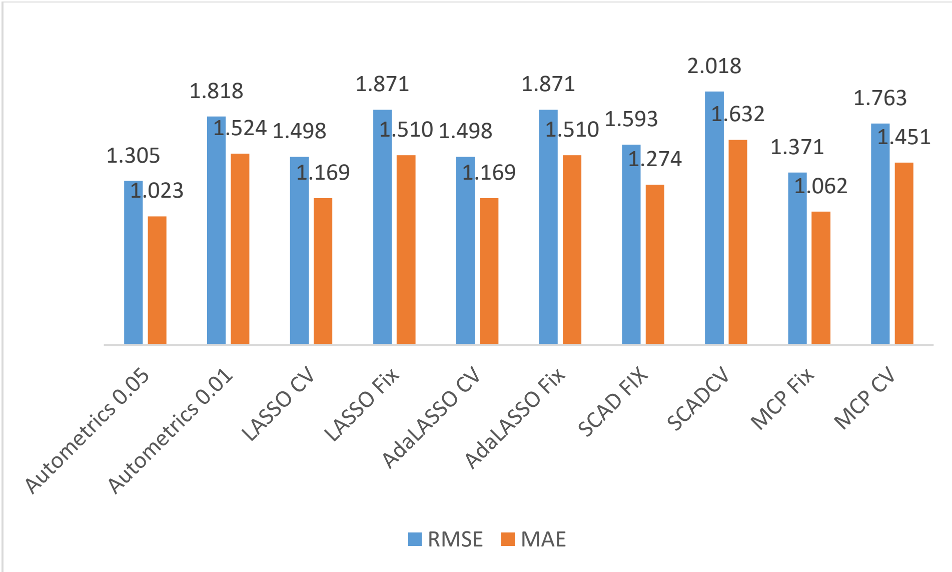



***Figure 5. 3*****:** RMSE and MAE of GDP growth

### 5.2.2. GDP Deflator

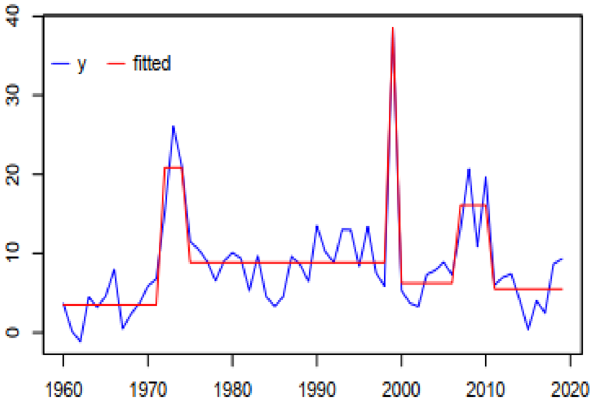
The graphical overview of break detection in GDP deflator of each considered technique with fixed tuning parameter, figure 5.4. The graphic visualization indicates that the fixed tuning parameter with LASSO and AdaLASSO underestimates the break coefficient comparatively to SCAD and MCP. However, on average, regularization techniques with LASSO and AdaLASSO select more irrelevant breaks.

The LASSO and AdaLASSO select the break compared to SCAD and MCP with 3.74 RMSE illustrated in table 5.3. In contrast, SCAD possesses the least RMSE, equal to 2.518 and even less than Autometrics. The analysis indicates that Autometrics missed the break in 2018; however, regularization techniques estimate it efficiently. The estimated breaks are relevant to historical events in Pakistan; the 1970 oil crisis, the 1999 break is the consequence of the 1998 atomic test, 2007 related to the global financial crisis, and the 2018 political instability. SCAD and MCP detect a slightly higher break than Autometrics but fewer irrelevant breaks than LASSO and AdaLASSO.

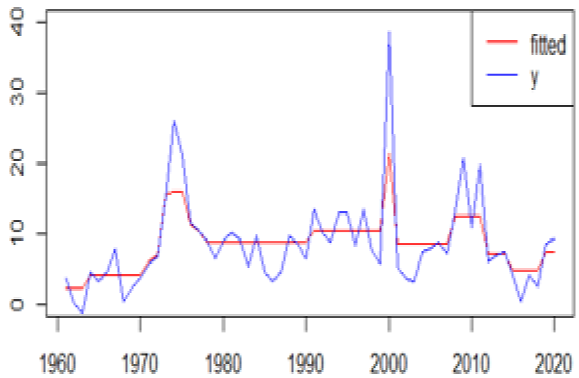
Regularization techniques with cross-validation tuning parameters possess a higher RMSE than fixed tuning parameters, as illustrated in table 5.4; see table 5.3 for comparison. Autometrics possess 3.74 RMSE with a 0.01 level higher than the RMSE of 0.05 significance level. The break detected via regularization techniques has down estimated break coefficients. Regularization techniques with cross-validation tuning parameters detect inconsistence breaks and possess higher RMSE than the fixed tuning parameter, figure 5.6.

The graphical overview of break detection in GDP deflator of each considered technique with cross-validation tuning parameter presented in figure 5.5. The graphic visualization indicates that regularization techniques with cross-validation underestimate the break coefficients. However, regularization techniques omit relevant breaks compared to Autometrics with 0.01 and 0.05 levels of significance and possess higher RMSE, figure 5.6.

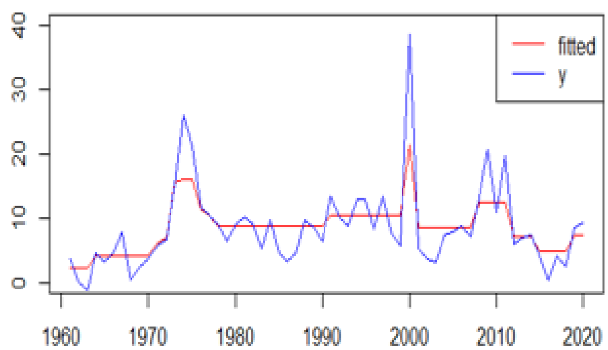
Autometrics (α=0.05)



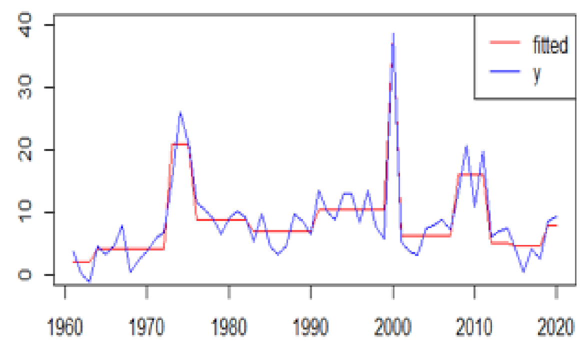

AdaLASSO Fixed




LASSO Fixed




SCAD Fixed




MCP Fixed

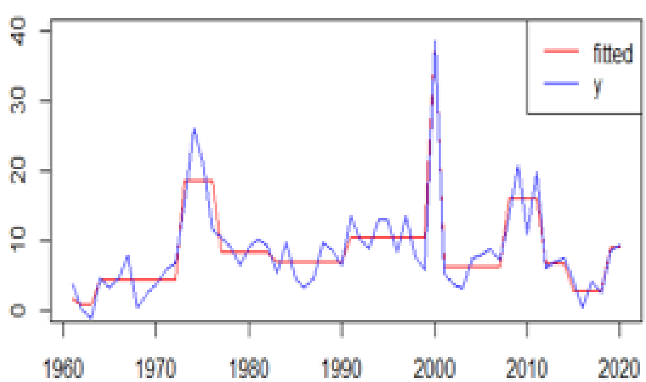



***Figure 5. 4:*** *Fitted GDP Deflator Plot under Different Regularization Techniques with Fixed Tuning Parameter* ***.***

**Table 5. 3:** GDP Deflator Break Detection via Regularization Techniques with Fixed Tuning Parameter

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Autometrics with 0.05 level of significance** | | | | | | | |
| | **Intercept** | **sis1972** | **sis1975** | **sis1999** | **sis2000** | **sis2007** | **sis2011** |
| **Coef** | 3.501 | 17.323 | -12.007 | 29.694 | -32.281 | 9.860 | -10.618 |
| **LASSO with Fixed tuning parameter** | | | | | | | |
| | **Intercept** | **sis1963** | **sis1970** | **sis1971** | **sis1972** | **sis1973** | **sis1975** |
| **Coef** | 2.195 | 2.021 | 1.938 | 0.822 | 8.539 | 0.462 | -4.639 |
| | **sis1976** | **sis1977** | **sis1978** | **sis1990** | **sis1999** | **sis2000** | **sis2007** |
| **Coef** | -0.983 | -1.512 | -0.161 | 1.686 | 11.066 | -12.929 | 4.129 |
| | **sis2011** | **sis2014** | **sis2018** | | | | |
| **Coef** | -5.507 | -2.221 | 2.448 | | | | |
| **AdaLASSO with Fixed tuning parameter** | | | | | | | |
| | **Intercept** | **sis1963** | **sis1970** | **sis1971** | **sis1972** | **sis1973** | **sis1975** |
| **Coef** | 2.195 | 2.020 | 1.938 | 0.822 | 8.539 | 0.462 | -4.639 |
| | **sis1976** | **sis1977** | **sis1978** | **sis1990** | **sis1999** | **sis2000** | **sis2007** |
| **Coef** | -0.983 | -1.512 | -0.161 | 1.686 | 11.066 | -12.929 | 4.129 |
| | **sis2011** | **sis2014** | **sis2018** | | | | |
| **Coef** | -5.507 | -2.221 | 2.448 | | | | |
| **SCAD with Fixed tuning parameter** | | | | | | | |
| | **Intercept** | **sis1963** | **sis1972** | **sis1975** | **sis1982** | **sis1990** | **sis1999** |
| **Coef** | 1.951 | 2.066 | 16.808 | -11.998 | -1.812 | 3.398 | 28.099 |
| | **sis2000** | **sis2007** | **sis2011** | **sis2014** | **sis2015** | **sis2018** | |
| **Coef** | -32.281 | 9.861 | -11.107 | -0.353 | -0.043 | 3.373 | |
| **MCP with Fixed tuning parameter** | | | | | | | |
| | **Intercept** | **sis1961** | **sis1963** | **sis1972** | **sis1976** | **sis1982** | **sis1990** |
| **Coef** | 1.646 | -0.835 | 3.493 | 14.191 | -10.104 | -1.384 | 3.405 |
| | **sis1999** | **sis2000** | **sis2007** | **sis2011** | **sis2014** | **sis2018** | |
| **Coef** | 28.103 | -32.284 | 9.861 | -9.309 | -4.036 | 6.221 | |

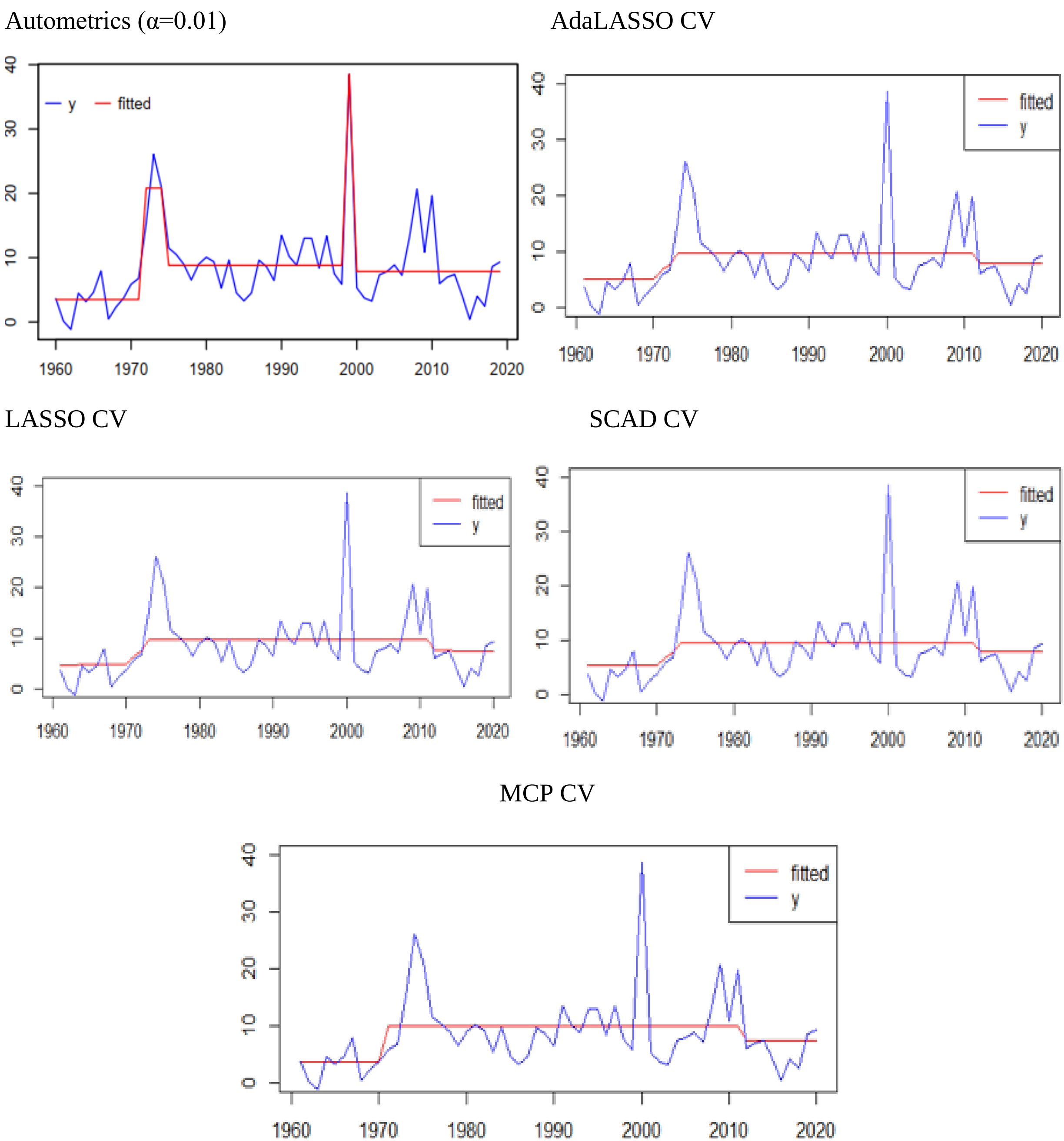



***Figure 5. 5:*** Fitted GDP Deflator Plot under Different Regularization Techniques with cross-validation.

**Table 5. 4:** GDP Deflator Break Detection via Regularization Techniques with cross-validation Tuning Parameter

| | | | | |
|---|---|---|---|---|
| **Autometrics with 0.01 level of significance** | | | | |
| | **Intercept** | **sis1972** | **sis1975** | **sis1999** |
| **Coef** | 3.501 | 17.324 | -12.007 | 29.694 |
| | **sis2000** | | | |
| **Coef** | -30.650 | | | |
| **LASSO with Cross-Validation tuning parameter** | | | | |
| | **Intercept** | **sis1963** | **sis1970** | **sis1971** |
| **Coef** | 3.877 | 0.847 | 1.698 | 0.927 |
| | **sis1972** | **sis2011** | **sis2014** | |
| **Coe** | 2.635 | -2.419 | -0.687 | |
| **AdaLASSO with Cross-Validation tuning parameter** | | | | |
| | **Intercept** | **sis1963** | **sis1970** | **sis1971** |
| **Coef** | 4.489 | 0.421 | 1.632 | 0.916 |
| | **sis1972** | **sis2011** | **sis2014** | |
| **Coef** | 2.362 | -2.096 | -0.426 | |
| **SCAD with Cross-Validation tuning parameter** | | | | |
| | **Intercept** | **sis1963** | **sis1970** | **sis1971** |
| **Coef** | 5.029 | 0.0476 | 1.596 | 0.852 |
| | **sis1972** | **sis2011** | **sis2014** | |
| **Coef** | 2.153 | -1.810 | -0.196 | |
| **MCP with Cross-Validation tuning parameter** | | | | |
| | **Intercept** | **sis1970** | **sis2011** | |
| **Coef** | 3.603 | 6.397 | -2.687 | |

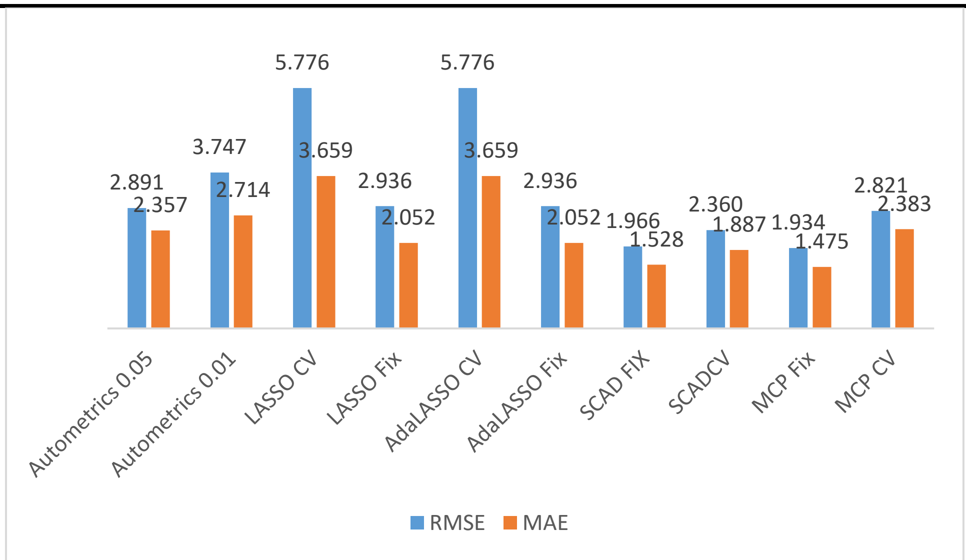



***Figure 5. 6:*** RMSE and MAE of GDP deflator

### 5.3. Real data analysis of IIS method for outlier detection in cross-sectional analysis

Coronavirus disease 2019 (COVID-19) is a global outbreak caused by coronavirus-2, which causes severe acute respiratory illness (SARS-CoV-2). The World Health Organization declared COVID-19 a pandemic in March 2020. As of December 1, 2021, COVID-19 had been found in over 2.614 billion people worldwide, with 5.2 billion deaths. On February 25, 2020, the first verified case of COVID-19 was reported in Pakistan. Pakistan, however, is not among the nations with the highest number of COVID-19 cases and fatalities. Up to December 1, 2021, 1.284 million COVID cases had been discovered, with 28,718 deaths.

Coronavirus pneumonia (COVID-19) is a worldwide health emergency because of its quick transmission and high death rate (Chatterjee et al., 2020). The clinical and physiological characteristics of SARS-COV-2, as well as diagnostic approaches, have been studied all over the world (Elshazli et al., 2020). During this pandemic, scientists and physicians face a global challenge in patient care and suitable treatment techniques, including creating an effective vaccine. Different diagnostic indicators have played a significant role in diagnosing and controlling the status of SARS-COV-2 patients (Y. Li et al., 2020). C-reactive protein (CRP) levels can be used as a biomarker to help diagnose pneumonia early, and individuals with severe lung infections have increased CRP levels (Stringer et al., 2021). Patients with COVID-19 have higher serum C-reactive protein (CRP) levels, which are used to help classify, diagnose, and prognostic the disease (Chen et al., 2020). This analysis aims to investigate the relationship between the length of hospital stay and CRP level, Gender, Age, Diabetes, Patient discharge status, and other comorbidities with permission of hospital authorities and consent of patient’s privacy. The data was gathered from Isolation Hospital and Infectious Treatment Center (IHITC) in Islamabad from July 2021 to 30 September 2021. A total of 275 patients agreed to participate in the study between July to

September. All the patients admitted they belonged to Rawalpindi and Islamabad regions. Figure 5.7 illustrates the correlation graph of considered variables; this indicates the positive correlation between a hospital stay and CRP level with a correlation equal to 0.2 and a negative correlation with other comorbidities with -0.1. However, patients' survival and age are positively associated with hospital stay with a correlation equal 0.2 and 0.1, respectively. Figure 5.8 illustrates the box plot of the hospital stay. It indicates that the minimum length of hospital stay equals 1 and the maximum 41, as the hospital stay is the dependent variable and contains an outlier, as shown in Figure 5.8. Furthermore, the residual plot of linear regression presented in figure 5.9 confirms outliers in model residuals. For the out-of-sample forecast, we randomly train the model on 90% of observations (233) and validate 10% of observations (26)(Franklin, 2005; James et al., 2013).

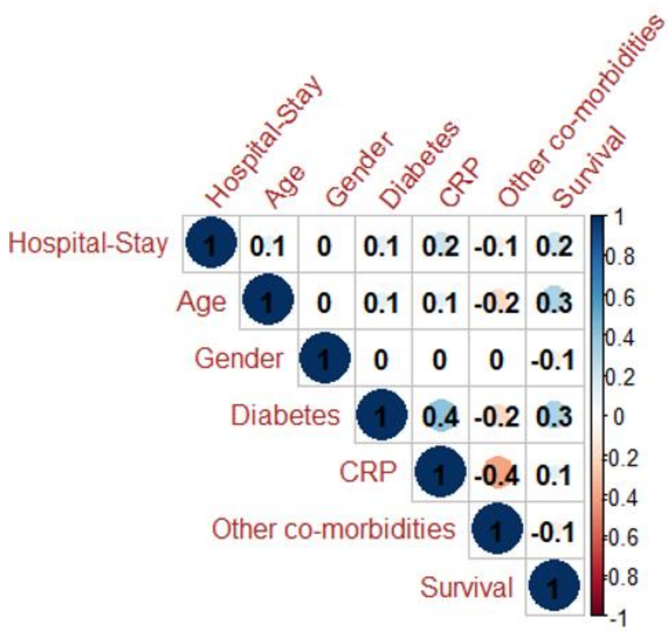



***Figure 5. 7:*** Correlation graph

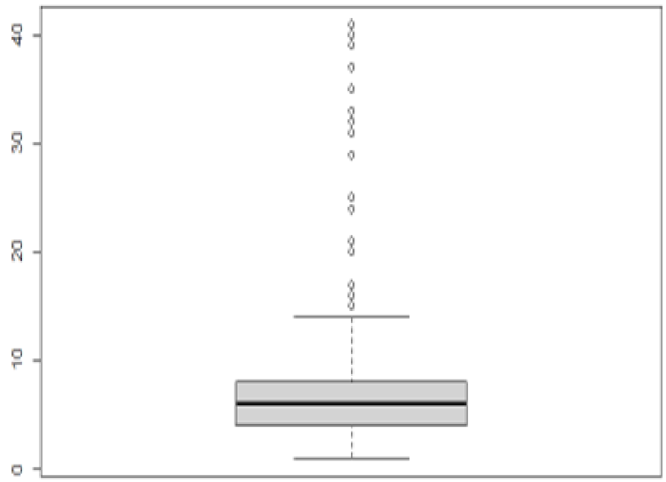



***Figure 5. 8:*** Box plot of hospital-stay

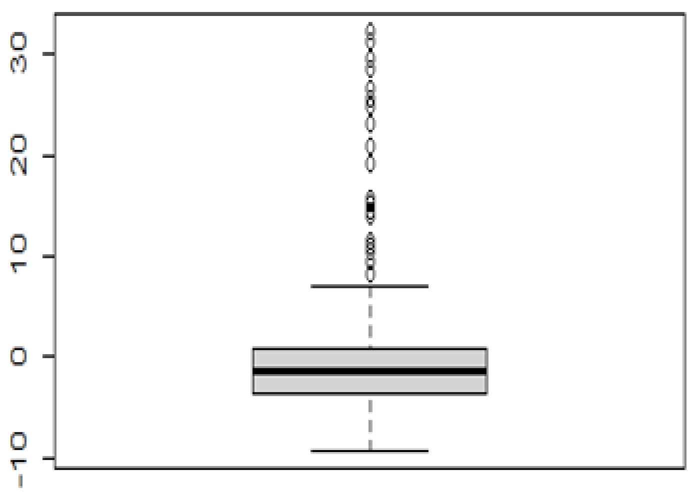



***Figure 5. 9*:** Residual box plot of linear regression

$$Hospital\ stay\ =\ \beta_0 +\ \beta_1 Gender +\ \beta_2 Age + \beta_3 Diabetes +\ \beta_4 CRP + \beta_5 Survival + \beta_6 Other\ Co - morbidities + \sum_{i=1}^{233} \gamma_i I_i +\ \varepsilon_i \quad (5.1)$$

We randomly train the model on 90% of observations (233) and validate 10% of observations (26). We report the RMSE of regularization techniques below figure 5.10.

**Table 5. 5:** Real data analysis with covariate selection and number of selected outliers

| | | | | |
|---|---|---|---|---|
| **SCAD Number of selected outliers (28)** | | | | |
| **Variable** | **Gender** | **CRP Level** | **Other comorbidities** | |
| **Coefficient** | 0.24463 | 0.00083 | 0.20533 | |
| **MCP Number of selected outliers (31)** | | | | |
| Variable | **Gender** | **CRP Level** | **Other comorbidities** | |
| Coefficient | 0.22493 | 0.0004 | 0.2585 | |
| **LASSO Number of selected outliers (204)** | | | | |
| **Variable** | **Age** | **Gender** | **CRP Level** | **Other comorbidities** |
| **Coefficient** | 0.00225 | 0.55747 | 0.00282 | 1.3966 |
| **Auto(0.05) Number of selected outliers (14)** | | | | |
| **Variable** | **CRP Level** | **Other comorbidities** | | |
| **Coefficient** | 0.00766 | 0.9653 | | |

The above table 5.5 indicates that SCAD and MCP perform similarly in covariate selection, as Gender, CRP level, and other comorbidities are significant variables which increases the length of hospital stay. However, SCAD selected 28 outliers, and MCP selected 31, slightly higher than SCAD. The real data analysis confirms that the LASSO estimates more covariates and outliers than other regularization techniques, which is aligned with our simulation findings. LASSO selects four covariates which are more than the covariates selected via SCAD and MCP. Autometrics with a 0.05 level of significance selects two covariates and 14 outliers. AdaLASSO and Autometrics with a 0.01 level of significance do not select any covariate, only retain outliers. Real data analysis indicates that Gender, CRP level, and other comorbidities are significant covariates. These indicator dummies can be interpreted as an observed heterogeneity of individuals, which prolonged hospital stay length.

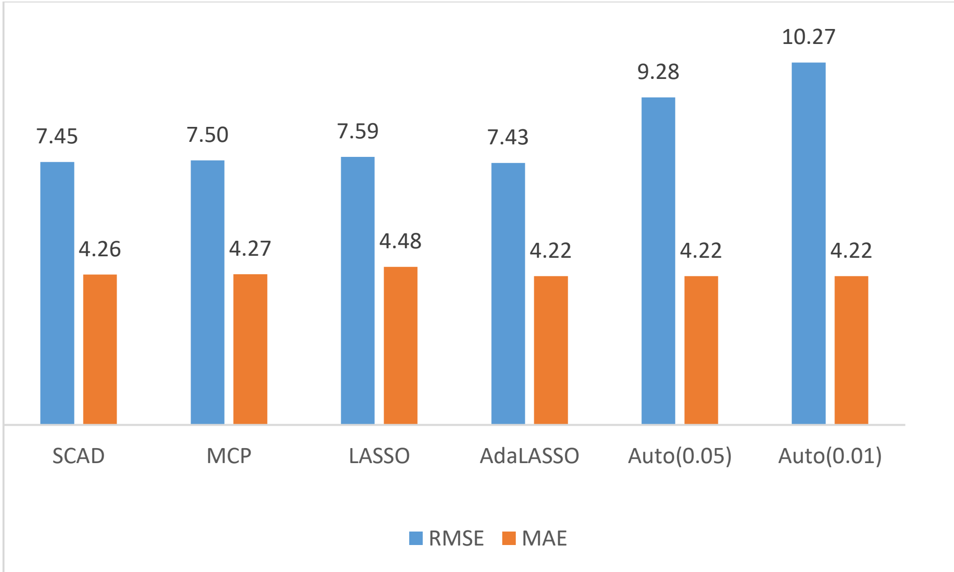



***Figure 5. 10:*** Out-of-sample RMSE of real data analysis

Figure 5.10 presented that SCAD and MCP outperform out-of-sample RMSE compared to all other techniques. As expected, the LASSO selected more indicator dummies and retained higher RMSE than other regularization techniques. Autometrics with a 0.01 level of significance retain the highest RMSE compared to all other techniques, not selecting covariates. Autometrics with tight significance levels omit relevant variables due to this RMSE increase (as observed from the simulation graph and table). In contrast, even Autometrics with a 0.05 level of significance possesses higher RMSE than regularization techniques.

### 5.4. Real data analysis for covariate and lag selection

For the real data analysis, we aim to probe the determinants of the trade balance for Pakistan and implement the considered techniques and assess their performance. Trade has played an important role in developing countries as a growth engine in various eras. The trade deficit or surplus is a term used to describe trade imbalances. Since independence, Pakistan has been in a trade deficit, except for three years: 1947-1948, 1950-1951, and 1972-1973 (Asif, 2014). According to economic literature, a variety of factors are thought to be responsible for long-term trade deficits in various economies, including ineffective public policies, shocks in major trading countries, oil price hikes if the economy is heavily reliant on oil imports, residents' socioeconomic conditions, and increased urbanization (Grupe & Rose, 2010; Manual & San, 2019). The existing studies in the case of Pakistan considered only a few macroeconomic variables as like GDP, exchange rate, broad money supply, inflation, and Foreign Direct Investment (Asif, 2014; Awan et al., 2011; Hussain & Muhammad, 2010; Kakar et al., 2010; Muhammad, 2010; Shahbaz et al., 2010). This study intakes the Generalized Unrestricted Model (GUM) that include each and every possible determinant of trade balance with 11 regressors namely Domestic Investment (log), Domestic Consumption(log), FDI(log), GDP(log), Inflation(log), Budget Deficit(log), Remittances(log), Exchange Rate(log), Population(log), Urban population(log), and Government expenditure(log).

We use annual frequency data from 1980 to 2020. The data has been compiled from World Data Indicator. The model contains 11 regressors (with a difference) and includes 5 lags of each covariate and the lags of the dependent variable. The GUM includes 71 covariates; due to differencing data and 5 lags of covariates, we have 35 observations. We train the model on 30 observations from 1985-2015 as the last 5 observations from 2016-2020 have been discarded for test data. Throughout the simulation experiment and real data analysis, we use BIC-based tuning parameters for regularization techniques, while for Autometrics, we select the model with 0.01 and 0.05 significance levels.

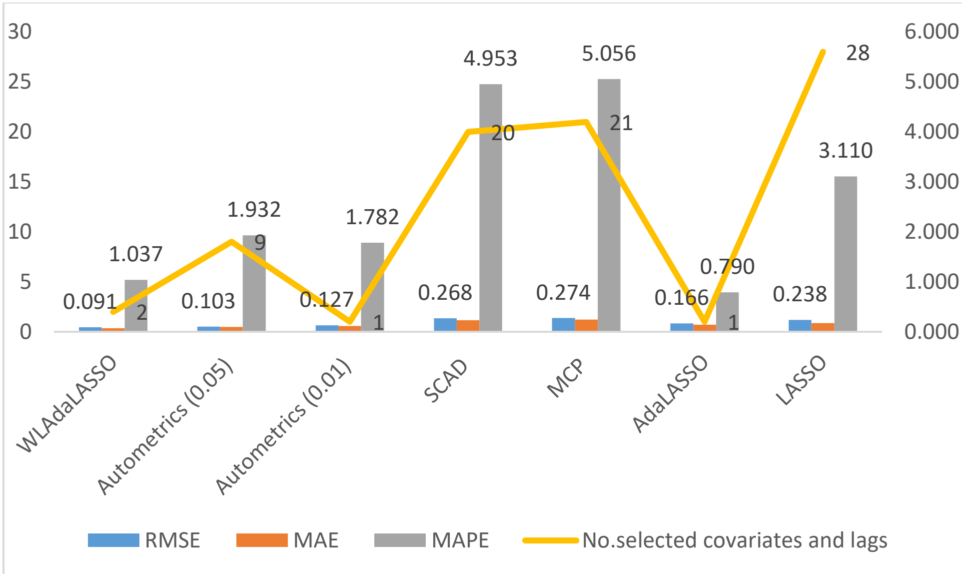



***Figure 5. 11:*** Real data analysis

The real data analysis illustrated in figure 5.11 verifies our simulation findings as WLAdaLASSO outperforms all other techniques with the least out-of-sample RMSE equal to 0.091, followed by Autometrics (0.01) with RMSE of 0.127. Autometrics with 0.05 possesses a higher RMSE equal to 0.103 than Autometrics with a 0.01 significance level. The finding is aligned with the simulation experiment as Autometrics with a 0.05 level of significance possesses a slightly higher average gauge with a higher RMSE than a 0.01 level of significance. SCAD, MCP, and LASSO estimate higher RMSE as the model selects more irrelevant covariates and lag than WLAdaLASSO and Autometrics. WLAdaLASSO selects three covariates, namely dupop (Difference of urban population), dlnGDP(-1) (Difference of log GDP lag 1), and dpop(-4) (Difference of log population lag 4). Autometrics with a 0.05 significance level selects five covariates and their lag, and with a 0.01 significance level, select three covariates. dlnGDP(-1) is a common covariate between WLAdaLASSO and Autometrics with 0.05 and 0.01 significance levels. SCAD, MCP, and LASSO select too many covariates and their lag, due to which these techniques possess higher

RMSE than WLAdaLASSO. However, AdaLASSO selects one covariate with RMSE is equal to 0.166, which is higher than WLAdaLASSO and Autometrics.

It is obvious that macroeconomic variables are correlated; the break detected in this study provides empirical evidence that a break in one variable impact other economic variables simultaneously or over time. Eventually, GDP growth, inflation, and unemployment rates possess common breaks, especially in 1970, 1971, and 2008. The break detected from the considered series is equitable; subsequently, the breaks correspond to significant economic events, the Oil crisis 1970, the Pakistan-India war of 1971, and the global financial crises. With empirical evidence, it can be inferred that international uncertainty, like the Oil crises in 1970 and the global financial crises, spontaneously decreases GDP growth, rising inflation, and increasing interest rates. On the other hand, the unemployment increment in 2006 was due to the economic slowdown after 2005. However, instability in GDP growth is because of international and national political uncertainties, which consequently impact the unemployment rate, inflation rate, and interest rate. The empirical result of break detection suggests political solidity to endorse a strong investment climate for national and international investors; extraordinary levels of human capital investment are needed to achieve sustainable development. Reducing dependency on crude oil can reduce import bills, as other events like the oil crisis would not impact the economy in the future. The break detected via the SIS method indicates that the rigid fiscal and monetary policy and significant structural changes were chosen as the principal policy instruments to attain these goals. Researchers and data analysts can adopt the SIS approach to arrive at valid results, leading to better policymaking and forecasting results. The empirical result from outlier detection via the IIS method estimated via SCAD and MCP possesses the least RMSE and MAE. This empirical analysis enhances the quality of estimation techniques for covariate selection and forecasting methods in cross-section analysis with multiple outliers without discarding extreme values from the model.

# Chapter 6

## Conclusion, Limitations, and Future Direction

Structural breaks in time series modeling, if go unreported, lead to parameter instability and poor out-of-sample forecast. SIS outperforms all other methods for structural break identification within the heart of current techniques, since it does not limit the break length, pre-specified number of breaks, break timing and breaks at the end or at the start of observations. SIS inherently uses Autometrics for break detection, while Autometrics is sensitive to the pre-specified significance level; with a nominal significance level, it selects more irrelevant breaks, and with a tight significance level (0.001) it omits relevant breaks. However, selecting relevant breaks is a crucial step; for this purpose, our first objective of this study is to compare different regularization techniques, SCAD, MCP, and AdaLASSO, for structural break detection. Comparison based on different scenarios, involving a break at the end of observations, a break at the first half of observations with different magnitudes, and multiple unknown breaks with different magnitudes. Methods are assessed on gauge, potency, in-sample RMSE, and MAE.

To control the shrinkage of regularization techniques, we use fixed tuning parameters and cross-validation tuning parameters. The simulation result indicates that all the considered methods performed well for break detection at the end of observation as the average potency approaches 1, average gauge approaches 0, and simultaneously retained the least average RMSE. All the methods found it easy to detect a single break at the end of observations. However, with a break in the first half of observations, the performance of regularization techniques decreases in average potency compared to the end of observations. SCAD with fixed tuning parameters performs identically to Autometrics in Potency and MAE. However, regularization techniques with cross-validation tuning parameters retain a higher average gauge than fixed tuning parameters. The gauge and potency of all methods are enhanced in terms of least gauge and higher potency as the magnitude shift equals 4.

Meanwhile, LASSO and AdaLASSO find it hard to detect the breaks with two multiple unknown shifts compared to SCAD and MCP. The LASSO and AdaLASSO with cross-validation possess the highest average potency 88.9%, at the cost of the higher average gauge of 10.3%. SCAD with fixed tuning parameters performs better than all other techniques, with the highest average potency of 73.4% and an average gauge of 1%. The overall simulation result indicates that among regularization techniques, SCAD and MCP with fixed tuning parameters perform better than LASSO and AdaLASSO. However, in terms of average potency, SCAD with fixed tuning parameter performs better than Autometrics in average potency retention.

For the empirical application of structural break detection, we use GDP growth and the GDP deflator of Pakistan. The empirical analysis indicates that Autometrics with a 0.01 level of significance omits the relevant break, specifically the break at the end of observation. However, regularization techniques detect the break efficiently. The real data analysis indicates tuning parameter SCAD and MCP perform near Autometrics. However, regularization techniques with cross-validation tuning parameters possess downward bias estimates and higher RMSE for break detection. Regularization techniques with fixed tuning parameters possess the least RMSE compared to the RMSE of cross-validation. Overall, LASSO and AdaLASSO contain higher RMSE and more irrelevant breaks than SCAD and MCP. The SCAD and MCP with fixed tuning parameters are close to Autometrics in real data analysis with the least RMSE and MAE. Overall, the empirical analysis is aligned with the simulation finding.

The second objective of this study is based on outlier detection in AR(1) series and multivariate static model. IIS is a well-known method for outlier detection; for this purpose, we use regularization techniques to estimate the model in the presence of outliers using the IIS method and compare its efficiency with Autometrics.

Overall analysis indicates that regularization techniques outperform than Autometrics in simulation study for covariate selection and out-of-sample forecasting. However, the IIS method estimated via SCAD and MCP retains the least (gauge, RMSE, and MAE) and high potency among other regularization techniques. Regularization techniques with DGP consisting of 5% outlying observations with 4 SD magnitude possess a higher average potency than DGP consisting of 5% outlying observations with 6 SD magnitude. Conversely, the DGP consisting of 5% outlying observations with 4 SD magnitude regularization technique possesses a higher average gauge than the DGP consisting of 5% outlying observations with 6 SD magnitude. The overall simulation experiment indicates that the higher magnitude of an outlier like 6 SD diminishes the average potency, increasing RMSE. Throughout the simulation experiment, LASSO, and AdaLASSO possess higher average gauge than all other considered regularization techniques.

Meanwhile, the AR(1) simulation results indicate that the BIC tuning parameter performs better than the fixed tuning parameter in all considered parameters. However, as the outlier's magnitude decreases to $\gamma$ equal 3 the average potency of overall techniques decreases compared to $\gamma$ equal to 5. Compared to all other techniques, MCP with BIC tuning parameter retains the highest average potency of 73.3% with the least RMSE (2.518) and MAE (2.301).

For the empirical analysis of outlier detection via the IIS method, we use COVID-19 hospitalized patients in Islamabad. The analysis confirms the simulation findings as the LASSO estimates more outliers and covariates than other regularization techniques. While the SCAD and MCP possess a minimum out-of-sample RMSE than Autometrics and LASSO. The real data analysis indicates that SCAD and MCP select three covariates, Gender, CRP Level, and other comorbidities, as indicators of the length of hospital stay and possess the least RMSE. Our study proves that the IIS method for outlier detection and covariate selection estimated via SCAD and MCP gives more precise results than Autometrics in orthogonal covariates and outlier presences.

The third objective of this study is based on dynamic time series modeling; however, for covariate and its lag selection. The use of regularization techniques in time series modeling has been prevalent in recent years due to the availability of massive data. We analyze the performance of the WLAdaLASSO with Autometrics for covariate selection and forecasting. The simulation study illustrates that the WLAdaLASSO, with the stronger linear dependency between predictors outperforms Autometrics and other regularization techniques. However, Autometrics with $\phi$ being equal 0.1, the performance of gauge approaches to $\alpha$ (0.05 or 0.01 level of significance), potency approaches to 1, and the Average RMSE also decrease, with sample size increment. On the contrary, the situation is limited to $\phi$ equal to 0.1; however, $\phi$ equal to 0.8, and increasing sample size does not significantly enhance the performance of Autometrics compared to WLAdaLASSO. Autometrics with a 0.05 significances level include irreverent covariates that increase the RMSE compared to 0.01 significance, and the finding is aligned with real data analysis. However, other than the WLAdaLASSO, all considered regularization techniques perform poorly in covariate selection and forecasting even with $\phi$ being equal to 0.1 and $T$ equal to 50, whereas the performance of considered techniques is improved with an increase in sample size; still, WLAdaLASSO outperformed others among all simulation experiments.

The simulation experiment and real data analysis are evidence that the WLAdaLASSO is a more robust technique than all other considered regularization techniques and Autometrics as well in out-of-sample forecasting and covariate selection even with the stronger linear dependence between predictors and small sample size.

### 6.1. Study Limitation and Future Study direction

The study has a few limitations, considering only linear models for dynamic time series analysis. However, the study is limited to orthogonal covariates for outlier detection via the IIS method. The future study can be developed to examine the performance of modern statistical and machine learning methods combined with the IIS approach in panel data. Additionally, it is possible to compare these tools in forecasting and variable selection in panel data ARDL models while considering the lagged variables. On the other hand, the study can be expanded to analyze the performance of neural net, random forest with SCAD, and MCP with the IIS approach in terms of predicting.

**References**

Andrews, D. W. K., & Fair, R. C. (1988). Inference in nonlinear econometric models with structural change. *The Review of Economic Studies*, *55*(4), 615–640.

Andrews, D. W. K. (1993). Tests for Parameter Instability and Structural Change With Unknown Change Point. *Econometrica*, *61*(4), 821. https://doi.org/10.2307/2951764

Andrews, D. W. K., Ploberger, W., Andrews, D., & Ploberger, W. (1994). Optimal Tests When a Nuisance Parameter Is Present Only under the Alternative. *Econometrica*, *62*(6), 1383–1414. https://econpapers.repec.org/RePEc:ecm:emetrp:v:62:y:1994:i:6:p:1383-1414

Anjum, M. I., & Sgro, P. M. (2017). A brief history of Pakistan's economic development. *Real-World Economics Review, 80*, 171–178.

Asif, K. (2014). Determinants of Trade Balance: A Comparison between Pakistan and India. *IBA Business Review*, *9*(1), 1-136.

Audrino, F., & Camponovo, L. (2013). Oracle Properties and Finite Sample Inference of the Adaptive Lasso for Time Series Regression Models. *SSRN Electronic Journal, 1-29*. https://arxiv.org/abs/1312.1473v1

Awan, A., Asghar, N., & Rehman, H. U. (2011). The impact of exchange rate, fiscal deficit and terms of trade on external debt of Pakistan. *Australian Journal of Business and Management Research*, *1*(3), 10, 10-24.

Banerjee, A., Lumsdaine, R. L., & Stock, J. H. (1992). Recursive and sequential tests of the unit-root and trend-break hypotheses: theory and international evidence. *Journal of Business & Economic Statistics*, *10*(3), 271-287.

Bai, J. (1994). Least squares estimation of a shift in linear processes. *Journal of Time Series Analysis*, *15*(5), 453–472.

Bai, J. (1997). Estimation of a change point in multiple regression models. *Review of Economics and Statistics*, *79*(4), 551–560. https://doi.org/10.1162/003465397557132

Bai, J., Lumsdaine, R. L., & Stock, J. H. (1998). Testing for and dating common breaks in multivariate time series. *The Review of Economic Studies*, *65*(3), 395-432.

Bai, J., & Perron, P. (1998). Estimating and Testing Linear Models with Multiple Structural Changes. *Econometrica*, *66*(1), 47. https://doi.org/10.2307/2998540

Bai, J., & Ng, S. (2002). Determining the Number of Factors in Approximate Factor Models. *Econometrica*, *70*(1), 191–221. https://doi.org/10.1111/1468-0262.00273

Bai, J., & Perron, P. (2003a). Computation and analysis of multiple structural change models. *Journal of Applied Econometrics*, *18*(1), 1–22. https://doi.org/10.1002/jae.659

Bai, J., & Perron, P. (2003b). Computation and analysis of multiple structural change models. *Journal of Applied Econometrics*, *18*(1), 1–22. https://doi.org/10.1002/jae.659

Bai, J., & Perron, P. (2003c). Critical values for multiple structural change tests. *The Econometrics Journal*, *6*(1), 72–78. https://doi.org/10.1111/1368-423x.00102

Bai, J., & Perron, P. (2006). Multiple structural change models: a simulation analysis. *Econometric Theory and Practice: Frontiers of Analysis and Applied Research*, *1*, 212–237.

Berk, R. (1990). *A primer on robust regression*.

Ben-David, D., & Papell, D. H. (1998). Slowdowns and meltdowns: postwar growth evidence from 74 countries. *Review of Economics and Statistics*, *80*(4), 561-571.

Birkes, D., & Dodge, Y. (2011). *Alternative methods of regression* (Vol. 190). John Wiley & Sons.

Bondell, H. D., & Reich, B. J. (2008). Simultaneous regression shrinkage, variable selection, and supervised clustering of predictors with OSCAR. *Biometrics*, *64*(1), 115–123.

Box, G. E. P. (1979). Robustness in the Strategy of Scientific Model Building. *Robustness in Statistics*, 201–236. https://doi.org/10.1016/B978-0-12-438150-6.50018-2

Breusch, T. S., & Pagan, A. R. (1980). The Lagrange Multiplier Test and its Applications to Model Specification in Econometrics. *The Review of Economic Studies*, *47*(1), 239. https://doi.org/10.2307/2297111

Brown, R. L., Durbin, J., & Evans, J. M. (1974). *Techniques for Testing the Constancy of Regression Relationships over Time*. https://doi.org/10.1111/j.2517-6161.1975.tb01532.x

Bühlmann, P., & Van De Geer, S. (2011). *Statistics for high-dimensional data: methods, theory and applications*. Springer Science & Business Media.

Bühlmann, P., Rütimann, P., van de Geer, S., & Zhang, C.-H. (2013). Correlated variables in regression: clustering and sparse estimation. *Journal of Statistical Planning and Inference*, *143*(11), 1835–1858.

Cantrell, R. S., Burrows, P. M., & Vuong, Q. H. (1991). Interpretation and use of generalized Chow tests. *International Economic Review*, 725–741.

Castle, J. L., Doornik, J. A., & Hendry, D. F. (2012). Model selection when there are multiple breaks. *Journal of Econometrics*, *169*(2), 239–246. https://doi.org/10.1016/j.jeconom.2012.01.026

Castle, J. L., Clements, M. P., & Hendry, D. F. (2013). Forecasting by factors, by variables, by both or neither? *Journal of Econometrics*, *177*(2), 305–319.

Castle, J. L., & Hendry, D. F. (2014). Model selection in under-specified equations facing breaks. *Journal of Econometrics*, *178*, 286–293.

Castle, J. L., Doornik, J. A., Hendry, D. F., & Pretis, F. (2015a). Detecting location shifts during model selection by step-indicator saturation. *Econometrics*, *3*(2), 240–264. https://doi.org/10.3390/econometrics3020240

Castle, J. L., Doornik, J. A., Hendry, D. F., & Pretis, F. (2015b). Detecting location shifts during model selection by step-indicator saturation. *Econometrics*, *3*(2), 240–264. https://doi.org/10.3390/econometrics3020240

Castle, J. L., & Hendry, D. F. (2019a). Key Concepts: A Series of Primers. *Modelling Our Changing World*, 5–36. https://doi.org/10.1007/978-3-030-21432-6_2

Castle, J. L., & Hendry, D. F. (2019b). Modelling our Changing World. *Modelling Our Changing World*, 67–84. https://doi.org/10.1007/978-3-030-21432-6.

Castle, J. L., Doornik, J. A., & Hendry, D. F. (2021). Modelling non-stationary 'big data.' *International Journal of Forecasting*, *37*(4), 1556–1575.

Castle, J. L., Doornik, J. A., & Hendry, D. F. (2021). Robust discovery of regression models. *Econometrics and Statistics*, Volume 26, *31-51*.

Christiano, L. J. (1992). Searching for a Break in GNP. *Journal of Business & Economic Statistics*, *10*(3), 237-250.

Chatterjee, P., Nagi, N., Agarwal, A., Das, B., Banerjee, S., Sarkar, S., Gupta, N., & Gangakhedkar, R. R. (2020). The 2019 novel coronavirus disease (COVID-19) pandemic: A review of the current evidence. *The Indian Journal of Medical Research*, *151*(2–3), 147.

Chen, N., Zhou, M., Dong, X., Qu, J., Gong, F., Han, Y., Qiu, Y., Wang, J., Liu, Y., & Wei, Y. (2020). Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *The Lancet*, *395*(10223), 507–513.

Chong, T. T. L. (1995). Partial parameter consistency in a misspecified structural change model. *Economics Letters*, *49*(4), 351-357.

Chow, G. C. (1960). Tests of Equality Between Sets of Coefficients in Two Linear Regressions. *Econometrica*, *28*(3), 591. https://doi.org/10.2307/1910133

Campos, J., Hendry, D. F., and Krolzig, H.-M. (2003). Consistent model selection by an automatic Gets approach. *Oxford Bulletin of Economics and Statistics, 65*, 803–819.

Craven, P., & Wahba, G. (1978). Smoothing noisy data with spline functions. *Numerische Mathematik 1979 31:4*, *31*(4), 377–403. https://doi.org/10.1007/BF01404567

Cui, X., Hu, J., Ma, Y., Wu, P., Zhu, P., & Li, H.-J. (2021). Investigation of stock price network based on time series analysis and complex network. *International Journal of Modern Physics B*, *35*(13), 2150171, 1-18.

De Mol, C., Giannone, D., Reichlin, L., De Mol, C., Giannone, D., & Reichlin, L. (2008). Forecasting using a large number of predictors: Is Bayesian shrinkage a valid alternative to principal components? *Journal of Econometrics*, *146*(2), 318–328. https://econpapers.repec.org/RePEc:eee:econom:v:146:y:2008:i:2:p:318-328

Doornik, J. A. (2009). *Econometric model selection with more variables than observations*. Working paper, Economics Department, University of Oxford, 1-33..

Doornik, J. A., & Hendry, D. F. (2015). Statistical model selection with “Big Data” *Cogent Economics & Finance*, *3*(1), 1045216, 1-15.

Dufour, J.-M. (1982). Generalized Chow tests for structural change: A coordinate-free approach. *International Economic Review*, 565–575.

Dufour, J.-M., Ghysels, E., & Hall, A. (1994). Generalized predictive tests and structural change analysis in econometrics. *International Economic Review*, 199–229.

Efron, B., Hastie, T., Johnstone, I., & Tibshirani, R. (2004). Least angle regression. *32*(2), 407–499. https://doi.org/10.1214/009053604000000067

Elshazli, R. M., Toraih, E. A., Elgaml, A., El-Mowafy, M., El-Mesery, M., Amin, M. N., Hussein, M. H., Killackey, M. T., Fawzy, M. S., & Kandil, E. (2020). Diagnostic and prognostic value of hematological and immunological markers in COVID-19 infection: A meta-analysis of 6320 patients. *PLOS ONE*, *15*(8), e0238160. https://doi.org/10.1371/JOURNAL.PONE.0238160

Engle, R. F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, *50*(4), 987. https://doi.org/10.2307/1912773

Epprecht, C., Guégan, D., Veiga, Á., & Correa da Rosa, J. (2021). Variable selection and forecasting via automated methods for linear models: LASSO/adaLASSO and Autometrics. *Communications in Statistics: Simulation and Computation*, *50*(1), 103–122. https://doi.org/10.1080/03610918.2018.1554104

Fan, J., & Li, R. (2001). Variable selection via nonconcave penalized likelihood and its oracle properties. *Journal of the American Statistical Association*, *96*(456), 1348–1360.

https://doi.org/10.1198/016214501753382273

Farley, J. U., & Hinich, M. J. (1970). A test for a shifting slope coefficient in a linear model. *Journal of the American Statistical Association*, *65*(331), 1320–1329.

Fernandez, D. G. (1997). Breaking trends and the money-output correlation. *Review of economics and statistics*, *79*(4), 674-679.

Fisher, I. (1930). *Theory of interest: as determined by impatience to spend income and opportunity to invest it*. Augustusm Kelly Publishers, Clifton, 29-31.

Franklin, J. (2005). The elements of statistical learning: data mining, inference and prediction. *The Mathematical Intelligencer*, *27*(2), 83–85.

Godfrey, L. G. (1978). Testing for Higher Order Serial Correlation in Regression Equations when the Regressors Include Lagged Dependent Variables. *Econometrica*, *46*(6), 1303. https://doi.org/10.2307/1913830

Grupe, C., & Rose, A. (2010). China, India, and the Socioeconomic Determinants of Their Competitiveness. *Economics Research International*, *2010*, 860425. https://doi.org/10.1155/2010/860425

Hansen, B. E. (1992). Testing for parameter instability in linear models. *Journal of Policy Modeling*, *14*(4), 517–533. https://doi.org/10.1016/0161-8938(92)90019-9

Hansen, B. E. (1997). Approximate asymptotic p values for structuras-change tests. *Journal of Business & Economic Statistics*, *15*(1), 60-67.

Hansen, B. E. (2001). The new econometrics of structural change: Dating breaks in U.S. labor productivity. *Journal of Economic Perspectives*, *15*(4), 117–128.

https://doi.org/10.1257/jep.15.4.117

Hansen, B. E. (2012). Tests for Parameter Instability in Regressions with 1(1) Processes. *Http://Dx.Doi.Org/10.1080/07350015.1992.10509908*, *10*(3), 321–335. https://doi.org/10.1080/07350015.1992.10509908

Hasan, P., Kemal, A. R., & Naseem, S. M. (1997). Learning from the past: A fifty-year perspective on Pakistan's development [with comments]. *The Pakistan Development Review*, *36*(4), 355–402.

Hauskrecht, M., Batal, I., Hong, C., Nguyen, Q., Cooper, G. F., Visweswaran, S., & Clermont, G. (2016). Outlier-based detection of unusual patient-management actions: An ICU study. *Journal of Biomedical Informatics*, *64*, 211–221. https://doi.org/10.1016/j.jbi.2016.10.002

Hendry, D. F., & Krolzig, H. M. (2004). We ran one regression. In *Oxford Bulletin of Economics and Statistics* (Vol. 66, Issue 5, pp. 799–810). John Wiley & Sons, Ltd. https://doi.org/10.1111/j.1468-0084.2004.102_1.x

Hendry, D. F., Johansen, S., & Santos, C. (2006). *Selecting a Regression Saturated by Indicators*, *1-17*.

Hoerl, A. E., & Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, *12*(1), 55–67.

Hoover, K. D., Perez, S. J., & ~kdhoover, /. (1999). Data mining reconsidered: encompassing and the general-to-specific approach to specification search. *Econometrics Journal*, *2*, 167–191.

Hu, J., Xia, C., Li, H., Zhu, P., & Xiong, W. (2020). Properties and structural analyses of USA's regional electricity market: A visibility graph network approach. *Applied Mathematics and*

*Computation*, *385*, 125434.

Hua, Y. (2011). *Macroeconomic Forecasting using Large Vector Auto Regressive Model*, 1-49.

Husain, I. (2000). Pakistan: The economy of an elitist state. *OUP Catalogue* Oxford University Press, number 9780195793796, 1-32.

Hussain, A., & Muhammad, S. D. (2010). The role of exchange rate on balance of trade: Empirical from Pakistan. *European Journal of Social Sciences, 14*(1), 150.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning* (Vol. 112). Springer.

Jansen, E. S., & Teräsvirta, T. (1996). TESTING PARAMETER CONSTANCY AND SUPER EXOGENEITY IN ECONOMETRIC EQUATIONS. *Oxford Bulletin of Economics and Statistics*, *58*(4), 735–763. https://doi.org/10.1111/J.1468-0084.1996.MP58004008.X

Jarque, C. M., & Bera, A. K. (1980). Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics Letters*, *6*(3), 255–259. https://doi.org/10.1016/0165-1765(80)90024-5

Jenkinson, G., Li, Y. I., Basu, S., Cousin, M. A., Oliver, G. R., & Klee, E. W. (2020). LeafCutterMD: an algorithm for outlier splicing detection in rare diseases. *Bioinformatics*, *36*(17), 4609–4615.

Johansen, S., & Nielsen, B. (2009). An analysis of the indicator saturation estimator as a robust regression estimator. *Castle, and Shephard (2009)*, *1*, 1–36.

Jyothi, P. N., Lakshmi, D. R., & Rama Rao, K. V. S. N. (2020). A supervised approach for detection of outliers in healthcare claims data. *Journal of Engineering Science and*

*Technology Review*, *13*(1), 204–213. https://doi.org/10.25103/jestr.131.25

Kakar, M. K., Kakar, R., & Khan, W. (2010). The Determinants of Pakistan's Trade Balance: An ARDL Cointegration Approach. *Lahore Journal of Economics*, *15*(1).

Khan, M. H. (2002). *When is Economic Growth Pro-poor?: Experiences in Malaysia and Pakistan* (Vol. 2). International monetary fund, 1-43.

Kim, H.-J., & Siegmund, D. (1989). The likelihood ratio test for a change-point in simple linear regression. *Biometrika*, *76*(3), 409–423.

Konzen, E., & Ziegelmann, F. A. (2016). LASSO-Type Penalties for Covariate Selection and Forecasting in Time Series. *Journal of Forecasting*, *35*(7), 592–612. https://doi.org/10.1002/for.2403

Krämer, W., Ploberger, W., & Alt, R. (1988). Testing for structural change in dynamic models. *Econometrica: Journal of the Econometric Society*, 1355–1369.

Krolzig, H. M., & Hendry, D. F. (2001). Computer automation of general-to-specific model selection procedures. *Journal of Economic Dynamics and Control*, *25*(6–7), 831–866. https://doi.org/10.1016/S0165-1889(00)00058-0

KULLBACK, S., & ROSENBLATT, H. M. (1957). ON THE ANALYSIS OF MULTIPLE REGRESSION IN k CATEGORIES. *Biometrika*, *44*(1–2), 67–83. https://doi.org/10.1093/BIOMET/44.1-2.67

Langford, I. H., & Lewis, T. (1998). Outliers in multilevel data. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *161*(2), 121–160.

Lee, S. (2015). An additive sparse penalty for variable selection in high-dimensional linear regression model. *Communications for Statistical Applications and Methods*, *22*(2), 147–157.

Li, J. (2012). Monetary Policy Analysis Based on Lasso-Assisted Vector Autoregression (Lavar). *SSRN Electronic Journal*, *1-30*. https://doi.org/10.2139/SSRN.2017877

Li, J., & Chen, W. (2014). Forecasting macroeconomic time series: LASSO-based approaches and their forecast combinations with dynamic factor models. *International Journal of Forecasting*, *30*(4), 996–1015. https://doi.org/10.1016/J.IJFORECAST.2014.03.016

Li, Y., Zhou, W., Yang, L., & You, R. (2020). Physiological and pathological regulation of ACE2, the SARS-CoV-2 receptor. *Pharmacological Research*, *157*, 104833. https://doi.org/10.1016/J.PHRS.2020.104833

Lo, A. W., & Newey, W. K. (1985). A large-sample Chow test for the linear simultaneous equation. *Economics Letters*, *18*(4), 351–353.

Lumsdaine, R. L., & Papell, D. H. (1997). Multiple trend breaks and the unit-root hypothesis. *Review of economics and Statistics*, *79*(2), 212-218.

Maddala, G. S., & Kim, I.-M. (1998). *Unit roots, cointegration, and structural change*.

Manual, V., & San, W. (2019). Dynamic relationship between trade balance and macroeconomic elements: Empirical evidence from emerging economies in Malaysia. *International Journal of Recent Technology and Engineering*, *7*(5), 47–54.

Marsilli, C. (2014). Variable Selection in Predictive MIDAS Models. *SSRN Electronic Journal*. https://doi.org/10.2139/SSRN.2531339

McConnell, M. M., & Perez-Quiros, G. (2000). Output fluctuations in the United States: What has changed since the early 1980's?. *American Economic Review*, *90*(5), 1464-1476.

Medeiros, M. C., & Mendes, E. F. (2012). *Estimating High-Dimensional Time Series Models, 1-34.* https://www.econstor.eu/handle/10419/176085

Muhammad, S. D. (2010). Determinant of balance of trade: Case study of Pakistan. *European Journal of Scientific Research*, *41*(1), 13–20.

Muthuramu, P., & Uma Maheswari, T. (2019). Tests for Structural Breaks in Time Series Analysis: A Review of Recent Development. *Shanlax International Journal of Economics*, *7*(4), 66–79. https://doi.org/10.34293/economics.v7i4.628

Nelson, C. R., & Plosser, C. R. (1982). Trends and random walks in macroeconmic time series: some evidence and implications. *Journal of monetary economics*, *10*(2), 139-162.

Nicholson, W. B., Matteson, D. S., & Bien, J. (2017). VARX-L: Structured regularization for large vector autoregressions with exogenous variables. *International Journal of Forecasting*, *33*(3), 627–651. https://doi.org/10.1016/J.IJFORECAST.2017.01.003

Noma, H., Gosho, M., Ishii, R., Oba, K., & Furukawa, T. A. (2020). Outlier detection and influence diagnostics in network meta-analysis. *Research Synthesis Methods*, *11*(6), 891–902.

Park, S. Bin. (1991). The Wald and LM Testing for Structural Change in a Linear Simultaneous Equation Model. *Korean Economic Review*, *7*(2), 29–47.

Park, H., & Sakaori, F. (2013). Lag weighted lasso for time series model. *Computational Statistics*, *28*(2), 493–504. https://doi.org/10.1007/s00180-012-0313-5.

Papell, D. H., Murray, C. J., & Ghiblawi, H. (2000). The structure of unemployment. *Review of Economics and Statistics*, *82*(2), 309-315.

Perron, P. (1989). The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis.

*Econometrica*, *57*(6), 1361. https://doi.org/10.2307/1913712

Perron, P. (2006). Dealing with structural breaks. *Palgrave Handbook of Econometrics*, *1*(2), 278–352.

Perron, P., & Vogelsang, T. J. (1992). Testing for a unit root in a time series with a changing mean: corrections and extensions. *Journal of Business & Economic Statistics*, *10*(4), 467–470.

Perron, P., & Vogelsang, T. J. (1992). Nonstationarity and level shifts with an application to purchasing power parity. *Journal of business & economic statistics*, *10*(3), 301-320.

Pesaran, M. H., Shin, Y., & Smith, R. J. (2001). Bounds testing approaches to the analysis of level relationships. *Journal of Applied Econometrics*, *16*(3), 289–326. https://doi.org/10.1002/JAE.616

Pesaran, M. H., & Smith, R. (1995). Estimating long-run relationships from dynamic heterogeneous panels. *Journal of Econometrics*, *68*(1), 79–113. https://doi.org/10.1016/0304-4076(94)01644-F

Ploberger, W. (1983). Testing the constancy of parameters in linear models. *Econometric Society European Meeting, Pisa*.

Ploberger, Werner, & Krämer, W. (1990). The local power of the CUSUM and CUSUM of squares tests. *Econometric Theory*, *6*(3), 335–347.

Ploberger, Werner, & Krämer, W. (1992). The CUSUM test with OLS residuals. *Econometrica: Journal of the Econometric Society*, 271–285.

Ploberger, Werner, Krämer, W., & Kontrus, K. (1989). A new test for structural stability in the linear regression model. *Journal of Econometrics*, *40*(2), 307–318.

Pretis, F., Schneider, L., Smerdon, J. E., & Hendry, D. F. (2017). Detecting volcanic eruptions in temperature reconstructions by designed break-indicator saturation. *Environmental Economics and Sustainability*, 7-37.

Pretis, F., Reade, J. J., & Sucarrat, G. (2018). Automated general-to-specific (GETS) regression modeling and indicator saturation for outliers and structural breaks. *Journal of Statistical Software*, *86*. https://doi.org/10.18637/jss.v086.i03

Qiao, H.-H., Deng, Z.-H., Li, H.-J., Hu, J., Song, Q., & Gao, L. (2021). Research on historical phase division of terrorism: An analysis method by time series complex network. *Neurocomputing*, *420*, 246–265.

Quandt, R. E. (1960). Tests of the Hypothesis That a Linear Regression System Obeys Two Separate Regimes. *Journal of the American Statistical Association*, *55*(290), 324–330. https://doi.org/10.1080/01621459.1960.10482067

Rao, C. R. (1952). *Advanced statistical methods in biometric research.*

Rao, C. R. (1965). The theory of least squares when the parameters are stochastic and its application to the analysis of growth curves. *Biometrika*, *52*(3/4), 447–458.

Rousseeuw, P. J. (1984). Least median of squares regression. *Journal of the American Statistical Association*, *79*(388), 871–880. https://doi.org/10.1080/01621459.1984.10477105

Rousseeuw, P. J., & Leroy, A. M. (2005). *Robust regression and outlier detection* (Vol. 589). John wiley & sons.

Sakurai, R., Ueki, M., Makino, S., Hozawa, A., Kuriyama, S., Takai-Igarashi, T., Kinoshita, K., Yamamoto, M., & Tamiya, G. (2019). Outlier detection for questionnaire data in biobanks.

*International Journal of Epidemiology*, *48*(4), 1305–1315.

Santos, C., Hendry, D. F., & Johansen, S. (2008). Automatic selection of indicators in a fully saturated regression. *Computational Statistics*, *23*(2), 317–335. https://doi.org/10.1007/s00180-007-0054-z

Schmidt, P., & Sickles, R. (1977). Some further evidence on the use of the Chow test under heteroskedasticity. *Econometrica: Journal of the Econometric Society*, 1293–1298.

Shahbaz, M., Jalil, A., & Islam, F. (2010). *Real exchange rate changes and trade balance in Pakistan: A revisit*.

Song, S., & Bickel, P. J. (2011). *Large Vector Auto Regressions, 1-28*. https://arxiv.org/abs/1106.3915v1

Stock, J. H., & Watson, M. W. (2010). *Modeling inflation after the crisis*. National Bureau of Economic Research, 1-61.

Stone, M. (1974). Cross-Validatory Choice and Assessment of Statistical Predictions. *Journal of the Royal Statistical Society: Series B (Methodological)*, *36*(2), 111–133. https://doi.org/10.1111/J.2517-6161.1974.TB00994.X

Stringer, D., Braude, P., Myint, P. K., Evans, L., Collins, J. T., Verduri, A., Quinn, T. J., Vilches-Moraga, A., Stechman, M. J., Pearce, L., Moug, S., McCarthy, K., Hewitt, J., Carter, B., Collaborators, C. S., Bruce, E., Einarsson, A., McGovern, A., Bisset, C., … Simmons, R. (2021). The role of C-reactive protein as a prognostic marker in COVID-19. *International Journal of Epidemiology*, *50*(2), 420–429. https://doi.org/10.1093/IJE/DYAB012

Tibshirani, R. (1996). Regression Shrinkage and Selection Via the Lasso. *Journal of the Royal*

*Statistical Society: Series B (Methodological)*, *58*(1), 267–288. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x

Toyoda, T. (1974). Use of the Chow test under heteroscedasticity. *Econometrica: Journal of the Econometric Society*, 601–608.

Uematsu, Y., & Tanaka, S. (2019). High-dimensional macroeconomic forecasting and variable selection via penalized regression. *The Econometrics Journal*, *22*(1), 34–56.

Verbanck, M., Chen, C., Neale, B., & Do, R. (2018). Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. *Nature Genetics*, *50*(5), 693–698.

Vilares, M. J. (1986). The suggested model. In *Structural Change in Macroeconomic Models* (pp. 79–118). Springer.

Wang, L. (2013). The L1 penalized LAD estimator for high dimensional linear regression. *Journal of Multivariate Analysis*, *120*, 135–151. https://doi.org/10.1016/j.jmva.2013.04.001

Westlund, A. H., & Törnkvist, B. (1989). On the identification of time for structural changes by MOSUM-SQ and CUSUM-SQ procedures. In *Statistical Analysis and Forecasting of Economic Structural Change* (pp. 97–126). Springer.

Wilcox, R. R. (2011). *Introduction to robust estimation and hypothesis testing*. Academic press.

Wilson, B., Tseng, C. L., Soroka, O., Pogach, L. M., & Aron, D. C. (2017). Identification of outliers and positive deviants for healthcare improvement: Looking for high performers in hypoglycemia safety in patients with diabetes. *BMC Health Services Research, 17*(1), 1–10. https://doi.org/10.1186/s12913-017-2692-3

Wu, X., Liang, R., & Yang, H. (2021). Penalized and constrained LAD estimation in fixed and high dimension. *Statistical Papers*. https://doi.org/10.1007/s00362-021-01229-0

Zaidi, S. A. (2005). Issues in Pakistan's economy. *OUP Catalogue* Oxford University Press, edition 2, number 9780195979145, 1-30.

Zaman, A., Rousseeuw, P. J., & Orhan, M. (2001). Econometric applications of high-breakdown robust regression techniques. *Economics Letters*, *71*(1), 1–8. https://doi.org/10.1016/S0165-1765(00)00404-3

Zhang, C. H. (2010). Nearly unbiased variable selection under minimax concave penalty. *Annals of Statistics*, *38*(2), 894–942. https://doi.org/10.1214/09-AOS729

Zivot, E., & Andrews, D. W. K. (2002). Further evidence on the great crash, the oil-price shock, and the unit-root hypothesis. *Journal of business & economic statistics*, *20*(1), 25-44.

Zou, H., & Hastie, T. (2003). Regression Shrinkage and Selection via the Elastic Net, with Applications to Microarrays. *Journal of the Royal Statistical Society: Series B*, *67*(1), 301–320. http://webdocs.cs.ualberta.ca/~mahdavif/ReadingGroup/Papers/10.1.1.9.6188.pdf

Zou, H. (2006). The adaptive lasso and its oracle properties. *Journal of the American Statistical Association*, *101*(476), 1418–1429.

## Appendix

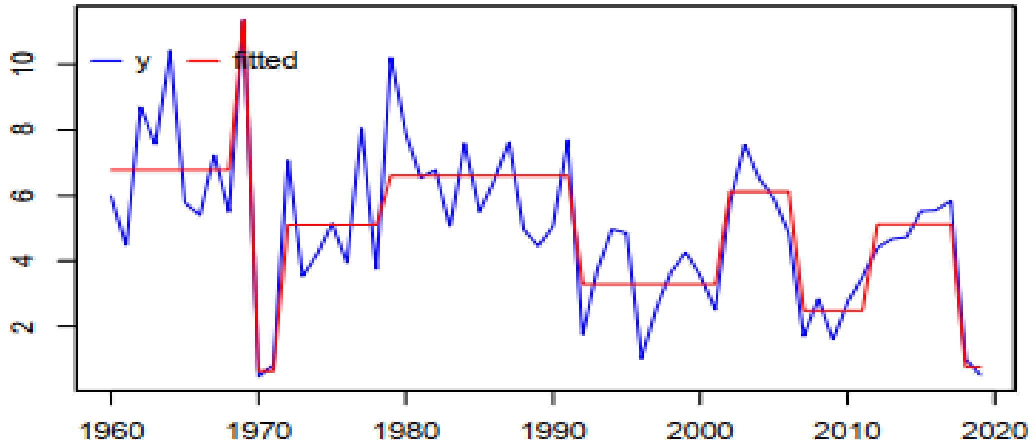



***Figure A. 1:*** The Estimated Four Breaks in the GDP Growth

**Table A. 1:** Empirical Results for the GDP growth (1960-2019)

| **Break Detection with 0.05 level of significance** | | | | | |
|---|---|---|---|---|---|
| | **mconst** | **sis1969** | **sis1970** | **sis1972** | **sis1979** |
| **Coef.** | 6.786 | 4.567 | -10.712 | 4.462 | 1.501 |
| **P-values** | (0.000)*** | (0.004)*** | (0.000)*** | (0.0003)*** | (0.029)** |
| | **sis1992** | **sis2002** | **sis2007** | **sis2012** | **sis2018** |
| **Coef.** | -3.320 | 2.829 | -3.635 | 2.641 | -4.362 |
| **P-values** | (0.000)*** | (0.0007)*** | (0.0002)*** | (0.003)*** | (0.0005)*** |
| **Diagnostic tests** | | | | | |
| **AR(1) Ljung-Box Test** | 0.556 (0.455) | | | | |
| **ARCH(1) Ljung-Box Test** | 0.138 (0.710) | | | | |
| **R-squared** | 0.704 | | | | |
| **Break Detection with 0.01 level of significance** | | | | | |
| | **mconst** | **sis1970** | **sis1972** | **sis1992** | |
| **Coef.** | 7.243 | -6.602 | 5.440 | -2.221 | |
| **P-values** | (0.000)*** | (0.000)*** | (0.000)*** | (0.000)*** | |
| **Diagnostic tests** | | | | | |
| **AR(1) Ljung-Box Test** | 0.99442 ( 0.3416) | | | | |
| **ARCH(1) Ljung-Box Test** | 0.99410 ( 0.3187) | | | | |
| **R-squared** | 0.427 | | | | |

*denotes the significance level (0.01***,0.05**,0.1*)

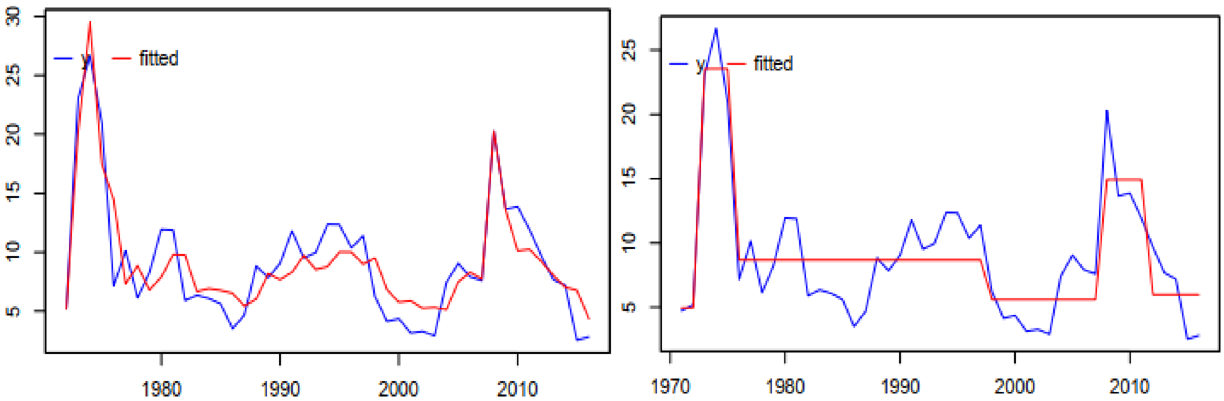



*Figure A. 2:* The Estimated Four Breaks in the Inflation Rates

**Table A. 2:** Empirical Results for the Inflation Rates (1971-2020)

**Break Detection with 0.05 level of significance**

| | mconst | ar1 | sis1973 | sis1975 | sis2008 | sis2009 |
|---|---|---|---|---|---|---|
| **Coef.** | 2.733 | 0.517 | 14.81 | -13.931 | 12.734 | -13.302 |
| **P-values** | (0.3163) | (0.000)*** | (0.000)*** | (0.000)*** | (0.000)*** | (0.000)*** |

**Diagnostics tests**

**AR(1) Ljung-Box Test** 0.165 (0.684)

**ARCH(1) Ljung-Box Test** 0.224 (0.635)

**R-squared** 0.778

**Break Detection with 0.05 level of significance**

| | mconst | sis1973 | sis1976 | sis1998 | sis2008 | sis2012 |
|---|---|---|---|---|---|---|
| **Coef.** | 4.957 | 18.589 | -14.857 | -3.076 | 9.321 | -8.949 |
| **P-values** | (0.014)** | (0.000)*** | (0.000)*** | (0.005)*** | (0.000)*** | (0.000)*** |

**Diagnostics tests**

**AR(1) Ljung-Box Test** 8.322( 0.004) **

**ARCH(1) Ljung-Box Test** 0.511 ( 0.474)

**R-squared** 0.764

**Break Detection with 0.01 level of significance**

| | mconst | ar1 | sis1973 | sis1975 | sis2008 | sis2009 |
|---|---|---|---|---|---|---|
| **Coef.** | 2.733 | 0.518 | 14.816 | -13.931 | 12.732 | -13.302 |
| **P-values** | 0.3163 | (0.000)*** | (0.000)*** | (0.000)*** | (0.000)*** | (0.000)*** |

**Diagnostics tests**

**AR(1) Ljung-Box Test** 0.165 (0.684)

**ARCH(1) Ljung-Box Test** 0.224 (0.635)

**R-squared** 0.778

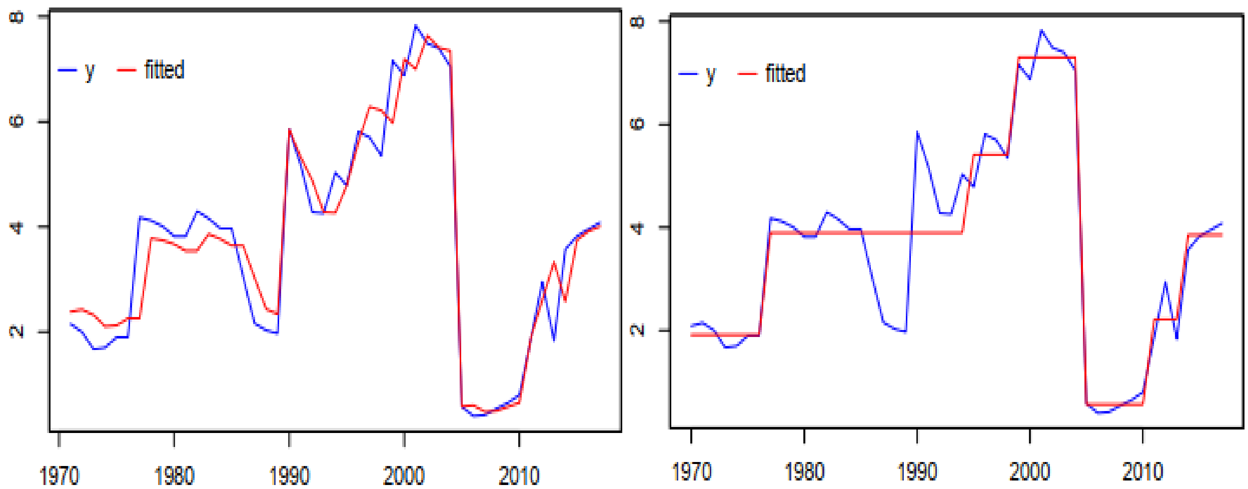



***Figure A. 3:*** The Estimated Six Breaks in the Unemployment Rates

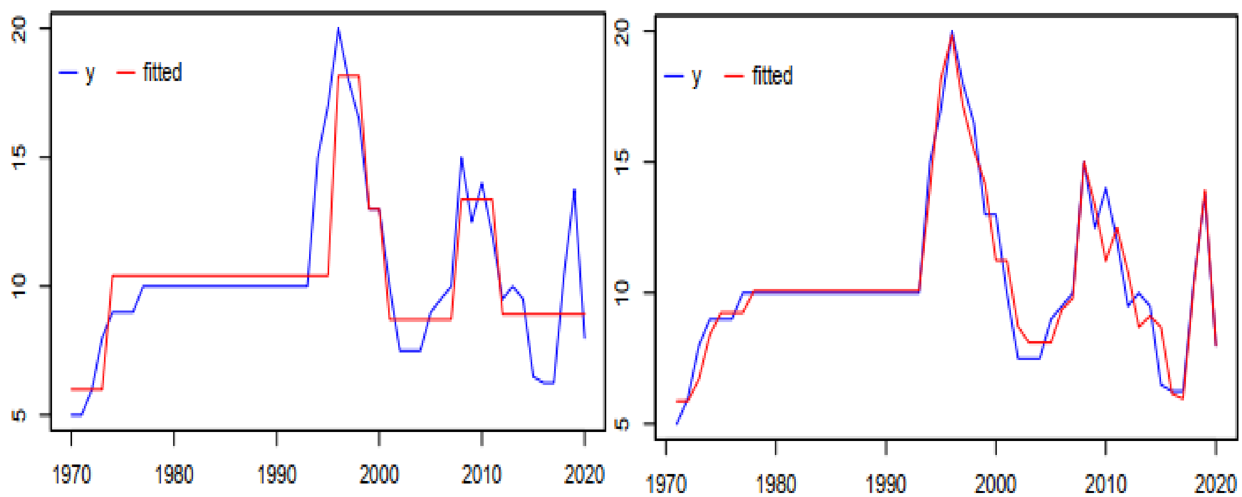



**Figure A. 4:** The Estimated Seven Breaks in the Interest Rates

**Table A. 3:** Empirical Results for the Unemployment Rates (1970-2017)

| Break Detection with 0.05 level of significance | | | | | | |
|---|---|---|---|---|---|---|
| | **mconst** | **ar1** | **sis1990** | **sis1991** | **sis1996** | **sis2005** |
| **Coef.** | 0.984 | 0.669 | -3.545 | 3.117 | 0.981 | -6.536 |
| **P-values** | (0.004)*** | (0.000)*** | (0.000)*** | (0.0001)*** | (0.013)** | (0.000)*** |
| | **sis2006** | **sis2011** | | | | |
| **Coef.** | 4.355 | 1.141 | | | | |
| **P-values** | (0.000)*** | (0.009)*** | | | | |
| **Diagnostics tests** | | | | | | |
| **AR(1) Ljung-Box Test** | 0.162 | (0.687) | | | | |
| **ARCH(1) Ljung-Box Test** | 0.068 | (0.793) | | | | |
| **R-squared** | 0.923 | | | | | |

| Break Detection with 0.05 level of significance | | | | | | |
|---|---|---|---|---|---|---|
| | **mconst** | **sis1977** | **sis1995** | **sis1999** | **sis2005** | **sis2011** |
| **Coef.** | 1.915 | 1.982 | 1.514 | 1.889 | -6.736 | 1.653 |
| **P-values** | (0.000)*** | (0.000)*** | (0.0004 )*** | (0.0002)*** | (0.000)*** | (0.002)*** |
| | **sis2014** | | | | | |
| **Coef.** | 1.638 | | | | | |
| **P-values** | (0.004)*** | | | | | |
| **Diagnostics tests** | | | | | | |
| **AR(1) Ljung-Box Test** | 5.697 | (0.017) * | | | | |
| **ARCH(1) Ljung-Box Test** | 34.88 | (3.504e-09) *** | | | | |
| **R-squared** | 0.895 | | | | | |

| Break Detection with 0.01 level of significance | | | | | | |
|---|---|---|---|---|---|---|
| | **mconst** | **ar1** | **sis1990** | **sis1991** | **sis2005** | **sis2006** |
| **Coef.** | 0.393 | 0.866 | 3.749 | -3.267 | -6.406 | 6.060 |
| **P-values** | 0.208 | (0.000)*** | (0.000)*** | (0.000)*** | (0.000)*** | (0.000)*** |
| **Diagnostics tests** | | | | | | |
| **AR(1) Ljung-Box Test** | 0.314 | (0.579) | | | | |
| **ARCH(1) Ljung-Box Test** | 0.039 | (0.844) | | | | |
| **R-squared** | 0.900 | | | | | |

*Denotes the significance level (0.01***,0.05**,0.1*)

**Table A. 4:** Empirical Results for the Interest Rates (1970-2020)

| | | | | | |
|---|---|---|---|---|---|
| **Break Detection with 0.05 level of significance** | | | | | |
| | **mconst** | **ar1** | **sis1994** | **sis1997** | **sis2003** |
| **Coef.** | 1.657 | 0.843 | 3.873 | -5.246 | 1.502 |
| **P-values** | (0.006)** | (0.000)** | (0.000)** | (0.000)** | (0.034)** |
| | **sis2008** | **sis2009** | **sis2018** | **sis2020** | |
| **Coef.** | 4.782 | -5.880 | 4.375 | -8.656 | |
| **P-values** | (0.000)** | (0.000)** | (0.000)** | (0.000)** | |
| **Diagnostics tests** | | | | | |
| **AR(1) Ljung-Box Test** | 3.195 | (0.073) . | | | |
| **ARCH(1) Ljung-Box Test** | 0.174 | (0.67) | | | |
| **R-squared** | 0.926 | | | | |
| **Break Detection with 0.05 level of significance** | | | | | |
| | **mconst** | **sis1974** | **sis1996** | **sis1999** | **sis2001** |
| **Coef.** | 6.00 | 4.409 | 7.757 | -5.166 | -4.285 |
| **P-values** | (0.000)*** | (0.000)*** | (0.000)*** | (0.003)*** | (0.005)*** |
| | **sis2008** | **sis2012** | | | |
| **Coef.** | 4.661 | -4.458 | | | |
| **P-values** | (0.002)*** | (0.000)*** | | | |
| **Diagnostics tests** | | | | | |
| **AR(1) Ljung-Box Test** | 8.561 | (0.003)** | | | |
| **ARCH(1) Ljung-Box Test** | 5.584 | (0.018) * | | | |
| **R-squared** | 0.70 | | | | |
| **Break Detection with 0.01 level of significance** | | | | | |
| | **mconst** | **ar1** | **sis1994** | **sis1997** | **sis2008** |
| **Coef.** | 2.357 | 0.767 | 4.242 | -4.719 | 5.453 |
| **P-values** | (0.000)*** | (0.000)*** | (0.000)*** | (0.000)*** | (0.000)*** |
| | **sis2009** | **sis2018** | **sis2020** | | |
| **Coef.** | -5.836 | 4.207 | -8.246 | | |
| **P-values** | (0.000)*** | (0.000)*** | (0.000)*** | | |
| **Diagnostics tests** | | | | | |
| **AR(1) Ljung-Box Test** | 0.287162 | (0.5920) | | | |
| **ARCH(1) Ljung-Box Test** | 0.052904 | (0.8181) | | | |
| **R-squared** | 0.91805 | | | | |

*Denotes the significance level (0.01***,0.05**,0.1*)

**Table A. 5:** Simulated result with lag length equal 2

| | Gauge | Potency | RMSE | MAE |
|---|---|---|---|---|
| | **ϕ equal 0.1** | | | |
| **WLAdaLASSO** | 0.2 | 0.691 | 1.202 | 1.009 |
| **Auto(0.05)** | 0.041 | 0.258 | 1.37 | 1.156 |
| **Auto(0.01)** | 0.011 | 0.15 | 1.355 | 1.142 |
| **SCAD** | 0.414 | 0.598 | 1.544 | 1.303 |
| **MCP** | 0.339 | 0.545 | 1.507 | 1.269 |
| **LASSO** | 0.415 | 0.647 | 1.476 | 1.258 |
| **AdaLASSO** | 0.259 | 0.482 | 1.343 | 1.134 |
| | **ϕ equal 0.5** | | | |
| **WLAdaLASSO** | 0.186 | 0.719 | 1.409 | 1.179 |
| **Auto(0.05)** | 0.047 | 0.356 | 1.655 | 1.39 |
| **Auto(0.01)** | 0.03 | 0.181 | 1.75 | 1.465 |
| **SCAD** | 0.399 | 0.484 | 1.525 | 1.276 |
| **MCP** | 0.339 | 0.43 | 1.504 | 1.262 |
| **LASSO** | 0.44 | 0.449 | 1.381 | 1.152 |
| **AdaLASSO** | 0.292 | 0.299 | 1.372 | 1.142 |
| | **ϕ equal 0.8** | | | |
| **WLAdaLASSO** | 0.198 | 0.638 | 1.75 | 1.563 |
| **Auto(0.05)** | 0.073 | 0.354 | 1.657 | 1.453 |
| **Auto(0.01)** | 0.044 | 0.299 | 1.549 | 1.345 |
| **SCAD** | 0.377 | 0.523 | 1.727 | 1.481 |
| **MCP** | 0.355 | 0.501 | 1.709 | 1.461 |
| **LASSO** | 0.569 | 0.731 | 1.664 | 1.431 |
| **AdaLASSO** | 0.387 | 0.557 | 1.575 | 1.35 |

**Table A. 6:** Diebold Statistics of covariate and its lag selection

| | **T=50(WAdaLASSO)** | | **T=100(WAdaLASSO)** | | **T=500(WAdaLASSO)** | |
|---|---|---|---|---|---|---|
| | t-statistics | p-value | t-statistics | p-value | t-statistics | p-value |
| | | | **$\phi$ equal 0.1** | | | |
| **SCAD** | -1.4597 | 0.1506 | -0.1829 | 0.8552 | 1.5372 | 0.1249 |
| **MCP** | -2.4783 | 0.0166 | -0.0115 | 0.9908 | 2.0121 | 0.0447 |
| **LASSO** | -1.8629 | 0.0684 | -2.7450 | 0.0072 | 0.2109 | 0.8330 |
| **AdaLASSO** | -1.1578 | 0.2524 | 0.5273 | 0.5992 | 1.6948 | 0.0907 |
| **Auto(0.05)** | -2.3337 | 0.0237 | -1.4993 | 0.1369 | 1.6948 | 0.0907 |
| **Auto(0.01)** | -1.7897 | 0.0796 | -1.7622 | 0.0811 | 1.1307 | 0.2587 |
| | | | **$\phi$ equal 0.5** | | | |
| **SCAD** | -2.4347 | 0.0185 | 1.0139 | 0.3131 | 1.0139 | 0.3131 |
| **MCP** | -3.3129 | 0.0017 | 1.0505 | 0.2960 | 1.0505 | 0.2960 |
| **LASSO** | -2.8976 | 0.0056 | 1.6089 | 0.1108 | 1.6089 | 0.1108 |
| **AdaLASSO** | -2.4764 | 0.0167 | 1.8629 | 0.0654 | 1.8629 | 0.0654 |
| **Auto(0.05)** | -1.8044 | 0.0772 | -0.9568 | 0.3410 | -0.9568 | 0.3410 |
| **Auto(0.01)** | -0.6604 | 0.5120 | -1.4552 | 0.1487 | -1.4552 | 0.1487 |
| | | | **$\phi$ equal 0.8** | | | |
| **SCAD** | -3.0823 | 0.0033 | -3.8363 | 0.0002 | 0.8196 | 0.4128 |
| **MCP** | -3.1853 | 0.0025 | -3.7379 | 0.0003 | 0.8835 | 0.3774 |
| **LASSO** | -4.5018 | 0.0000 | -1.6079 | 0.1110 | 0.2814 | 0.7785 |
| **AdaLASSO** | -3.7075 | 0.0005 | -2.8995 | 0.0046 | 0.7558 | 0.4501 |
| **Auto(0.05)** | -5.2710 | 0.0000 | -5.7281 | 0.0000 | -6.0708 | 0.0000 |
| **Auto(0.01)** | -4.8281 | 0.0000 | -4.9780 | 0.0000 | -6.0461 | 0.0000 |