

University of Toronto Mississauga
Introduction to Statistical Learning Learning
STA314H5E, Fall 2023
Dr. Masoud Ataei

Assignment 2

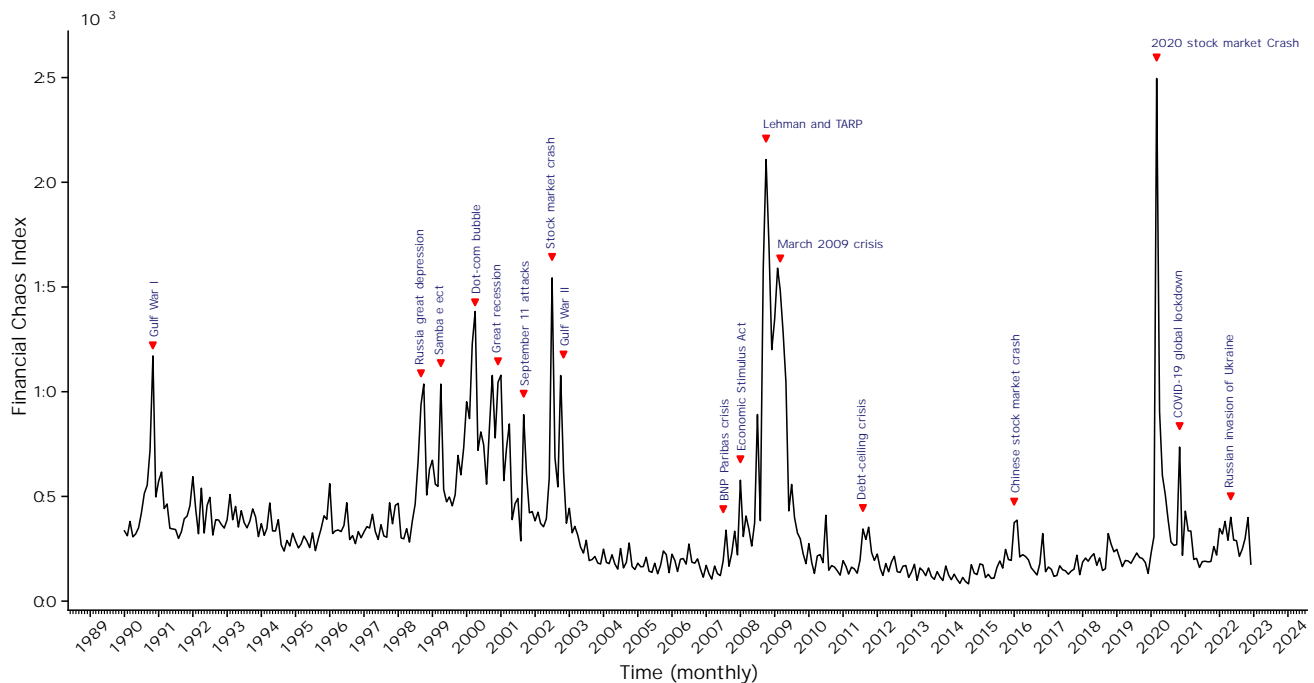
Due November 9, 2023 at 11:59 pm

***Analysis of the Stock Market Fluctuations, Anomalies
and Fear Index Using Data-driven Methodologies***

Forces Driving the Market Volatility

There are several different metrics to quantify market volatility; however, the most well-known and followed metrics are realized and implied volatility. The realized volatility gauges the fluctuations of underlying securities or indices by measuring price changes over predetermined periods, while implied volatility is a forward-looking metric representing future expectations of the market's uncertainty. The most important member of the latter family is the Chicago Board Options Exchange's (CBOE) VIX index that can be considered an estimator of the equity market's implied volatility.

In this assignment, your task is to build a regression model to predict the one-step-ahead value of the monthly VIX using publicly available economic and financial information that are obtained using text mining algorithms. One of the predominant indices of such kind is the [equity market volatility \(EMV\) tracker presented in \(Baker, Bloom, Davis, and Kost, 2019\)](#). More specifically, EMV is constructed by obtaining daily counts of articles containing at least one term in the categories economy or economic; uncertain or uncertainty; and one or more words from the equity market, equity price, stock market and stock price. It is noted that the U.S.-related articles used in the construction of EMV exceed 1000 newspapers and are retrieved from the Access World News' NewsBank service. In the following, we will only consider the monthly category-specific EMV uncertainty indexes for the purpose of identifying the economic and financial factors that trigger regime switches and state transitions in the stock market.



Phase I.

Load the dataset containing monthly values of VIX and all categories of monthly EMV trackers from January 1990 - December 2022, and develop a regression model not only for the purpose of predicting the future values of VIX. But, more importantly, the results of the regression analysis should possibly lead to new information by revealing the economic, financial and political factors that affect the behaviour of VIX over time. For implementation purposes, you are free to either use your previous program codes, the glmnet library, or any other packages available on open source. Also, note that the last column corresponds to values of VIX at month $t+1$, whereas its covariates at each row pertain to month t .

Solve this problem using OLS, LASSO, Ridge Regression and Elastic Net Regression, and compare their performances by providing a thorough interpretation of their results. For each method, you may discuss its advantages and drawbacks, such as presence of multicollinearity, and then elaborate on their other aspects like feature selection, sparsity, etc.

Can you provide a coherent picture of mechanisms underlying the sudden volatility changes in stock markets in the past 23 years?

Phase II.

Most often, the forces driving the market movements themselves undergo changes when a transition occurs from one segment of time into its following one(s). Use elastic nets to construct predictive regression models for each segment of stock market and then interpret the results with the goal of extracting knowledge about the phenomena causing the regime switches and apparent transitions.

Can you characterize the nature of the most dominant forces responsible for driving the stock market movements within each segment? Are there any pattern on sparsity of forces when comparing chaotic (pink) and normal (blue) segments of the stock market?

Can you provide a coherent picture of mechanisms underlying the sudden volatility changes in stock markets in the past 23 years, by further considering the relatively homogeneous segments of the market?

Segment	1	2	3	4	5	6
Period	1990-01 to 1998-06	1998-07 to 2003-03	2003-04 to 2007-12	2008-01 to 2009-09	2009-10 to 2019-12	2022-01 to 2022-12

