

# **Three Essays on Volatility and Tail-Risk**

by

Moses Oluwafolajimi Dada

A thesis submitted for the degree of

Doctor of Philosophy

Essex Business School

at the University of Essex

October, 2024

Doctoral Committee:

Dr. Luiz Vitellio, University of Essex, Essex Business School

Dr. Vivek Nawosah, University of Essex, Essex Business School

Dr. Athanasios Triantafyllou, IESEG School of Management

Moses Oluwafolajimi Dada

[md20759@essex.ac.uk](mailto:md20759@essex.ac.uk); [mosesdadaphd@outlook.com](mailto:mosesdadaphd@outlook.com)

© Moses Oluwafolajimi Dada, 2025

I hereby certify that the thesis I have submitted for examination for the MPhil/PhD degree at the University of Essex is entirely my own work, except where explicitly stated otherwise. In cases where work was carried out jointly, the contributions of other individuals are clearly acknowledged. The copyright of this thesis resides with me as the author. Quotations from the thesis are permitted with proper attribution, but it may not be reproduced in any form without my prior written consent. I confirm, to the best of my knowledge, that this authorisation does not infringe upon the rights of any third party.

Moses Oluwafolajimi Dada - 04/10/2024

## Acknowledgements

First and foremost, I would like to express my deepest gratitude to my PhD supervisors, Dr Luiz Vitellio, Dr Vivek Nawosah, and Dr Athanasios Triantafyllou. Your unwavering support, insightful guidance, and relentless encouragement have been invaluable throughout this journey. You have not only helped shape this thesis but also contributed to my growth as a thinker, researcher, and scholar. I am truly fortunate to have had the opportunity to learn from Dr Luiz Vitellio.

To my beloved mother, who has always been my pillar of strength and greatest advocate—your unconditional love, wisdom, and sacrifices have made everything possible. Your faith in God and in me has been my guiding light. I owe this accomplishment to the values you instilled in me and the endless encouragement you provided. This achievement is as much yours as it is mine.

To my sister Ayomide, whose constant encouragement and understanding have been a source of comfort and joy—thank you for always being there to listen.

To my wife Oluwatosin, my partner —your understanding is truly appreciated.

To my unborn child, as I write this, you're not yet in my arms, but I already think about the

future we will share and the person you will become. One of the greatest gifts I hope to pass on to you is the love of learning. Through my own studies, I've learned that knowledge is powerful—but wisdom comes from how you apply that knowledge. You'll acquire a lot of information over the years, but what will matter most is how you use it to improve your life and the lives of others.

Lastly, I would like to close with a quote that resonates deeply with the themes of this thesis and the journey of pursuing this PhD:

*"In the real world, the most important things we deal with—randomness, uncertainty, opacity, and incomplete understanding—are all driven by a much thicker layer of complexity."* — Nassim Nicholas Taleb, *The Black Swan: The Impact of the Highly Improbable*.

## Table of Contents

<b>Acknowledgements</b>	<b>iii</b>
<b>List of Figures</b>	<b>ix</b>
<b>List of Tables</b>	<b>x</b>
<b>Abstract</b>	<b>xii</b>
<b>Publication</b>	<b>xiv</b>
<b>Chapter 1: INTRODUCTION</b>	<b>1</b>
<b>Chapter 2: AN EX-ANTE AND EX-POST ANALYSIS OF OPTION- IMPLIED PROBABILITY DENSITY FUNCTIONS IN TIMES OF MACROECONOMIC UNCERTAINTY</b>	<b>8</b>
2.1 Introduction . . . . .	14
2.2 Literature Review . . . . .	17
2.2.1 Construction of PDF . . . . .	17
2.2.2 Risk-Neutral Density Estimation Methods . . . . .	18
2.2.3 Impact of Macroeconomic Uncertainty . . . . .	20
2.2.4 Economic Theory and Interpretation . . . . .	22
2.3 Data Description . . . . .	24
2.4 Methodology . . . . .	26
2.4.1 Risk-Neutral Densities . . . . .	26
2.4.2 Non-Parametric Methods . . . . .	29
2.4.2.1 Shimko's Spline Method . . . . .	29
2.4.2.2 Edgeworth Density . . . . .	30
2.4.3 Parametric Methods . . . . .	33
2.4.3.1 Log-normal Density Implied by the Black-Scholes-Merton Model . . . . .	33
2.4.3.2 The Mixture of Two Log-normals . . . . .	35

2.4.3.3	Generalised Beta of the Second Kind (GB2) . . . . .	36
2.4.4	Comparing Forecasting Density Accuracy Using the Diebold-Mariano Test . . . . .	38
2.4.5	Assessing Uncertainty Shocks . . . . .	40
2.4.5.1	The Proxy SVAR Model . . . . .	41
2.4.6	Estimating the Tail Loss Measure . . . . .	43
2.4.7	Decomposing Volatility into High and Low States . . . . .	45
2.4.8	Probit Regression Model . . . . .	46
2.5	Empirical Analysis . . . . .	48
2.5.1	Comparing Parametric and Non-Parametric Methods in Recovering the Risk Neutral Density . . . . .	48
2.5.1.1	Real Data Study . . . . .	48
2.5.2	Interpreting Densities . . . . .	50
2.5.2.1	Evolution of the Densities . . . . .	50
2.5.2.2	Statistical Analysis . . . . .	55
2.5.2.3	Whole Dataset Analysis . . . . .	58
2.5.3	Evolution of the Tail shape Parameter . . . . .	62
2.5.3.1	News Events Selection . . . . .	63
2.5.4	Probit Model Results . . . . .	71
2.6	Conclusion . . . . .	76
<b>Chapter 3:</b>	<b>EXTREME VALUE THEORY-MANAGED PORTFOLIOS</b>	<b>91</b>
3.1	Introduction . . . . .	96
3.2	Literature Review . . . . .	102
3.3	Data Description . . . . .	110
3.4	Methodology . . . . .	113
3.4.1	EVT-VaR Peaks-Over-Threshold . . . . .	113
3.4.1.1	Summary Steps to Compute EVT-VaR . . . . .	116
3.4.2	EVTM Methodology . . . . .	117
3.4.2.1	Constructing EVTm Portfolios . . . . .	118
3.4.2.2	Single-Factor EVTm Portfolios . . . . .	121
3.4.2.3	Multi-Factor EVTm Portfolios . . . . .	125
3.4.2.4	Summary Steps to Compute EVTm portfolios . . . . .	126
3.4.3	Volatility Managed Portfolios . . . . .	129
3.4.3.1	Single-Factor Volatility Managed Portfolios . . . . .	129
3.4.3.2	Multi-Factor VM Portfolios . . . . .	130

3.4.3.3	Summary Steps to Compute VM portfolios . . . . .	131
3.4.4	Transaction Costs . . . . .	132
3.5	Empirical Analysis . . . . .	134
3.5.1	Single-Factor EVTm Portfolios . . . . .	137
3.5.1.1	Sharpe Ratio and Sortino Ratio Optimised Portfolios . . . . .	140
3.5.1.2	Spanning Regressions . . . . .	143
3.5.2	Multi-Factor EVTm Portfolios . . . . .	146
3.5.2.1	Spanning Regressions . . . . .	149
3.5.3	Break-Even Transaction Costs . . . . .	150
3.6	Conclusion . . . . .	152
<b>Chapter 4:</b>	<b>A ROUGH GARCH-TYPE LSTM REALISED VOLATILITY MODEL</b>	<b>155</b>
4.1	Introduction . . . . .	161
4.2	Literature Review . . . . .	165
4.3	Data Description . . . . .	174
4.4	Methodology . . . . .	178
4.4.1	Fractional Brownian and Hurst Exponent . . . . .	178
4.4.2	GARCH-Type Models . . . . .	181
4.4.2.1	Estimating GARCH-Type Model Parameters . . . . .	183
4.4.3	Long Short-Term Memory . . . . .	184
4.4.3.1	Neural Network Models . . . . .	184
4.4.4	Realised Volatility . . . . .	188
4.4.5	Development of the Models . . . . .	188
4.4.6	Loss Functions and Tests . . . . .	192
4.4.7	Three-State First-Order Markov Switching Regression Model . . . . .	194
4.4.8	Markov Switching Regression Model Specification . . . . .	195
4.5	Empirical Analysis . . . . .	197
4.5.1	Model Configuration . . . . .	197
4.5.1.1	Training Set Percentages . . . . .	197
4.5.1.2	Number of LSTM Layers . . . . .	201
4.5.1.3	$n$ Days-Ahead Forecasts . . . . .	203
4.5.1.4	Model Comparison . . . . .	206
4.5.1.5	Model Tests . . . . .	207
4.5.2	Political and Financial Events and Volatility States . . . . .	210
4.6	Conclusion . . . . .	217

<b>Chapter 5: CONCLUSION</b>	<b>220</b>
<b>Bibliography</b>	<b>224</b>
<b>Appendices</b>	<b>234</b>

## List of Figures

Figure 2.1	GB2 Distribution Family Tree . . . . .	37
Figure 2.2	Evolution of Densities of Political Uncertainty . . . . .	52
Figure 2.3	Evolution of Densities of Financial Uncertainty . . . . .	53
Figure 2.4	Dodd-Frank-Act signed by Obama . . . . .	56
Figure 2.5	Obama re-elected . . . . .	57
Figure 2.6	Tail Shape Parameter from 1996-2019 . . . . .	65
Figure 2.7	Tail Shape Parameter for Various Dates Spanning from 1996-2007 . . . . .	66
Figure 2.8	Tail Shape Parameter for Various Dates Spanning from 2007-2013 . . . . .	67
Figure 2.9	Tail Shape Parameter for Various Dates from 2013-2017 . . . . .	68
Figure 2.10	Monthly Time-Series of Tail Shape Parameter from 1996-2021 . . . . .	69
Figure 2.11	$\xi\sigma_t - \tau_t$ is Decomposed in to $\delta_t^{\text{high}}(\lambda)$ and $\delta_t^{\text{low}}(\lambda)$ Volatility for the Dates 1996-2021 . . . . .	70
Figure 2.12	Low Volatility vs. Debt-To-GDP . . . . .	76
Figure 3.1	Fama-French charts . . . . .	112
Figure 3.2	$EVT^\alpha$ Density Plots of Fama-French Factors . . . . .	136
Figure 3.3	Plots of Weights of MVE, Sortino and Sharpe Optimised EVT Portfolios . . . . .	148
Figure 4.1	S&P 500 index daily close price $P_t$ time-series . . . . .	175
Figure 4.2	S&P 500 index daily logarithmic return $R_t$ time-series . . . . .	175
Figure 4.3	S&P 500 index intraday logarithmic return $R_\tau$ time-series . . . . .	175
Figure 4.4	LSTM Cell Illustration . . . . .	187
Figure 4.5	Sliding Window Design for Training and Test Set and Predictions over the Entire Period . . . . .	191
Figure 4.6	RV True Values vs. RV Predicted Values of Forecasting Models . . . . .	200
Figure 4.7	Multivariate Diebold-Mariano(DM) Test . . . . .	209
Figure 4.8	Markov Switching Model Volatility States . . . . .	212

## List of Tables

Table 2.1	Summary Statistics for Options Data . . . . .	25
Table 2.2	Diebold-Mariano Test Applied to Shimko’s Spline Method, Edgeworth Density, Black-Scholes-Merton Model, Mixture of Two Log-normals, and GB2 Forecasts of S&P500 European Call Prices. . . . .	49
Table 2.3	Ex-ante and Ex-post Calls and Puts of Financial and Political Events	60
Table 2.4	Ex-ante and Ex-post Calls and Puts of Financial and Political Events with Gold Proxy . . . . .	61
Table 2.5	Descriptive Analysis . . . . .	70
Table 2.6	Volatility and Crisis . . . . .	73
Table 2.7	Volatility and Crisis Using Different Lag Lengths . . . . .	74
Table 2.8	Ex-ante and Ex-post MAE of Calls of Financial and Political Events	78
Table 2.9	Ex-ante and Ex-post MAE of Calls of Financial and Political Events	79
Table 2.10	Ex-ante and Ex-post MAPE of Calls of Financial and Political Events	80
Table 2.11	Ex-ante and Ex-post MAPE of Calls of Financial and Political Events	81
Table 2.12	Ex-ante and Ex-post RMSE of Calls of Financial and Political Events	82
Table 2.13	Ex-ante and Ex-post RMSE of Calls of Financial and Political Events	83
Table 2.14	Ex-ante and Ex-post MAE of Puts of Financial and Political Events	84
Table 2.15	Ex-ante and Ex-post MAE of Puts of Financial and Political Events	85
Table 2.16	Ex-ante and Ex-post MAPE of Put of Financial and Political Events	86
Table 2.17	Ex-ante and Ex-post MAPE of Put of Financial and Political Events	87
Table 2.18	Ex-ante and Ex-post RMSE of Puts of Financial and Political Events	88
Table 2.19	Ex-ante and Ex-post RMSE of Puts of Financial and Political Events	89
Table 2.20	List of Financial and Political Events . . . . .	90
Table 3.1	Summary Statistics for Data . . . . .	111
Table 3.2	$\widehat{P}_{t+1}(\alpha) - \kappa_t(\alpha)$ Confusion Matrix . . . . .	120
Table 3.3	EVTM Optimal Threshold $\kappa^*(\alpha)$ Estimates . . . . .	137
Table 3.4	EVTM Logit Estimates . . . . .	139

Table 3.5	Unmanaged, VM, Sharpe Optimised EVT <sub>M</sub> , and Sortino Optimised EVT <sub>M</sub> Portfolio Performance Measures for Factors Mkt-RF, SMB, CMA, HML, and RMW . . . . .	142
Table 3.6	Single-Factor Portfolio Alphas . . . . .	145
Table 3.7	EVT <sub>M</sub> Sharpe Ratio and Sortino Ratio Optimised Performance Measures	147
Table 3.8	Multi-Factor Portfolio Alphas . . . . .	150
Table 3.9	Break-Even Transaction Costs for Single-Factor and Multi-Factor Portfolios . . . . .	151
Table 4.1	Pairwise Comparisons of Forecasting Models . . . . .	170
Table 4.2	Summary Statistics for Data . . . . .	176
Table 4.3	Input Variables for Models . . . . .	192
Table 4.4	Comparison of $RV_t$ Forecasting Errors of Various Models over Different Time Horizons for Various Training Set Percentages . . . . .	198
Table 4.5	Comparison of $RV_t$ Forecasting Errors of Various Models over Different Time Horizons for Various Layers . . . . .	202
Table 4.6	Optimal configurations for forecasts . . . . .	202
Table 4.7	Comparison of $RV_t$ MAE Forecasting Errors of Various Models over Different Time Horizons for Various $n$ Days-Ahead . . . . .	204
Table 4.8	Comparison of $RV_t$ RMSE Forecasting Errors of Various Models over Different Time Horizons for Various $n$ Days-ahead . . . . .	205
Table 4.9	Single Models vs. Hybrid GARCH-Type LSTM Comparison . . . . .	207
Table 4.10	DM Test Pair-Wise Comparison of GARCH-Type LSTM Models . . . . .	208
Table 4.11	Wilcoxon Signed-Rank Test(W <sub>S</sub> ) Pair-Wise Comparison of GARCH-type LSTM Models . . . . .	210
Table 4.12	Markov Switching Model Regression Results . . . . .	211
Table 4.13	Low Volatility State Events . . . . .	214
Table 4.14	Medium Volatility State Events . . . . .	215
Table 4.15	High Volatility State Events . . . . .	216
Table A.1	Diebold-Mariano Test Applied to Shimko’s Spline Method, Edgeworth Density, Black-Scholes-Merton Model, Mixture of Two Log-normals, and GB2 Forecasts of S&P500 European Call Prices. . . . .	235
Table A.2	Time Expense of Density Extraction Methods . . . . .	236

## Abstract

This thesis examines three interconnected topics: Option-implied Probability Density Functions (PDFs) during macroeconomic uncertainty, Extreme Value Theory (EVT) volatility-managed portfolios, and rough GARCH-type LSTM models for volatility forecasting. Using advanced econometric and machine learning methods, the study provides deeper insights into financial market dynamics.

Chapter 2 investigates option-implied PDFs, which reveal investor expectations about future asset prices. Unlike traditional methods, PDFs provide a richer perspective by capturing a broader range of investor beliefs. The study employs non-parametric and parametric techniques to analyse the impact of macroeconomic uncertainty on these densities. It uses proxy Structural Vector Autoregressive (SVAR) models to study uncertainty shocks, while EVT tail shape parameters assess distribution tail decay. Findings highlight the link between anticipated uncertainty and its resolution, showing how uncertainty influences risk-taking behaviours. A Probit model indicates that low volatility periods often precede financial crises.

Chapter 3 focuses on a volatility-managed portfolio strategy aimed at reducing tail risk and enhancing the Sharpe ratio. Built on the Fama-French Five-Factor Model, the strategy is compared with benchmarks like buy-and-hold and Moreira & Muir (2017) strategies. The

chapter introduces EVT to better capture extreme loss risks by focusing on distribution tails, addressing limitations of Value at Risk (VaR). The analysis shows that EVT-based strategies effectively mitigate tail risk and improve returns, offering a novel volatility timing approach.

Chapter 4 presents rough hybrid Long Short-Term Memory (LSTM) models, including rGARCH-LSTM, rEGARCH-LSTM, and rGE-LSTM, to improve volatility forecasting. These models combine financial time series roughness, LSTM predictive power, and GARCH-type model robustness. Tested on intraday SPX data, the hybrid models outperform conventional methods, enhancing volatility forecasts and risk management strategies.

This thesis advances our understanding of volatility, tail risk, and uncertainty in financial markets, providing valuable insights for investors, policymakers, and financial institutions in risk management and investment decisions.

## Publication

Portions of this thesis are slated for publication, reflecting significant contributions to the field of finance.

- **AN EX-ANTE AND EX-POST ANALYSIS OF OPTION-IMPLIED PROBABILITY DENSITY FUNCTIONS IN TIMES OF MACROECONOMIC UNCERTAINTY.**

Authors: M. O. Dada, L. Vitellio, V. Nawosah, A. Triantafyllou.

Target Journal: To Be Announced (Pending Submission)

This study provides a detailed investigation into the behaviour of option-implied probability density functions before and after major economic events, offering novel insights into low volatility as an early predictor of macroeconomic instability. The research explores the implications of these pricing mechanisms for financial stability and risk management strategies. The findings from this research were presented at prestigious academic venues, highlighting the robustness and relevance of the work:

– **International Symposium in Finance (ISF2022), Greece.**

This presentation facilitated rigorous academic discourse and provided valuable feedback from leading experts in the field, further refining the research and positioning it for successful publication in a high-impact journal.

• **EXTREME VALUE THEORY-MANAGED PORTFOLIOS.**

Authors: M. O. Dada, L. Vitellio, V. Nawosah.

Target Journal: To Be Announced (Pending Submission)

This study introduces a trading strategy that harnesses volatility to mitigate portfolio tail risk while enhancing the Sharpe ratio. It makes a significant contribution by encouraging the adoption of Extreme Value Theory (EVT) within the Volatility-Managed portfolio literature. The research underscores the advantages of EVT in strengthening risk management and optimising volatility timing, resulting in improved financial performance, particularly in the face of market volatility and extreme events.

## Chapter 1

### INTRODUCTION

This research aims to deepen our understanding of how financial markets react to various economic and geopolitical events by exploring the evolution of option-implied PDFs, implementing volatility-managed strategies to optimise portfolios while managing tail risks, and forecasting realised volatility.

Chapter 2 focusses on the evolution of option-implied PDFs during times of macroeconomic uncertainty and offers valuable insights into market expectations and risk assessments. These PDFs, which reflect the risk-neutral probabilities of future asset prices, are derived from market-traded options and are crucial for understanding how investors perceive risk and uncertainty. An option-implied Probability Density Function (PDF) depicts estimates of future movements of an asset's price as priced by investors. PDFs are significant as PDFs are forward-looking and based on market prices; in this Chapter, I compute PDFs using option prices. By analysing the ex-ante, current, and ex-post states of these PDFs, we can observe how market sentiments evolve in response to anticipated and realised economic events.

The theoretical foundation for this analysis is built upon the risk-neutral valuation principle,

as established by Cox & Ross (1976) and Ross (1976) which posits that the price of an option is equal to the expected present value of its payoff, discounted under the Risk-Neutral Density (RND). The mathematical framework developed by Breeden & Litzenberger (1978) further allows for the extraction of these RNDs by calculating the second derivative of the option price with respect to the strike price. This method provides a direct link between option prices and market expectations of future asset prices.

The study applies a range of parametric and non-parametric methods to estimate RNDs, ultimately selecting the mixture of log-normals model for its superior accuracy in forecasting European call option prices, as determined by the Diebold-Mariano test from Diebold & Mariano (1995). Through this approach, the empirical relationship between ex-ante uncertainty arising from scheduled and unscheduled macroeconomic news, and the ex-post resolution of uncertainty in financial markets is explored. This relationship is further analysed by implementing a proxy SVAR model, incorporating macro economic news events, estimating the Extreme Value Theory Tail Loss Measure (TLM), and applying Probit regression models to investigate the predictive power of low volatility in forecasting financial crises. The results are further examined with an array of loss functions.

Chapter 3 focuses on portfolio optimisation, particularly in managing tail risk while improving the Sharpe ratio, a key measure of risk-adjusted return. The study constructs a portfolio using the Fama-French Five-Factor Model, which extends the original three-factor model by incorporating profitability and investment factors. This model is compared against benchmark portfolios, including buy-and-hold strategies and volatility-managed portfolios.

A central theme in this analysis is the application of Extreme Value Theory (EVT) to manage tail risk, as EVT is particularly effective in assessing the risk of extreme losses, which traditional Value-at-Risk (VaR) measures might underestimate. By focusing on the tails of the distribution, EVT provides a more accurate risk assessment in scenarios where return distributions deviate significantly from normality. The study demonstrates that incorporating EVT into the portfolio optimisation process can lead to a more resilient portfolio, better equipped to handle extreme market events.

Portfolios managed with EVT are critical due to investor behaviour, as outlined by Prospect Theory from Kai-Ineman & Tversky (1979), which highlights that individuals are more sensitive to losses than gains. In times of market volatility or turbulence, economic agents often prioritise stability and shift towards less volatile assets to avoid potential losses. Conversely, during periods of market calm, when economic agents are experiencing gains, they become more willing to take on additional risks in pursuit of higher returns. By utilising Extreme Value Theory-Managed (EVTM) portfolios to dynamically adjust allocations based on market conditions, investors can optimise risk-adjusted returns while mitigating the negative impacts of volatility. This adaptive approach enhances portfolio resilience, balancing risk and reward across different market environments.

The study underscores the importance of integrating EVT with portfolio optimisation techniques, such as the Sharpe and Sortino ratios to improve performance by managing downside risk and maximising risk-adjusted returns. These findings are especially valuable for portfolio managers and risk professionals, as they provide insights into managing the risks associated with extreme market movements. They also offer credit risk professionals a

tool for assessing the impact of extreme events like defaults and downgrades. Overall, the results support adopting the EVT strategy in both single-factor and multi-factor portfolios, presenting a robust method for achieving superior financial outcomes amid market volatility and extreme events.

Chapter 4 is the final aspect of this research which analyses volatility forecasting as financial market volatility is a key indicator of financial market stability and investor confidence. Accurate forecasting of realised volatility is essential for effective risk management, portfolio optimisation, and policy formulation. This research examines the use of various models, including GARCH-type models and machine learning approaches such as Long Short-Term Memory (LSTM) networks to forecast realised volatility based on historical data.

The study places particular emphasis on the Hurst exponent, GARCH-type models, and LSTM networks for their respective strengths in capturing long-term dependencies, volatility clustering, and complex non-linear relationships in time-series data. The Hurst exponent, originally developed to analyse the presence of long-range dependence in hydrology, is employed here to assess the long-term memory of financial time-series, providing insights into market behaviour over extended periods. GARCH and EGARCH models are utilised to capture the volatility clustering and leverage effects observed in financial markets, while LSTM networks are leveraged for their ability to model and predict non-linear patterns in data that traditional models might neglect. We analyse the rough GARCH-type LSTM models with the DM and WS statistical tests, which test if one model is a better predictor of  $RV_t$  in comparison to another, and if there are significant differences in the models, respectively.

Eventually, we decompose the  $RV_t$  to low, mid, and high states for financial and political events using a three-state first-order Markov switching regression model. We map the identified volatility states to the event days and proceed with forecasting using the LSTM models. We compute the one day-ahead forecasts corresponding to or immediately following the event date. Subsequently, we evaluate the forecasting accuracy, by calculating the Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE).

By combining these models into a hybrid framework, the research aims to enhance the accuracy of realised volatility forecasts, as supported by empirical findings from Hu & Ni & Wen (2020) and Kim & Won (2018). The rough GARCH-type LSTM models developed in this study integrates roughness, GARCH and EGARCH outputs with LSTM networks, offering a novel approach to forecasting realised volatility in financial markets.

My main contribution in relation to Chapter 2 is analysing how low volatility is a strong predictor of financial crisis. By implementing various modelling approaches, and estimating the RND to examine the forecastability of crashes. Ultimately, we observe how the evolution of the tail-shape parameter and how the PDF conforms to Keynes (1937) and Minsky (1977) economic theories in how economic agents observing low financial risk tend to increase risk-taking, leading to a crisis.

The primary contribution of Chapter 3 lies in its comprehensive analysis and validation of portfolio optimisation strategies that incorporate Extreme Value Theory (EVT) alongside traditional risk-adjusted performance measures such as the Sharpe ratio and Sortino ratio. The Chapter demonstrates that portfolios managed using Extreme Value Theory optimised

by Sharpe and Sortino methods (referred to as EVT<sub>M</sub> portfolios) consistently outperform Volatility-Managed (VM) portfolios from Moreira & Muir (2017) and unmanaged portfolios across several key performance metrics. The study provides compelling evidence that EVT<sub>M</sub> portfolios, particularly those optimised using the Sortino ratio, achieve significantly higher risk-adjusted returns compared to VM portfolios. The Sortino-optimised EVT<sub>M</sub> portfolios, for instance, exhibit a Sortino ratio 3.46 times higher than that of VM portfolios, underscoring the effectiveness of incorporating downside risk into portfolio management. By employing Conditional Value at Risk (CVaR) as a measure of extreme downside risk, the analysis reveals that EVT<sub>M</sub> portfolios manage extreme risks more effectively than VM portfolios at high confidence levels (99% and 95%). This is particularly significant during extreme market events, where the EVT<sub>M</sub> strategy shows a more favourable risk-return profile, as indicated by positive skewness and more moderate kurtosis compared to VM portfolios. The study also addresses the practical aspect of transaction costs, showing that despite higher break-even transaction costs associated with EVT<sub>M</sub> portfolios, these portfolios maintain superior performance metrics, stemming from higher Sharpe ratios and alphas. This highlights the resilience of EVT<sub>M</sub> strategies to transaction costs, making them a viable option in real-world portfolio management. Overall, this Chapter makes a significant contribution by advocating for the integration of EVT into portfolio optimisation, demonstrating its value in enhancing risk management and volatility-timing, and achieving superior financial outcomes, particularly in the context of market volatility and extreme events.

Regarding Chapter 4, my contribution is the implementation of novel rough hybrid long short-term memory models (rGARCH-LSTM, rEGARCH-LSTM, and rGE-LSTM) designed for

forecasting realised volatility, integrating roughness, the LSTM model, and various Generalised Autoregressive Conditional Heteroscedasticity (GARCH)-type models. These hybrids are compared against standard models like GARCH-LSTM and EGARCH (exponential GARCH)-LSTM models. Our findings reveal that roughness, combined with the GARCH model and/or EGARCH model, within an LSTM framework, exhibits significant predictions. The chapter's primary contribution is on the emphasis of rough GARCH-type LSTM models, which learn sequential patterns and enhances prediction accuracy in stock market realised volatility by integrating a neural network model with multiple econometric qualities.

We assess the rough GARCH-type LSTM models in the context of macroeconomic events. We evaluate various loss functions computed on three volatility states (low, medium, and high volatility states) generated from a three-state first-order Markov switching regression model. The analysis confirms rough GARCH-type LSTM models prove to outperform non-rough LSTM types, and this comparative analysis reveals an increase in forecasting errors associated with macroeconomic events. Therefore, we show empirical support to the phenomenon of volatility clustering, which implies that high volatility affects and increases forecasting errors.

## Chapter 2

# AN EX-ANTE AND EX-POST ANALYSIS OF OPTION-IMPLIED PROBABILITY DENSITY FUNCTIONS IN TIMES OF MACROECONOMIC UNCERTAINTY

## Abstract

An option-implied Probability Density Function (PDF) provides estimates of future movements in an asset's price as priced by investors. PDFs are significant because they are forward-looking and based on market prices. In this Chapter, I compute PDFs using option prices. This approach is comprehensive, as it incorporates information from the entire distribution of investor beliefs, rather than relying solely on the mean expectation. Therefore, PDFs can explain a range of future possibilities. I explore uncertainty by examining proxy Structural Vector Autoregressive (SVAR) models to analyse uncertainty shocks and the tail shape parameter to observe the rate of decay in the distribution's tails. Ultimately, I estimate the empirical relationship between ex-ante uncertainty and the ex-post resolution of uncertainty in financial markets. I implement both non-parametric and parametric methods to construct the PDFs and subsequently observe the statistical results and the economic impact of uncertain events on the shape of the PDF in its ex-ante and ex-post states.

In conclusion, my contribution lies in revealing how macroeconomic uncertainty affects the evolution of option-implied PDFs in ex-ante and ex-post states - which could be driven by a change in investors' risk aversion, and how the analysis of a lagged TLM probit model supports the proposition that low volatility encourages excessive risk-taking behaviour.

## Abbreviation

<b>PDF</b>	Probability Density Function
<b>SDE</b>	Stochastic Differential Equation
<b>PDE</b>	Partial Differential Equation
<b>EW</b>	Edgeworth Density
<b>ATM</b>	At-the-money
<b>ITM</b>	In-The-Money
<b>OTM</b>	Out-The-Money
<b>BSM</b>	Black-Scholes-Merton
<b>GB2</b>	Generalised Beta of the Second Kind (GB2)
<b>RND</b>	Risk-Neutral Density
<b>TLM</b>	Tail Loss Measure
<b>SVAR</b>	Structural Vector Autoregressive Model
<b>VAR</b>	Vector Autoregressive Model
<b>MAE</b>	Mean Absolute Error
<b>EVT</b>	Extreme Value Theory

**CDF** Cumulative Distribution Function

**GPD** Generalised Pareto Distribution

## Nomenclature

$t$	Current time
$T$	Maturity time of option
$T - t$	Time left till option expiry
$r$	Risk-free interest rate
$S_t$	Asset price at current time
$C_t$	Call option price at current time
$P_t$	Put option price at current time
$S_T$	Asset price at maturity of option
$K$	Strike price
$q_t$	Conditional risk-neutral probability density of the underlying price $S$ at time $t$
$Q_t$	Conditional risk-neutral probability distribution function of $S_t$
$\mathcal{L}(a_i, b_i, S_T)$	Is the $i^{\text{th}}$ log-normal density function with parameters $a_i, b_i$
$\mathbf{B}(p, q)$	Is the Beta function of the GB2, and $p$ and $q$ are shape parameters controlling the behaviour of the tails
$\mathbf{y}_t$	SVAR vector

- $I$  Is the proxy SVAR impact matrix that relates reduced-form innovations to structural shocks
- $\epsilon_t$  Is the proxy SVAR vector of structural shocks
- $\mathbf{A}(L)$  Is a lag matrix polynomial capturing the autoregressive component of the model
- $\xi$  Tail shape parameter

## 2.1 Introduction

The evolution of option-implied probability density functions (PDFs) during periods of macroeconomic uncertainty offers a unique window into market expectations and risk perception. This chapter examines the dynamic evolution of these PDFs by analysing three distinct states, the ex-ante, current, and ex-post states within a risk-neutral framework. Under the risk-neutral measure, the current price of a security reflects the discounted expectation of its future prices, making risk-neutral probabilities an essential tool for estimating fair option prices. A fundamental assumption underpinning this framework is the absence of arbitrage, as initially formalised by Cox & Ross (1976) and Ross (1976).

Building on the seminal work of Breeden & Litzenberger (1978) and Banz & Miller (1978), who demonstrated that the second derivative of the European call price with respect to the strike price yields the discounted risk-neutral density, this research advances the literature by focusing on how macroeconomic uncertainty shapes option-implied PDFs. This theoretical foundation not only underpins our analysis but also highlights the importance of extracting and observing the entire distribution of future asset prices rather than relying solely on point estimates.

Motivated by the need for robust risk management and to assess pricing models, our study employs both non-parametric and parametric methods to construct option-implied PDFs. We use the Diebold-Mariano test (Diebold & Mariano (1995)) to compare the density forecasting accuracy of different methods, finding that a mixture of log-normals offers superior

performance in estimating European call option prices.

Furthermore, by examining the empirical relationship between ex-ante uncertainty, current states, and the subsequent ex-post resolution of uncertainty (driven by scheduled and unscheduled macroeconomic news, which are economic and political announcements), this research bridges a gap between macroeconomic risk and market dynamics. A notable contribution of our work is estimating and decomposing the Tail Loss Measure (TLM), which captures the severity of extreme market downturns. Using TLM within a probit regression framework, we demonstrate that periods of low volatility can paradoxically encourage excessive risk-taking, a finding that aligns with recent studies Hamidieh (2014), Vilkov & Xiao (2013) and Danielsson & Valenzuela & Zer (2018).

In summary, our major contributions include; providing a detailed analysis of the evolution of option-implied PDFs under macroeconomic uncertainty, empirically linking ex-ante uncertainty, and current states, to ex-post market outcomes, and offering insights into how low volatility environments may foster conditions conducive to financial crises.

This work not only builds on foundational theories in risk-neutral pricing and option valuation but also contributes novel empirical evidence on the interplay between macroeconomic uncertainty and market risk as reflected in option markets.

This Chapter is structured as follows: in the next section, Section 2.2, we present the literature review and discuss several methods for recovering risk-neutral densities, and the TLM. Our results show that the mixed log-normal distribution is the best method for estimating SP500 European call options. We discuss the density accuracy test which is based on a comparative

analysis using the Diebold-Mariano (DM) test from Diebold & Mariano (1995).

In Section 2.3, the data are presented. The dataset consists of CBOE European options prices, the interest rate is the 3-Month Treasury Constant Maturity Rate (DGS3MO), public debt as a percentage of gross domestic product is sourced from the Federal Reserve Economic Data (FRED), and volatility surface data is obtained from OptionMetrics and is used to compute the evolution of the tail shape parameter.

The methodology section, Section 2.4, offers a detailed explanation of the procedures and techniques employed to recover and compare risk-neutral densities, evaluate uncertainty shocks using gold as a proxy, analyse the evolution of the tail shape parameter, decompose volatility into high and low volatility states, and compute the Probit regression model.

In Section 2.5, we analyse the ex-ante and ex-post PDFs along with statistical analysis. Following Piffer & Podstawski (2018), gold is used as the most favourable safe-haven asset for constructing the proxy. We conduct statistical analysis within this section, and assess the use of gold as a safe-haven proxy to determine whether it provides more significance in understanding the economic and statistical properties of financial and political events. We estimate the Tail Loss Measure and discuss its evolution, and decompose volatility into high and low states using a probit regression model. Our aim here is to examine the relationship between volatility, as measured by the tail shape parameter, and the probability of a future crisis.

Finally, in Sections 2.6, we conclude with our findings which infer that economic impact is the highest if the economy stays in the low volatility environment.

## 2.2 Literature Review

### 2.2.1 Construction of PDF

The adoption of PDFs to predict events or the probability of extreme movements has been explored in various ways. PDFs implied from option prices, used to assess market expectations about future uncertainty, asymmetry, and the probability of extreme movements in asset prices, are discussed in Lynch & Panigirtzoglou (2008). Bahra (1997) shows how the information contained in implied RND functions can supplement forward-looking information available to policy-makers, aiding in assessing monetary operations and identifying anomalous market prices. Jackwerth (2000) advance this framework by addressing several practical challenges in extracting RNDs from observed option prices. Their methods allow researchers and practitioners to infer the market's risk-neutral expectations about future asset prices providing valuable insights into investor sentiment, risk premia, and the pricing of risk. The shape of the RND, which illustrates its skewness, kurtosis, and tail behaviour offers information on how the market perceives extreme events and uncertainty.

We explore the dynamics of an option-implied PDF by observing macroeconomic uncertainty, examining the states prior to the event, on the day of the event, and after the event to explain the significance of the uncertainty. The events are categorised as financial or political news to differentiate which type of news, whether anticipated or unanticipated, has a greater impact on or after the event. It is shown that financial news has more of an impact on the PDFs that are recovered. We also explore several types of RNDs to determine which method is the most preferred in producing the risk-neutral density used to estimate the option-implied density.

### 2.2.2 Risk-Neutral Density Estimation Methods

RNDs are categorised into three groups: parametric, semi-parametric, and non-parametric. We focus on parametric and non-parametric estimation methods. The estimation methods we investigate include Shimko's spline method, Edgeworth's expansion, the log-normal Black-Scholes-Merton model, a mixture of two log-normals, and the generalised beta of the second kind.

Shimko's spline method is computed by smoothing the volatility smile. Bu & Hadri (2007) explain the impact of this method on estimating the RND, particularly within the tails. Smoothing the volatility smile is preferred over interpolating the call pricing function because it is more challenging to accurately fit the shape when interpolating the call pricing function. Moreover, small errors from interpolating the call price function tend to have significant effects on the RND.

Jarrow & Rudd (1982) proposed a method to value European options. Their method is an option pricing formula derived from an Edgeworth series expansion of the distribution  $F$  about an approximating distribution  $A$ . Given the ubiquity and prominence of the log-normal distribution in option pricing theory and practice, Jarrow & Rudd (1982) suggest the log-normal distribution as a suitable approximating distribution.

Jackwerth (1999) discusses how mixture methods achieve greater flexibility in producing different probabilities from several simple distributions. A limitation of the mixture procedure

is how the number of parameters increases exponentially. This is explained by how mixing three log-normal distributions results in eight parameters, as two parameters are used for each log-normal and two mixing probabilities are added. Melick & Thomas (1997) use mixtures of three log-normal distributions and apply their method to American options on crude oil futures, accounting for the optionality of early exercise. Ritchey (1990) analyses option prices where the risk-neutral distribution of the log returns is a mixture of normal distributions. Vitiello & Poon (2014) present closed-form European option pricing formulas within a general equilibrium framework for assets following an  $N$ -mixture of transformed normal distributions. These component distributions, while not needing to be from the same class, must all be transformed normal. They prove this mixture can explain the seemingly abnormal non-monotonic pricing kernel for the S&P 500, while maintaining a logically consistent monotonic decreasing marginal utility for the representative agent. The authors demonstrate that a mixture of two log-normal distributions is sufficient to generate this outcome, including various implied volatility smile patterns.

Bookstaber & McDonald (1987) introduce the adoption of an extremely flexible generalised distribution, the GB2 distribution. This distribution encompasses a wide array of well-known distributions. The advantage of this flexibility is that it allows for the representation of different degrees of fat tails in the distribution.

We conclude that the mixed log-normal distribution is the preferred method for estimating S&P 500 European call options. This conclusion is based on a comparative test using the Diebold-Mariano (DM) test from Diebold & Mariano (1995) detailed in the Appendix A.1. The uniqueness of the DM test lies in its capacity to evaluate the best predictive model when

multiple density forecasts are available. This result is also supported by Liu et al. (2007), who compare parametric densities with more flexible non-parametric spline densities. Liu et al. (2007) parametric method consisted of the mixture of log-normals and the generalised beta of the second kind, while they also computed a non-parametric spline method. Their results imply that parametric densities provide more accurate predictive densities for real-world observed index levels. Parametric densities also prove to be more accurate, as the observed likelihoods are higher than those of the spline densities for both risk-neutral and real-world density comparisons. Liu et al. (2007) made comparisons of log-likelihoods, ranked cumulative probabilities, and the moments of sets of densities.

### **2.2.3 Impact of Macroeconomic Uncertainty**

Piffer & Podstawski (2018) highlight the use of gold as the most favourable safe-haven asset for constructing a proxy. We explore using gold as a safe-haven proxy to assess whether using gold as a proxy would show more significance in understanding the economic and statistical properties of uncertainty shocks. Finally, we observe the impact of uncertainty before, during, and after the event.

According to Hamidieh (2014) and Vilkov & Xiao (2013), the estimation of the Tail Loss Measure (TLM) allows one to observe a portion of the probability density function instead of examining the full distribution of asset prices. Here, we concentrate on the left tail alone; the TLM allows us to use the left tail to capture the severity of a crash instead of relying on an estimated probability. The construction of the tail loss measure involves the use of

Extreme Value Theory (EVT). EVT centres on the distributional properties of extreme, low probability, or uncertain events.

We take the TLM as a volatility measure a step further by applying Probit regression models to assess a fundamental notion in economics. The proposition is whether low volatility explains financial crises. To examine if low volatility is a strong predictor of a crisis, we decompose volatility into high and low components following the methodology of Danielsson & Valenzuela & Zer (2018). Danielsson & Valenzuela & Zer (2018) construct historical volatility for 60 countries, some dating back to 1800, using monthly stock prices as daily prices are scarce. In this Chapter, volatility based on the TLM is derived from daily S&P 500 prices. Another difference in this Chapter is that we use a Probit regression model, as opposed to Danielsson & Valenzuela & Zer (2018), who use a panel-Logit regression, which better accounts for the panel data composed of a large number of countries. Ultimately, high and low volatilities are defined as deviations of volatility from above and below its trend, respectively. The cycle and trend is estimated through a one-sided Hodrick & Prescott (1997) filter.

#### 2.2.4 Economic Theory and Interpretation

Jackwerth (2000) assert the changes in risk-neutral PDFs may be driven by shifts in future earnings expectations and or alterations in investors' risk aversion. Jackwerth (2000) identifies a significant departure in post-crash S&P 500 Index option pricing from the Black-Scholes formula, which presents as a left-skewed, highly leptokurtic risk-neutral distribution for the future underlying asset price. This deviation prompts the question of its underlying economic causes. Jackwerth (2000) offers four potential explanations. First, the leverage effect suggests that when stock prices fall, a firm's debt-to-equity ratio rises because the denominator decreases faster than the numerator, thereby amplifying the impact of asset returns on stock returns and increasing volatility; this effect influences both individual stock options and index options, the latter being convex combinations of constituent stock returns. Second, the correlation effect posits that significant declines in stock index prices lead to higher correlations among individual stock returns, reducing the benefits of diversification and elevating overall volatility, as observed during the 1987 market crash. Third, the wealth effect contends that falling stock prices diminish investors' wealth, increasing their risk aversion and causing them to react more strongly to new information, which in turn boosts volatility. Finally, the risk effect reverses the causality of the wealth effect by suggesting that an exogenous increase in market risk leads risk-averse investors to demand a higher expected return, thereby lowering current stock prices.

Danielsson & Valenzuela & Zer (2018) support the view that economic agents change their

risk-taking behaviour in times of uncertainty. This notion has a long history in economic literature. Low volatility can lead to excessive credit buildups and balance sheet leverage in the financial system; this implies that agents take more risks during periods when they perceive risk to be low, supporting the maxim that “stability is destabilising.”

This theory is supported by Minsky (1977) instability hypothesis, which asserts that economic agents observing low financial risk are inclined to increase risk-taking, ultimately leading to a crisis. Furthermore, discussions of the adverse effects of low volatility on financial stability are consistent with Brunnermeier & Sannikov (2014) “volatility paradox.” This theory argues that low volatility can paradoxically increase the probability of a systemic event. Additionally, Bhattacharya et al. (2015) examine Minsky’s hypothesis in a model with endogenous defaults, where economic agents update their optimistic expectations during good times, thus increasing risk-taking.

It is conventional to assume that during low-volatility periods, excessive risk-taking connotes overoptimism. Hence, as suggested by Danielsson & Valenzuela & Zer (2018), during quiet market periods when perceived risk is low, economic agents may behave irrationally and take excessive risks. Excessive lending and leverage can incentivise a tendency to take on even more risk.

Danielsson & Valenzuela & Zer (2018) investigate whether low volatility predicts crises by decomposing volatility into high and low components. Their primary contribution demonstrates that low volatility is a strong predictor of financial crises. Moreover, they show that low volatility over a prolonged period can lead to increased risk-taking, measured by

high credit-to-GDP ratios and financial sector leverage. As a result, low volatility encourages greater risk-taking, which progressively leads to riskier investments. Over time, loan losses increase, eventually culminating in a crisis.

### 2.3 Data Description

The data consist of CBOE bid and ask prices for European call and put options on the S&P 500 Index, sourced from OptionMetrics. The data range spans from January 4th, 1996, to December 31st, 2019. The criteria for selecting option expiration dates include expirations approximately within one month. The midpoint, which is the average of the bid and ask prices for each option, is taken to obtain a single price. All puts with a bid price of zero are removed. The S&P 500 Index option prices and dividend yield data are downloaded from the OptionMetrics database.

Table 2.1 depicts key descriptive statistics for the option best bid and ask prices, implied volatility (IV), and Delta. In our dataset, the average call option bid price is \$292.68 and the average ask price is \$296.78. The dispersion around these averages is high, with standard deviations of \$346.27 for the bid and \$348.81 for the ask. For call options, the implied volatility averages 27%. Notably, the distribution exhibits a high positive skew (skewness = 3.5885), indicating a pronounced long right tail where extreme high volatility values occur. Additionally, a kurtosis of 19.44 suggests that the tails are extremely heavy compared to a normal distribution. Regarding put options, the implied volatility averages 26%. The distribution for puts shows a moderate positive skew (skewness = 1.5192) and a kurtosis

Table 2.1: Summary Statistics for Options Data

Index		Mean	Median	Standard Dev	Min	Max	Skewness	Kurtosis
SPX Option Price - Call	Best Bid	292.6847	184.5000	346.2690	0.0500	2613.7000	1.9680	4.8341
	Best Offer	296.7845	188.8000	348.8089	0.0500	2620.8000	1.9510	4.7550
	Delta	0.6667	0.8721	0.3701	0.0014	0.9992	-0.7883	-1.0418
	Implied Volatility	0.2712	0.1952	0.2479	0.0503	2.9929	3.5885	19.4399
SPX Option Price - Put	Best Bid	55.7253	6.0000	113.7725	0.0500	1859.4000	3.9058	22.6844
	Best Offer	57.5297	6.5000	115.9400	0.1000	1863.3000	3.8263	21.7046
	Delta	-0.2785	-0.0779	0.3486	-0.9990	-0.0003	-1.0226	-0.5702
	Implied Volatility	0.2552	0.2250	0.1426	0.0277	1.6948	1.5192	3.9366

This table presents the summary statistics for our call and put option dataset. It comprises 89,292 call option observations and 86,709 put option observations, all with a 30-day maturity from January 4th, 1996, to December 31st, 2019.

of 3.94, reflecting moderately heavy tails relative to a normal distribution. Together, these summary statistics provide a robust foundation for understanding the market conditions captured in the dataset.

The interest rate used is the 3-Month Treasury Constant Maturity Rate (DGS3MO), which is downloaded from the Federal Reserve Economic Data (FRED).

The volatility surface data are obtained from OptionMetrics and are used to compute the evolution of the tail shape parameter. The descriptive statistics for the tail shape parameter is shown in Table 2.5. Total public debt as a percentage of gross domestic product ( $\Delta PD/GDP$ ) is downloaded from the Federal Reserve Economic Data (FRED).

The collated events are a subset of a database of events initially identified by Bloom (2009) and later expanded by Piffer & Podstawski (2018). These dates span a horizon of years from 1996 to 2015.

## 2.4 Methodology

The methodology section provides a comprehensive description of the steps and techniques used to recover and compare risk-neutral densities, assess uncertainty shocks using gold as a proxy, analyse the evolution of the tail shape parameter, decompose volatility into high-volatility and low-volatility states, and the compute of Probit regression model.

### 2.4.1 Risk-Neutral Densities

There are several methods and distributions for estimating implied PDFs. PDFs are primarily obtained using parametric or nonparametric methods. In this section, we compare and contrast the statistical properties of various probability density functions of the S&P 500 Index. The advantage of PDFs is that they can be applied to practically any financial market, and they provide critical benefits over methods based solely on conventional time-series data.

Exploring the evolution of option-implied PDFs during periods of macroeconomic uncertainty is an exciting domain. The evolution can be analysed by observing three states: the ex-ante state, the current state, and the ex-post state of the PDF. We examine these states within a risk-neutral framework. In relation to the risk-neutral function, this is a PDF for which the security's current price is equal to the discounted expectation of its future prices. In other words, risk-neutral probabilities represent future outcomes adjusted for risk and are used to estimate fair option prices. A key assumption when computing risk-neutral probabilities is the absence of arbitrage.

Cox & Ross (1976) and Ross (1976) assert that in a dynamically complete, arbitrage-free market, and when applying the risk-neutral valuation principle, the price of a contingent claim  $v_t$  is given by the expected present value of the payoff, computed under the Risk-Neutral Density (RND),

$$v_t(S_t, K) = e^{-r(T-t)} \mathbb{E}_t^Q[h(S_T)], \quad (2.1)$$

where  $Q$  is the unique probability measure where absence of arbitrage exists in a complete market,  $h(s)$  is the general payoff,  $S$  is the underlying price,  $K$  the strike price,  $T$  the expiry date,  $T - t$  the time to maturity, and  $r$  the instantaneous risk-free interest rate for that maturity. Therefore, for a European call option, we have,

$$h_c(S_T) = \max[S_T - K, 0], \quad (2.2)$$

and for the European put,

$$h_p(S_T) = \max[K - S_T, 0]. \quad (2.3)$$

The price at time  $t$  of a European call option  $C_t(S_t, K)$  is expressed as follows:

$$\begin{aligned} C_t(S_t, K) &= e^{-r(T-t)} \mathbb{E}_t^Q[\max[S_T - K, 0]], \\ &= e^{-r(T-t)} \int_K^\infty (S_T - K) q_t(S_T) dS_T, \end{aligned} \quad (2.4)$$

where  $q_t$  denotes the conditional risk-neutral probability density of the underlying price  $S$  at time  $t$ . Therefore, the price of a European put option,

$$\begin{aligned}
P_t(S_t, K) &= e^{-r(T-t)} \mathbb{E}_t^Q[\max[K - S_T, 0]], \\
&= e^{-r(T-t)} \int_0^K (K - S_T) q_t(S_T) dS_T,
\end{aligned} \tag{2.5}$$

Differentiating the European call price from Equation 2.4 with respect to the strike price  $K$  yields:

$$\begin{aligned}
\frac{\partial C_t(S_t, K)}{\partial K} &= e^{-r(T-t)} \int_K^\infty q_t(S_T) dS_T, \\
&= -e^{-r(T-t)} [1 - Q_t(K)].
\end{aligned} \tag{2.6}$$

Here,  $Q_t$  denotes the conditional risk-neutral probability distribution function of  $S_t$ . Calculating the second derivative of the European call price specified in Equation 2.4 with respect to  $K$ , we finally obtain:

$$\frac{\partial^2 C_t(S_t, K)}{\partial K^2} = e^{-r(T-t)} q_t(S_T) |_{S_T=K}. \tag{2.7}$$

Equation 2.7 is the prominent result originally presented in Breeden & Litzenberger (1978) and Banz & Miller (1978). This equation states that the second derivative of the European call price function with respect to the strike is equal to the discounted risk-neutral probability density of  $S_T$ . A similar mathematical expression is derived for European put options.

An option-implied PDF provides estimates of future movements in an asset's price as perceived by investors. One advantage of using PDFs is that they do not require an extensive historical time series for accurate estimations.

## 2.4.2 Non-Parametric Methods

Non-parametric methods provide greater flexibility in fitting the risk-neutral distribution to option prices. The two methods we investigate are Shimko's spline method and the Edgeworth density.

### 2.4.2.1 Shimko's Spline Method

Shimko (1993) suggests a relatively simple way to interpolate the calculated implied volatilities for the same maturity and different exercise prices. To generate the estimated density, Shimko calculates implied volatilities using observed market prices. Thereafter, the implied volatilities are interpolated, generating a volatility curve. Furthermore, following Garman & Kohlhagen (1983) and Jondeau & Poon & Rockinger (2007), option prices are calculated for each point on the interpolated volatility curve, and finally, the estimated risk-neutral density is generated. The volatility structure is represented as a parabola of best least-squared fit and modelled by the following equation:

$$\sigma(K) = a_0 + a_1K + a_2K^2. \quad (2.8)$$

The smoothed volatility structure gives a value of  $\sigma$  for every  $K$ , where  $K$  is the strike price. Using the Black-Scholes equation,  $\sigma$  is then used to generate call option prices. The constant term in the spline  $a_0$  represents the baseline level of implied volatility,  $a_1$  is the linear term, capturing the slope or the first-order sensitivity of implied volatility with respect

to changes in strike price or moneyness. The quadratic term  $a_2$  accounts for the curvature or the second-order sensitivity, which is crucial for capturing the "smile" or "skew" in the implied volatility structure.

### 2.4.2.2 Edgeworth Density

Jarrow & Rudd (1982) and Jondeau & Poon & Rockinger (2007) employ the generalised Edgeworth series expansion to obtain an approximate option valuation formula. Jarrow & Rudd (1982) proposed a method to value European style options when the underlying security price  $S$  at option expiration follows a distribution,  $F$ , which is known through its moments. They derive an option pricing formula from an Edgeworth series expansion of the distribution  $F$  about an approximating distribution,  $A$ . Their analysis yielded several variations, but we present their simplest expression for an approximate option price.

$$\begin{aligned}
 C(F) = C(A) &- e^{-rt} \frac{\kappa_3(F) - \kappa_3(A)}{3!} \frac{da(K)}{dS_t} \\
 &+ e^{-rt} \frac{\kappa_4(F) - \kappa_4(A)}{4!} \frac{d^2a(K)}{dS_t^2} dS_t + \epsilon(K),
 \end{aligned} \tag{2.9}$$

In Equation 2.9 the term on the left,  $C(F)$ , represents the call option price computed using the stock price distribution  $F$ . On the right, the first term,  $C(A)$ , is the call price determined from a known or approximating distribution  $A$ . This is then adjusted by additional terms that depend on the cumulants of the actual distribution  $F$  and the approximating distribution

A. The cumulants of a distribution relate to its central moments as follows:

$$\kappa_2(F) = \mu_2(F), \quad \kappa_3(F) = \mu_3(F), \quad \kappa_4(F) = \mu_4(F), \quad \kappa_2(F) = \mu_2(F) - 3\mu_2^2(F), \quad (2.10)$$

where  $\mu_2$ ,  $\mu_3$ , and  $\mu_4$  are the second, third, and fourth central moments, respectively. The density function of the approximating distribution  $A$  is denoted by  $a(S_t)$ , where  $S_t$  is the stock price at option expiration. Its derivatives are evaluated at the strike price  $K$ . The remainder term,  $\epsilon(K)$ , extends the Edgeworth expansion by incorporating higher-order cumulants and their corresponding derivatives. Jarrow & Rudd (1982) argue that with an appropriate choice for  $A$  is typically the lognormal distribution, which is central to option pricing where the higher-order terms in  $\epsilon(K)$  can be considered negligible. In this case,  $C(A)$  reduces to the familiar Black-Scholes call price formula.

This formulation corresponds to the first option pricing approximation method proposed by Jarrow & Rudd (1982). Their approach involves selecting the approximating distribution  $A$  such that its second cumulant matches that of  $F$  (i.e.,  $\kappa_2(F) = \kappa_2(A)$ ), where econometric evidence, discussed in Jarrow & Rudd (1982) empirical results section, supports the preference for this first-order approximation. By omitting the remainder term  $\epsilon(K)$  in Equation 2.9, the call option price can be succinctly written as:

$$C(F) = C(A) + \lambda_1 Q_3 + \lambda_2 Q_4, \quad (2.11)$$

where the terms,  $\lambda_j$  and  $Q_j$ , above are defined as follows:

$$\lambda_1 = \gamma_1(F) - \gamma_1(A), \quad Q_3 = -(S_0 e^{-rt})^3 (e^{\sigma^2 t} - 1)^{3/2} \frac{e^{-rt}}{3!} \frac{da(K)}{dS_t}, \quad (2.12)$$

$$\lambda_2 = \gamma_2(F) - \gamma_2(A), \quad Q_4 = -(S_0 e^{-rt})^4 (e^{\sigma^2 t} - 1)^2 \frac{e^{-rt}}{4!} \frac{d^2 a(K)}{dS_t^2}. \quad (2.13)$$

In Equation 2.12 and Equation 2.13,  $\gamma_1(F)$  and  $\gamma_1(A)$  are skewness coefficients for the distributions  $F$  and  $A$ , respectively. Similarly,  $\gamma_2(F)$  and  $\gamma_2(A)$  are excess kurtosis coefficients. Skewness and excess kurtosis coefficients are defined in terms of cumulants by:

$$\gamma_1(F) = \frac{\kappa_3(F)}{\kappa_2^{3/2}(F)}, \quad \gamma_2(F) = \frac{\kappa_4(F)}{\kappa_2^2(F)}. \quad (2.14)$$

In Equation 2.12 and Equation 2.13, the terms,  $Q_3$  and  $Q_4$ , represent skewness and kurtosis deviations from log-normality.

### 2.4.3 Parametric Methods

#### 2.4.3.1 Log-normal Density Implied by the Black-Scholes-Merton Model

The Black–Scholes–Merton model is a differential equation used to price for options. We assume that the spot price  $S_t$  follows a geometric Brownian motion,

$$S_t = S_0 \exp(\alpha t + \sigma W_t), \quad (2.15)$$

where  $S_0$  is an initial condition at time  $t = 0$ ,  $\alpha$  and  $\sigma$  are constants, and  $W_t$  is a standard Brownian motion.

This process is log-normal because:

$$\log(S_t) = \log(S_0) + \alpha t + \sigma W_t, \quad (2.16)$$

is normally distributed. Ito's lemma implies that,

$$dS_t = \mu S_t dt + \sigma S_t dW_t, \quad (2.17)$$

with drift,

$$\mu = (\alpha + \sigma^2/2), \quad (2.18)$$

and a deterministic volatility  $\sigma$ .

After some solving, we arrive at the prominent Partial Differential Equation (PDE) that was originally proposed and solved by Black & Scholes (1973):

$$\frac{\partial V}{\partial t} + \frac{\partial V}{\partial S_t} r S_t + \frac{1}{2} \frac{\partial^2 V}{\partial S_t^2} \sigma^2 S_t^2 = rV. \quad (2.19)$$

From the Black–Scholes equation, a partial differential equation, one can deduce the Black–Scholes formula when solved using Call and Put boundary conditions. Thus, the Black–Scholes pricing formulas for the European call and put options are:

$$C_t(S_t, t) = S_t \Phi(d_1) - K e^{-r(T-t)} \Phi(d_2), \quad (2.20)$$

$$P_t(S_t, t) = K e^{-r(T-t)} \Phi(-d_2) - S_t \Phi(-d_1), \quad (2.21)$$

with,

$$d_1 = \frac{\ln(S_t/K) + (r + \sigma^2/2)(T-t)}{\sigma \sqrt{(T-t)}}, \quad (2.22)$$

$$d_2 = d_1 - \sigma \sqrt{(T-t)}, \quad (2.23)$$

and  $\Phi(\cdot)$  denotes the cumulative standard normal distribution. Feynman-Kac analysis provides the link between the typically numerical PDE approach and stochastic processes.

We calculate the transition density as log-normal:

$$f(S_T) = \frac{1}{S_T \sqrt{2\sigma^2(T-t)}} \exp \left[ -\frac{[\ln(\frac{S_T}{S_t}) - (r - \sigma^2/2)(T-t)]^2}{2\sigma^2(T-t)} \right]. \quad (2.24)$$

### 2.4.3.2 The Mixture of Two Log-normals

A mixture of two log-normal densities is a very popular parametric RND specification that was first proposed by Ritchey (1990). He assumes that the implied density function  $f(S_T)$ , of the underlying asset terminal price  $S_T$ , comprises a weighted sum of  $z$  individual log-normal density functions:

$$f(S_T) = \sum_{i=1}^z [\theta_i \mathcal{L}(a_i, b_i, S_T)], \quad (2.25)$$

where  $a_i = \ln(S_t) + (\mu_i - \sigma_i^2/2)(T - t)$  and  $b_i = \sigma_i \sqrt{(T - t)}$ . Therefore, where  $\mathcal{L}(a_i, b_i, S_T)$  is the  $i^{\text{th}}$  log-normal density function with parameters  $a_i, b_i$ :

$$\mathcal{L}(a_i, b_i, S_T) = \frac{1}{S_T b_i \sqrt{2\pi}} \exp\left(-(\ln(S_T) - a_i)^2 / 2b_i^2\right), \quad (2.26)$$

from Equation 2.26, we can see how the weighted sum of  $z$  individual log-normal density functions in Equation 2.25 relates to Equation 2.24.

Where we have a risk-free interest rate  $r$  that is constant and asset price is  $S_t$ , the price of the call is the discounted expected payoff:

$$C(K) = e^{-r(T-t)} \int_K^{\infty} f(S_T)(S_T - K) dS_T, \quad (2.27)$$

where  $f(S_T)$  is the risk-neutral probability density function of the asset price at time  $T$ .

Similarly, the payoff on a European put is  $\max[K - S_T, 0]$  and its current price is:

$$P(K) = e^{-r(T-t)} \int_0^K f(S_T)(K - S_T)dS_T. \quad (2.28)$$

Under the assumption that the probability density function is a mixture of two log-normals with weights  $\theta$  and  $(1 - \theta)$ , the equations for call and put prices can be expressed as:

$$C(K) = e^{-r(T-t)} \int_K^\infty [\theta \mathcal{L}(a_1, b_1, S_T) + (1 - \theta) \mathcal{L}(a_2, b_2, S_T)](S_T - K)dS_T, \quad (2.29)$$

$$P(K) = e^{-r(T-t)} \int_0^K [\theta \mathcal{L}(a_1, b_1, S_T) + (1 - \theta) \mathcal{L}(a_2, b_2, S_T)](K - S_T)dS_T. \quad (2.30)$$

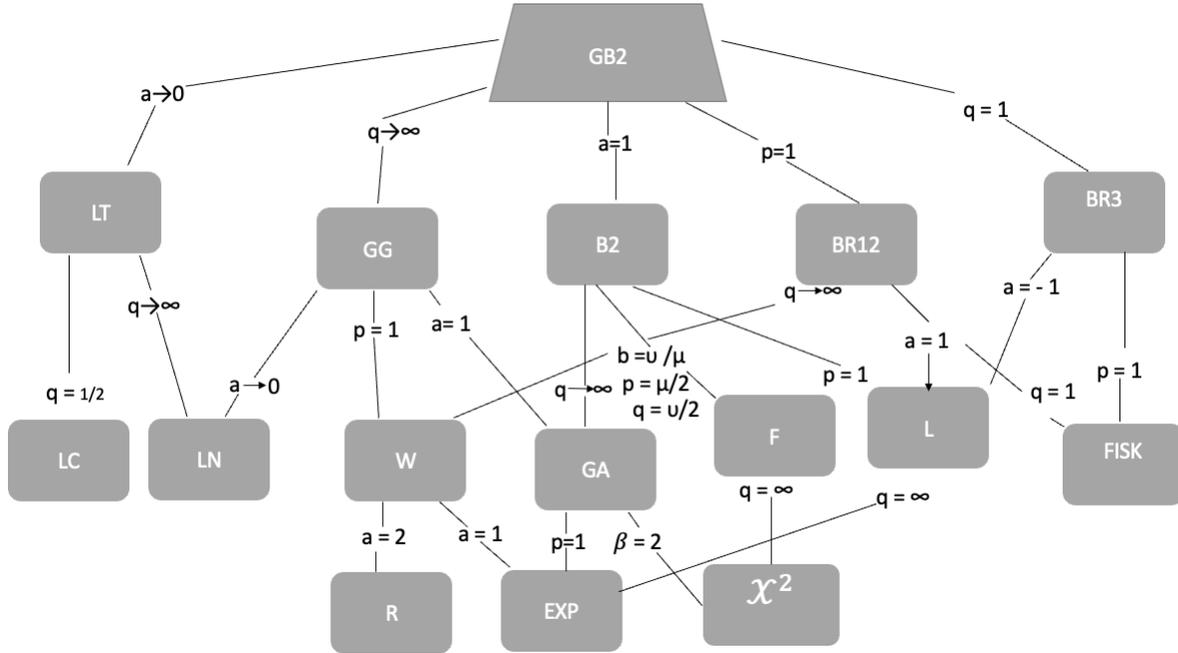
Gemmill & Saflekos (2000) represent  $\theta$  as the probability of an event with  $\theta \mathcal{L}(a_1, b_1, S_T)$  as the density conditional on this event, and  $(1 - \theta)$  as the probability of an event not occurring with  $(1 - \theta) \mathcal{L}(a_2, b_2, S_T)$  as the density conditional on the event not occurring.

### 2.4.3.3 Generalised Beta of the Second Kind (GB2)

The third parametric method we consider is the Generalised Beta of the second kind (GB2). In Figure 2.1, Bookstaber & McDonald (1987) illustrate that the GB2 is valuable because it is a flexible distribution that encompasses many other distributions.

Jondeau & Poon & Rockinger (2007) define the GB2 distribution as follows. The GB2 involves four positive variables, which are grouped into the vector  $\theta = (a, b, p, q)$ . In the equation below  $\mathbf{B}(p, q)$  is the Beta function which is a normalisation factor which controls the shape

Figure 2.1: GB2 Distribution Family Tree



The parameters are explained as the following: LT = log  $t$ ; B2 = Beta of the Second Kind; BR12 = Singh-Maddala or Burr type 12; BR3 = Burr type 3, LC = log Cauchy; LN = log-normal; W = Weibull; GA = gamma; L = Lomax; R = Rayleigh; EXP = exponential.

of the distribution,  $b$  is a scale parameter, and  $p$  and  $q$  are shape parameters controlling the behaviour of the tails. The  $a$  parameter determines how quickly the tails of the distribution approach the X-axis, where the higher the value for  $a$ , the quicker the distribution approaches the axis. The density of the GB2 is given by:

$$f(s|\theta) = \frac{as^{ap-1}}{b^{ap}\mathbf{B}(p, q)[1 + (\frac{s}{b})^a]^{(p+q)}} \quad \text{for } s > 0, \quad (2.31)$$

McDonald & Bookstaber (1991) show that if the RND is GB2, then the price of the European

call option is as follows:

$$C(K|\theta) = e^{-r(T-t)} \int_K^{\infty} (S_T - K)^+ f(S_T|\theta) dS_T, \quad (2.32)$$

from Equation 2.31, the price of the European call option becomes:

$$C(K|\theta) = e^{-r(T-t)} \int_K^{\infty} (S_T - K) \frac{aS_T^{ap-1}}{b^{ap}\mathbf{B}(p, q)[1 + (\frac{S_T}{b})^a]^{(p+q)}} dS_T. \quad (2.33)$$

#### 2.4.4 Comparing Forecasting Density Accuracy Using the Diebold-Mariano Test

The problem in selecting a preferred model arises when one is required to examine and evaluate predictive densities in the presence of multiple density forecasts. In empirical applications, it is often the case that two or more time-series models are available for forecasting a particular variable of interest.

To compare the forecasts using the DM test, the process is as follows:

- Actual values  $\{y_t; t = 1, \dots, T\}$
- Two forecasts  $\{\hat{y}_{1t}; t = 1, \dots, T\}, \quad \{\hat{y}_{2t}; t = 1, \dots, T\}$

The forecast errors are defined as:

$$e_{it} = \widehat{y}_{it} - y_t; \quad i = 1, 2. \quad (2.34)$$

The loss associated with forecast  $i$  is assumed to be a function of the forecast error  $e_{it}$ , and is denoted by  $g(e_{it})$ . The function  $g(\cdot)$  is a loss function and we let the associated forecast errors be  $e_{it}$ . The loss function will be a direct function of the forecast error, therefore,  $g(y_t, \widehat{y}_{it}) = g(e_{it})$ . We define the loss differential between the two forecasts by  $D_t = g(e_{1t}) - g(e_{2t})$  and state that the two forecasts have equal accuracy if and only if the loss differential has zero expectation for all  $t$ . However, the forecasting models must be non-nested.

To test the null hypothesis:

$$H_0 : E(D_t) = 0 \quad \forall t, \quad (2.35)$$

versus the alternative hypothesis:

$$H_1 : E(D_t) \neq 0. \quad (2.36)$$

The null hypothesis is that the two forecasts have the same accuracy. The alternative hypothesis is that the two forecasts have different levels of accuracy. In this case, we investigate the event where the alternate hypothesis;  $\widehat{y}_{1t}$  is more accurate than  $\widehat{y}_{2t}$ . From

Diebold & Mariano (1995), the DM test statistic for testing the null hypothesis is:

$$DM = \frac{\bar{D}}{\sqrt{\frac{2\pi\widehat{f}_D(0)}{T}}}, \quad (2.37)$$

where  $\bar{D} = \frac{1}{T} \sum_{t=1}^T [g(e_{1t}) - g(e_{2t})]$  and  $\widehat{f}_D(0)$  is an estimate of  $f_D(0)$ .

#### 2.4.5 Assessing Uncertainty Shocks

In this Subsection, we present a method to evaluate uncertainty shocks using gold within a proxy SVAR, along with a comprehensive list of unanticipated events identified by Bloom (2009).

Piffer & Podstawski (2018) identify a proxy SVAR model to evaluate the economic impact of uncertainty shocks. They propose a unique strategy to identify uncertainty shocks by expanding on the proxy SVAR methodology developed by Stock & Watson (2012) and Mertens & Ravn (2013) to identify structural VARs using external instruments. The proxy for the uncertainty shock is computed as a vector, taking a value equal to the percentage variation in the price of gold around the event when an event occurred, and equal to zero otherwise. By construction, the price of a safe-haven asset should capture the uncertainty-related component of the events. Following Piffer & Podstawski (2018), gold is used as the most reliable safe-haven asset for constructing the proxy. This proxy is chosen because of the Granger causality properties of gold with several measures of uncertainty: the proxy based on the price of gold causes several measures of uncertainty which suggests a high informational

content of uncertainty dynamics, additionally, the proxy based on the price of gold is more correlated with the VXO residuals (which are (the differences between the observed values of the VXO S&P 100 implied volatility index and the values predicted by a statistical model) from the VAR model estimated.

#### 2.4.5.1 The Proxy SVAR Model

The framework for the identification of structural VARs via external instruments is developed in Stock & Watson (2012) and Mertens & Ravn (2013). The reduced form model is given by:

$$\mathbf{y}_t = \delta + \mathbf{A}(L)\mathbf{y}_{t-1} + \mathbf{u}_t, \quad (2.38)$$

where  $\mathbf{y}_t$  is a  $k \times 1$  vector including the endogenous variables,  $\delta$  includes constant terms, and  $\mathbf{A}(L)$  is a lag matrix polynomial capturing the autoregressive component of the model. The reduced form shocks, captured by the  $k \times 1$  vector  $\mathbf{u}_t$ , are assumed to be linearly related to the underlying structural shocks through the equation:

$$\mathbf{u}_t = \mathbf{I}\epsilon_t, \quad (2.39)$$

where  $\epsilon_t$  is a  $k \times 1$  vector of structural shocks with a variance-covariance matrix normalised to the identity matrix, and  $\mathbf{I}$  is the impact matrix that relates reduced-form innovations to structural shocks.

Constructing the proxy for the uncertainty shock involves collecting an array of events that

are potentially affected by economic uncertainty in a way unrelated to other macroeconomic shocks. Subsequently, we use variations in the price of safe-haven assets around these events to update the proxy. We select a collection of events that generated or reduced uncertainty. These events are unanticipated and exogenous with respect to other relevant macroeconomic shocks. Specifically, we choose the events shown in Table 2.20 identified by Bloom (2009) through peaks in the VXO and exclude all events that are potentially related to other relevant macroeconomic shocks.

Following Piffer & Podstawski (2018), we define the proxy for the uncertainty shock as a vector that takes the value of the percentage change in the price of gold around the occurrence of an event, and zero otherwise. Specifically, we compute this proxy as the percentage variation in gold prices surrounding selected events. Given an event  $E_j$ , with  $j = 1, \dots, N$  (where  $N$  is the total number of events considered), let  $\gamma^j$  represent the time when event  $E_j$  became known to the market. For each event, we calculate  $\Delta p^j$  as the percentage change in the price of gold between the last available auction price before  $\gamma^j$  and the first available auction price after  $\gamma^j$ . Following the approach of C. D. Romer & D. H. Romer (2004), these  $N$  realisations of  $\Delta p^j$  are then aggregated into a monthly time series by summing the daily proxies within each month. Appendix A.2 discusses the details of the Structural VAR methodology, specifying the necessary assumptions to identify the proxy of the uncertainty shock and provides an economic interpretation. From the Appendix A.2 we assert the identified shock can be interpreted as an convolution between a change in uncertainty and a news shock, thus, a potential change in risk appetite. While a news shock can lead to a change in risk appetite, it is important to note a news shock is the event or information update, and the shift in risk

appetite is one possible reaction by economic agents.

#### 2.4.6 Estimating the Tail Loss Measure

This Subsection focuses solely on the tail of the distribution. The tail shape parameter for risk-neutral density functions is the key parameter that quantifies the rate at which the tail of a distribution decays; the larger the tail shape parameter, the higher the probabilities of extreme values. Hamidieh (2014) explains that recovering the tail shape parameter is based on an asymptotic theorem from extreme value theory, where the generalised Pareto random variable approximates the distribution of the excess of a random variable over a large threshold. When the option strike is large, the difference can be approximated by a generalised Pareto random variable, producing a pricing formula based on the first moment of the generalised Pareto random variable. A similar formula is obtained for put options. The moment term contains the tail shape parameter, allowing us to estimate the tail shape parameter.

The Generalised Pareto Distribution (GPD) standard Cumulative Distribution Function (CDF) is defined as follows:

$$\mathbb{P}(Y \leq y) = H_{\beta, \xi}(y) = \begin{cases} 1 - \left(1 + \xi \frac{y}{\beta}\right)^{-1/\xi} & \text{if } \xi \neq 0, \\ 1 - \exp(-y/\beta) & \text{if } \xi = 0, \end{cases} \quad (2.40)$$

where,  $\beta > 0$ ,  $y \geq 0$  when  $\xi \geq 0$ , and  $0 \leq y \leq -\beta/\xi$  when  $\xi < 0$ . Here,  $\beta$  and  $\xi$  represent the scale and the tail shape parameter, respectively.

The Out-of-the-Money (OTM) put option pricing formula is as follows:

$$P = e^{-rT} \mathbb{E}[\max(K' - S_T, 0)] = e^{-rT} \mathbb{E}[K' - S_T \mid S_T < K'] \mathbb{P}(S_T < K'). \quad (2.41)$$

Here  $K'$  will be a low strike value in the left tail of a OTM option. The notation  $K'$  is used to differentiate this strike from the strike  $K$  for the call options.

$$P^* = e^{-rT} \frac{\beta(K')}{1 - \xi} \mathbb{P}(S_T < K'). \quad (2.42)$$

A similar reasoning leads to the pricing formula for the call options.<sup>1</sup>

The uniqueness of the Tail Loss Measure (TLM) is that it does not identify the probability but the severity of a crash as a measure. Markose & Alentorn (2011) estimate the tail shape parameter using the generalised distribution for the entire density of asset returns, whereas Hamidieh (2014) method is applicable irrespective of the underlying distribution. Furthermore, Hamidieh (2014) shows that when using 10%, 15%, or 20% of the lowest strike puts, the pattern for the evolution of the tail shape parameter is the same regardless of the proportions used.

Therefore, the estimation of the tail shape is not dependent on the proportion of data used to generate the tail shape. For this analysis, we focus on using 15% of the lowest strike puts. Observing the tail shape reveals how the risk-neutral density's tail evolves, and tail shape provides valuable information about the hedging price against significant moves during

---

<sup>1</sup> $C^* = e^{-rT} \frac{\beta(K)}{1 - \xi} \mathbb{P}(S_T > K)$

tumultuous times.

### 2.4.7 Decomposing Volatility into High and Low States

This Subsection builds on the expansion of the tail shape parameter discussed in Subsection 2.4.6. To examine whether low volatility is a strong predictor of a crisis, we decompose volatility into high and low components using the methodology of Danielsson & Valenzuela & Zer (2018).

High and low volatilities are defined as the deviations of volatility from above and below its trend, where the trend is estimated through a one-sided Hodrick & Prescott (1997) filter. Following Danielsson & Valenzuela & Zer (2018), we apply a smoothing parameter of  $\lambda = 5000$  to decompose volatility into the trend and deviations from the trend, or in other words, the cycle. For consistency, we choose the tail shape parameter  $\xi$  with a delta of 15% to represent the volatility  $\xi\sigma$ . The volatility trend is obtained from the following optimisation problem:

$$\min_{\{\tau_t(\lambda)\}_{t=1}^T} \sum_{t=1}^T [\xi\sigma_t - \tau_t(\lambda)]^2 + \lambda \sum_{t=2}^{T-1} \{[\tau_{t+1}(\lambda) - \tau_t(\lambda)] - [\tau_t(\lambda) - \tau_{t-1}(\lambda)]\}^2. \quad (2.43)$$

To identify the high and low volatility time-series, we further separate the deviation of volatility  $\xi\sigma_t$  from its trend  $\tau_t$  into two components, high and low volatilities, denoted as

$\delta_t^{\text{high}}(\lambda)$  and  $\delta_t^{\text{low}}(\lambda)$ , respectively.

$$\begin{aligned} \delta_t^{\text{high}}(\lambda) &= \begin{cases} \xi\sigma_t - \tau_t(\lambda) & \text{if } \xi\sigma \geq \tau_t(\lambda), \\ 0 & \text{otherwise,} \end{cases} \\ \delta_t^{\text{low}}(\lambda) &= \begin{cases} \xi\sigma_t - \tau_t(\lambda) & \text{if } \xi\sigma < \tau_t(\lambda), \\ 0 & \text{otherwise.} \end{cases} \end{aligned} \tag{2.44}$$

#### 2.4.8 Probit Regression Model

A Probit model is a popular specification for a binary response model. In this case, we regress the binary crisis indicator  $C_t$ , which shows the occurrence of an event on day  $t$ , on various specifications of volatility. The indicator takes the value 0 for the day of the event and 1 for a day when no event occurred. We first analyse the impact of volatility on market uncertainty by including the level of volatility as the main regressor. We then investigate whether the impact of volatility on market uncertainty is asymmetric by considering the absolute value of the deviation of volatility from its trend  $|\xi\sigma_t - \tau_t|$  as a regressor. Finally, we examine the effect of high and low volatilities separately by including  $\delta^{\text{high}}$  and  $\delta^{\text{low}}$  as regressors. For daily  $t$ , we estimate the following Probit regression:

$$\text{Probit}(J_t) = \alpha \bar{J}_{t-1 \text{ to } t-L} + \beta \bar{\Gamma}_{t-1 \text{ to } t-L}(\lambda) + \varepsilon_t, \tag{2.45}$$

where  $\Gamma$  is one of  $\xi\sigma$ ,  $|\xi\sigma - \tau|$  or  $[\delta^{\text{high}} \ \delta^{\text{low}}]'$  variables.  $\lambda$  is the HP filter smoothing parameter,

and  $\varepsilon_t$  the error term. Instead of regressing the crisis indicator on lags of the explanatory variables, following Reinhart & Rogoff (2011), we implement backward-looking moving averages of explanatory variables over  $L$  lags, from  $t - 1$  to  $t - L$ .

The moving average variables are constructed as:

$$\bar{Z}_{t-1 \text{ to } t-L} = \frac{1}{L} \sum_{j=1}^L Z_{t-j}, \quad Z = C, \xi\sigma, |\xi\sigma - \tau|, \delta. \quad (2.46)$$

The choice of using lagged moving averages instead of lags is to reduce collinearity between the explanatory variables and to smooth out temporary volatility spikes.

## 2.5 Empirical Analysis

The empirical analysis section presents a thorough investigation in models discussed in the methodology. It includes ex-ante and ex-post RND estimations, the evolution of densities, the decomposition of the TLM, and the application of a Probit regression model to assess whether low or high volatility serves as a strong predictor of financial crises.

### 2.5.1 Comparing Parametric and Non-Parametric Methods in Recovering the Risk Neutral Density

In Subsections 2.4.3 and 2.4.2, parametric and non-parametric methods are discussed mathematically. In this Subsection we compare densities using the Diebold-Mariano (DM) test.

#### 2.5.1.1 Real Data Study

To compute the DM test, we randomly select a date to assess densities to predict actual S&P 500 call option prices. There are various loss functions that can be used. These methods include squared errors, absolute errors, squared proportional errors (useful if errors are heteroscedastic), and absolute scaled errors.

For illustrative purposes, the randomly selected date is the 8th of January, 2015, with an expiration date of 57 days and this date provides 166 observations. As discussed, we use this date to generate forecasts with a forecasting horizon of 1 day ( $h = 1$ ) to compare with actual call option prices of the 9th of January, 2015. Chen & Wan & Y. Wang (2014) note

Table 2.2: Diebold-Mariano Test Applied to Shimko’s Spline Method, Edgeworth Density, Black-Scholes-Merton Model, Mixture of Two Log-normals, and GB2 Forecasts of S&P500 European Call Prices.

Methods	DM Test Statistic
Shimko’s Spline Method vs. Edgeworth Density	-21.268**
Shimko’s Spline Method vs. Black-Scholes-Merton Model	-16.188**
Shimko’s Spline Method vs. GB2	-22.20**
Shimko’s Spline Method vs. Mixture of Two Log-normals	7.09
Edgeworth Density vs. Black-Scholes-Merton Model	-7.48**
Edgeworth Density vs. GB2	-13.08**
Edgeworth Density vs. Mixture of Two Log-normals	22.41
Black-Scholes-Merton Model vs. GB2	-11.63**
Black-Scholes-Merton Model vs. Mixture of Two Log-normals	16.47
Mixture of Two Log-normals vs. GB2	-22.48**

We compute the DM test to see whether  $\hat{y}_{1t}$  is more accurate than  $\hat{y}_{2t}$  in predicting  $y_t$ , where  $y_t$  is the S&P500 European call prices. We have 166 observations and a forecasting horizon of 1. Here, stars are only intended to flag levels of significance. If a p-value is  $\leq 0.05$  it is flagged with one star (\*), if a p-value is  $\leq 0.01$  it is flagged with two stars (\*\*), subsequently, if a p-value is  $\leq 0.001$  it is flagged with three stars (\*\*\*)

that both squared-error loss and absolute-error loss are symmetric around the origin point.

Additionally, larger errors are penalised more severely by squared-error loss. Therefore, we use an absolute-error loss function. The loss function is defined as  $g(e_{it}) = |e_{it}|$ , with a forecast horizon of 1.

Table 2.2 provides an information of the DM test applied to various methods for forecasting S&P 500 European option call prices. Based on the statistical significance, the results show that the Mixture of Two Log-normals performs best when tested against all other methods for forecasting S&P 500 European option call prices. When the Black-Scholes-Merton Model, Shimko’s Spline Method, and Edgeworth Density are tested against the Mixture of Two Log-normals we obtain a positive test statistic, however, the result is not significant. When

mixture of two Log-normals is tested against the GB2, we obtain a negative test statistic which is also statistically significant.

A more thorough investigation for the choice of RND selection is in the Appendix Table A.1, where we consider 5,000 density forecasts and show that the Mixture of Two Log-normals method yields a better density forecast in 68.58% of cases when compared with Shimko's Spline method, 55.23% of cases when compared with the GB2 method, and 93.58% of cases when compared with the Edgeworth method. In Appendix Table A.2, we compute the average computational time expense of the estimation methods. We show the average CPU time required by each method over approximately 5000 density computations and the results explain that the Edgeworth Density method requires the least amount of time to compute a density estimation, whereas the GB2 method is the most time-consuming.

## 2.5.2 Interpreting Densities

This Section elaborates on the full distribution of the RNDs using the mixture of two Log-normals.

### 2.5.2.1 Evolution of the Densities

To compare the evolution of densities for an uncertain event, we observe across a cross-section of specific dates. We examine the day before the uncertain event  $t_{i-1}$ , the day of the uncertain event  $t_i$ , and the day after the uncertain event  $t_{i+1}$ , in other words, the ex-ante, current, and ex-post dates, respectively. The range of dates helps to see how expectations of option prices change due to the behavioural changes of risk-seeking and risk-averse economic agents. This

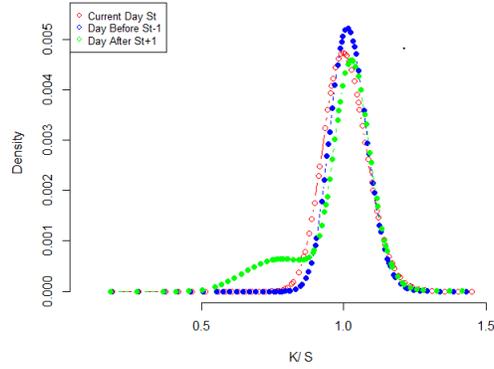
technique is used to illustrate the impact of uncertain financial events or policy announcements. If the day of observation (current day  $t_0$ ) falls on a weekend or public holiday, we use the next weekday that is not a public holiday as the new current day  $t_0$ . Below in Figure 2.2 and Figure 2.3, we present the superimposed RNDs of uncertain events, encompassing the day before the uncertain event  $t_{i-1}$ , the day of the uncertain event  $t_i$ , and the day after the uncertain event  $t_{i+1}$ . We choose a maturity of approximately 30 days.

From the observed dates, which include the 'Dodd-Frank Act signed by Obama', and 'Obama re-elected', in both cases, the location and shape of the PDFs have not significantly changed, as uncertainty is already priced into the PDF. However, for the 'US Government shutdown', the shape of the PDF on the day of the event and the ex-ante resolution has significantly changed. On the day before the event, we can see that the density has a wider distribution which implies that there's a larger spread or variance in the possible outcomes, which directly corresponds to higher volatility. Additionally, as the density is more spread out, it indicates that the market anticipates a broader range of future price levels. This broader range translates to more uncertainty because there's less consensus about where the price might settle. From the 'AIG asks for emergency lending + Lehman Brothers' and 'Northern Rock receives liquidity support by BoE' event, we can see that ex-post the event the market is pricing in greater uncertainty and there is increased probability of large price movements in the underlying asset. The current day density of the event, has a lower range of wider events than the ex-post, the ex-ante density has a narrower distribution, implying that uncertainty attached to this event has increased on and after the event.

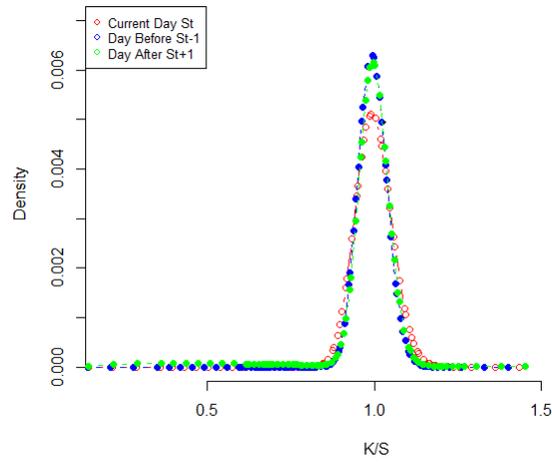
Furthermore, we acknowledge as discussed in the literature review economic theory Subsection

Figure 2.2: Evolution of Densities of Political Uncertainty

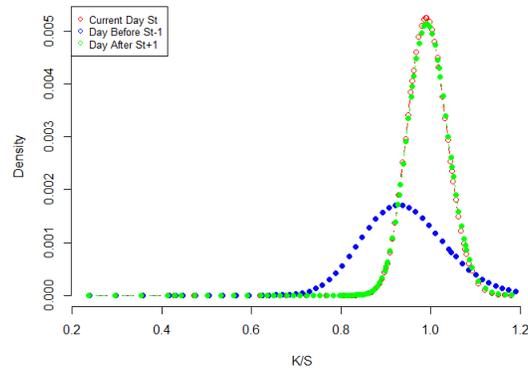
(a) Dodd-Frank-Act signed by Obama on 21<sup>st</sup> June, 2010



(b) Obama re-elected on 7<sup>th</sup> of November, 2012



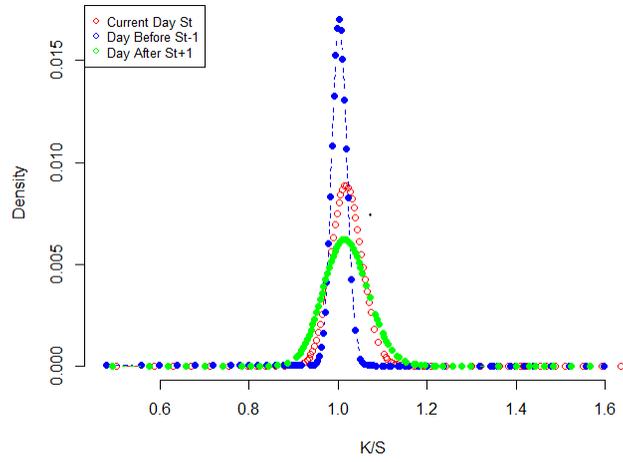
(c) US government shutdown on 1<sup>st</sup> October, 2013



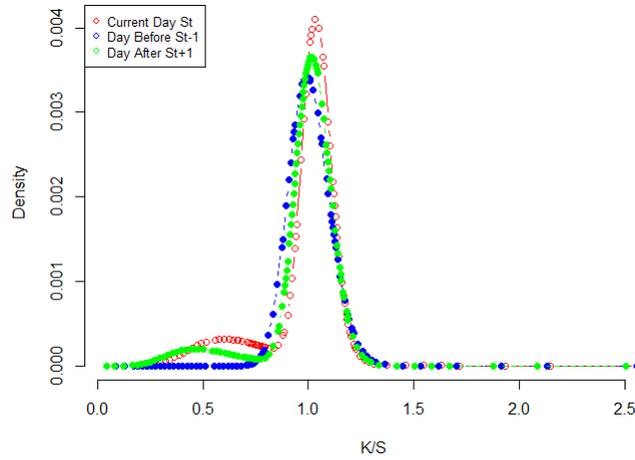
These illustrations show the mixture of two Log-normals densities of the day before political uncertainty  $t_{i-1}$ , on the day of political uncertainty  $t_i$ , and on the day after political uncertainty  $t_{i+1}$ .

Figure 2.3: Evolution of Densities of Financial Uncertainty

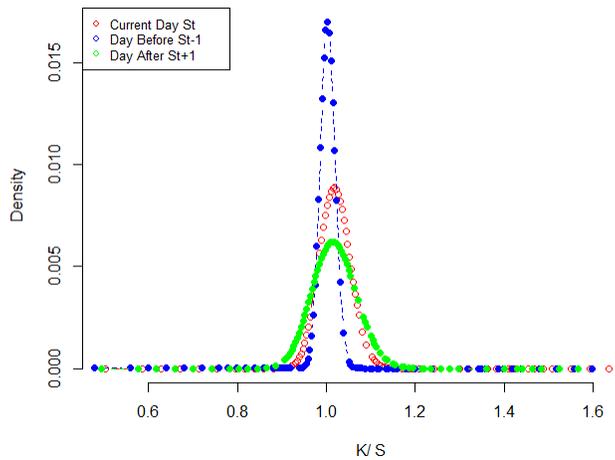
(a) AIG asks for emergency lending + Lehman Brothers on 15<sup>th</sup> September, 2008



(b) Swiss franc pegged on 6<sup>th</sup> September, 2011



(c) Northern Rock receives liquidity support by BoE on 14<sup>th</sup> September, 2007



These illustrations show the mixture of two Log-normal densities of the day before financial uncertainty  $t_{i-1}$ , on the day of financial uncertainty  $t_i$ , and on the day after financial uncertainty  $t_{i+1}$ .

2.2.4, the changes in risk-neutral PDFs may also be driven by shifts in future earnings expectations and or alterations in investors' risk aversion. This result aligns with early theoretical work, which suggests that risk affects economic decisions. This idea is explained in Keynes (1937) notion of animal spirits and Minsky (1977) instability hypothesis.

### 2.5.2.2 Statistical Analysis

The Mean Absolute Residual (MAR) is a common measure of forecasting errors in time-series analysis. Using the Mixture of Two Log-normals, we calculate the Mean Absolute Residual of the residuals for two events to examine how uncertainty affects the deviation of predicted prices from actual prices. The MAR gives an idea of the average error in the predictions. If the Mean Absolute Residual is small, it indicates that the model predictions are close to the actual market prices.

From Figure 2.4, it is shown that both ex-ante and ex-post the 'Dodd-Act Signing', both calls and puts had a smaller Mean Absolute Error with the exception of the ex-ante put (which is 0.03 higher than the put of day of the event), this implies that the event's date has a larger forecasting error due to uncertainty attached to this event. However, Figure 2.5 shows for 'Obama Re-election', the Mean Absolute Error was largest for both puts and calls ex-ante the announcement; this could mean that the anticipation of election results was more significant than learning who won the election on the actual election day. The ex-ante Mean Absolute Error is also high which also implies the uncertainty after the event.

Figure 2.4: Dodd-Frank-Act signed by Obama

(a) Dodd-Frank-Act signed by Obama on day  $t_{i-1}$

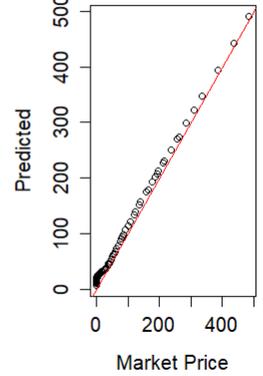
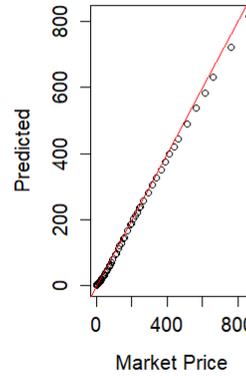
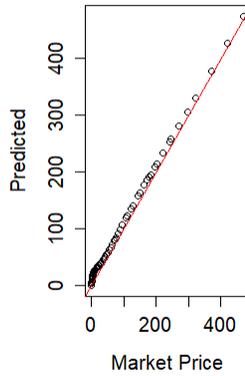
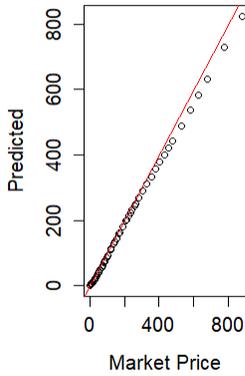
(b) Dodd-Frank-Act signed by Obama on day  $t_i$

MixLNorm, Calls, mean|res| = 3.46

MixLNorm, Puts, mean|res| = 3.72

MixLNorm, Calls, mean|res| = 4.32

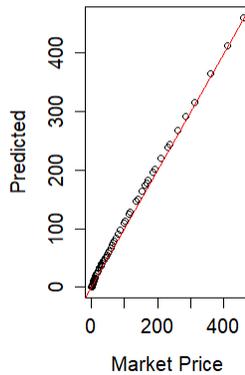
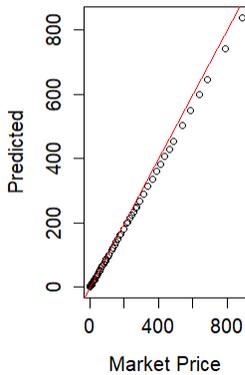
MixLNorm, Puts, mean|res| = 3.69



(c) Dodd-Frank-Act signed by Obama on day  $t_{i+1}$

MixLNorm, Calls, mean|res| = 2.34

MixLNorm, Puts, mean|res| = 3.43



This plot shows scatter plot comparing the predicted call and put option prices against the actual market call and put option prices using the Mean Absolute Residuals. We choose a maturity of approximately 30 days. The red line depicts the "perfect fit" line, where the predicted values equal the actual market values. Each point on the plot represents one option, with its market price on the x-axis and the corresponding predicted price on the y-axis.

Figure 2.5: Obama re-elected

(a) Obama re-elected on day  $t_{i-1}$

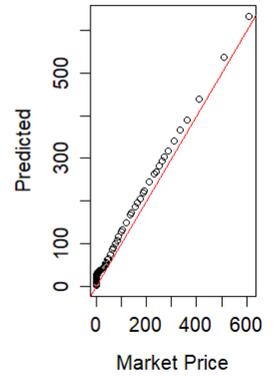
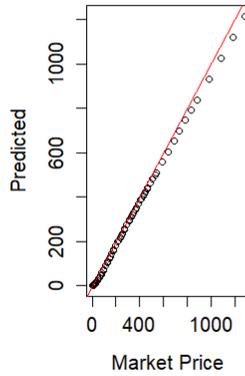
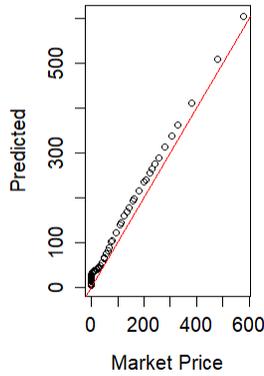
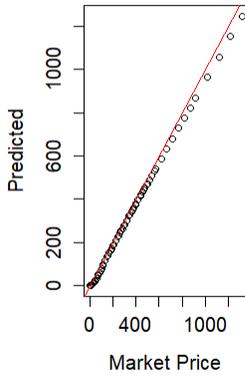
(b) Obama re-election day  $t_i$

MixLNorm, Calls, mean|res| = 5.13

MixLNorm, Puts, mean|res| = 5.12

MixLNorm, Calls, mean|res| = 4.7

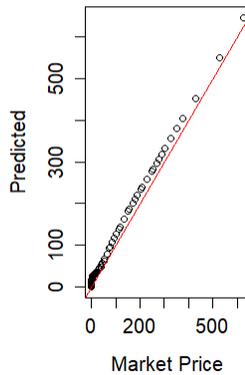
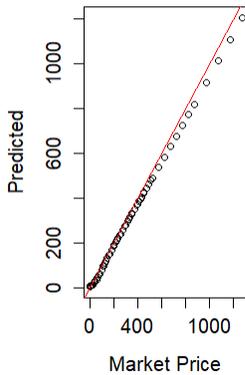
MixLNorm, Puts, mean|res| = 4.74



(c) Obama re-election day  $t_{i+1}$

MixLNorm, Calls, mean|res| = 4.78

MixLNorm, Puts, mean|res| = 5.48



This plot shows scatter plot comparing the predicted call and put option prices against the actual market call and put option prices using the Mean Absolute Residuals. We choose a maturity of approximately 30 days. The red line depicts the "perfect fit" line, where the predicted values equal the actual market values. Each point on the plot represents one option, with its market price on the x-axis and the corresponding predicted price on the y-axis.

### 2.5.2.3 Whole Dataset Analysis

For the complete dataset, the computed densities encompass all moneyness states (ITM, ATM and OTM), and this affects both Tables 2.3 and 2.4. The RNDs which are estimated from option prices across a range of strike prices are spread equally across all events. As shown in Table 2.4a, we can see that for political and financial events, both the mean and the variance are higher for the ex-post densities compared to the ex-ante densities. This variation also implies that there are more outliers in the data or larger data points on average after the uncertainty is observed. This observation shows how low volatility can paradoxically increase the probability of a systemic event. The ex-ante densities have higher skewness and kurtosis compared to the ex-post densities, indicating that the ex-ante density has heavier tails.

Generally, this also applies to the density with the gold proxy. For the gold proxy ex-ante financial density means, we compute the gold proxy SVAR discussed in Subsection 2.4.5.1, and the results are presented in Table 2.5a. The results show the ex-ante financial density variance and mean are lower than the ex-post density variance and means. This could imply that the surprise element is not anticipated before the event. Furthermore, the ex-post skewness and kurtosis of financial events are both negative, which implies that the tail on the left side of the distribution is longer or fatter in relation to the tail on the right side, and the shape of the curve is platykurtic. However, ex-ante mean and variance for political events are higher than the ex-post means and variance. Similarly to results presented in Table 2.4a, we also notice from Table 2.5a that the political ex-ante mean is higher than the financial ex-ante mean, which could imply that it is harder to capture political uncertainty.

For the complete dataset in both Tables 2.3 and 2.4, the mean of the ex-ante MAE, MAPE, and RMSE for both puts and calls are higher than the ex-post MAE, MAPE, and RMSE results. This implies that the forecasting errors decrease after the uncertain event.

Table 2.3: Ex-ante and Ex-post Calls and Puts of Financial and Political Events

Type of Event	Statistical Moments	Ex-ante Mean of Densities	Ex-post Mean of Densities
Financial	Mean	0.002	0.144
	Variance	0.000	0.282
	Skewness	4.857	3.635
	Kurtosis	26.874	14.553
Political	Mean	0.017	0.074
	Variance	0.014	0.142
	Skewness	7.402	5.471
	Kurtosis	57.839	32.179

(a) Statistical Moments of the Densities of Financial and Political Events

Type of Event	Statistical Moments	Ex-ante MAE Calls	Ex-post MAE Calls	Ex-ante MAPE Calls	Ex-post MAPE Calls	Ex-ante RMSE Calls	Ex-post RMSE Calls
Financial	Mean	36.3	34.1	13.8	12.0	44.0	41.6
	Variance	1369.1	1411.1	651.2	535.9	1852.8	1898.1
	Skewness	1.5	1.6	2.3	2.5	1.4	1.5
	Kurtosis	1.6	2.0	6.2	7.5	1.2	1.6
Political	Mean	41.5	38.8	20.9	17.2	49.8	46.6
	Variance	1452.8	1506.7	2250.7	924.4	1889.9	1976.4
	Skewness	1.4	1.5	4.6	2.243	1.3	1.3
	Kurtosis	1.4	1.4	28.0	5.6	0.8	0.8

(b) Statistical Moments of MAE, MAPE and RMSE of Calls of Financial and Political Events

Type of Event	Statistical Moments	Ex-ante MAE Puts	Ex-post MAE Puts	Ex-ante MAPE Puts	Ex-post MAPE Puts	Ex-ante RMSE Puts	Ex-post RMSE Puts
Financial	Mean	22.2	19.3	6.3	2.9	31.0	26.9
	Variance	277.7	247.0	222.1	12.0	753.6	665.8
	Skewness	1.0	1.1	4.2	1.6	1.3	1.4
	Kurtosis	0.5	0.4	21.8	2.1	1.1	1.5
Political	Mean	26.4	24.2	8.9	8.3	36.2	33.1
	Variance	455.1	474.0	269.6	273.0	1006.3	1049.2
	Skewness	1.6	1.8	3.1	3.9	1.4	1.6
	Kurtosis	2.3	3.7	10.4	18.6	1.5	2.7

(c) Statistical Moments of MAE, MAPE and RMSE of Puts of Financial and Political Events

These tables contain statistical moments of ex-ante and ex-post densities, ex-ante and ex-post density loss functions for calls, and ex-ante and ex-post density loss functions for puts.

Table 2.4: Ex-ante and Ex-post Calls and Puts of Financial and Political Events with Gold Proxy

Type of Event	Statistic	Ex-ante Mean of Densities	Ex-post Mean of Densities
Financial	Mean	$9.584 \times 10^{-4}$	$9.849 \times 10^{-4}$
	Variance	$5.0 \times 10^{-7}$	$5.3 \times 10^{-7}$
	Skewness	0.217	-0.090
	Kurtosis	-1.990	-1.213
Political	Mean	0.003	0.001
	Variance	$5.486 \times 10^{-5}$	$1.6 \times 10^{-7}$
	Skewness	2.460	-0.589
	Kurtosis	8.919	0.263

(a) Statistical Moments of the Densities of Financial and Political Events with Gold Proxy

Type of Event	Statistical Moments	Ex-ante MAE calls	Ex-post MAE calls	Ex-ante MAPE Calls	Ex-post MAPE Calls	Ex-ante RMSE Calls	Ex-post RMSE Calls
Financial	Mean	52.9	49.3	23.5	21.8	65.7	61.1
	Variance	1272.5	1302.7	632.8	441.4	1893.5	1959.8
	Skewness	0.4	0.5	0.5	0.4	0.5	0.5
	Kurtosis	-1.8	-1.9	-1.8	-1.8	-1.4	-1.5
Political	Mean	37.8	34.9	17.1	17.4	44.6	41.3
	Variance	1488.1	1642.8	779.6	817.8	1932.7	2138.0
	Skewness	1.7	1.7	1.5	1.4	1.6	1.6
	Kurtosis	4.2	4.1	2.2	1.6	3.5	3.5

(b) Statistical Moments of MAE, MAPE and RMSE of Calls of Financial and Political Events with Gold Proxy

Type of Event	Statistical Moments	Ex-ante MAE Puts	Ex-post MAE Puts	Ex-ante MAPE Puts	Ex-post MAPE Puts	Ex-ante RMSE Puts	Ex-post RMSE Puts
Financial	Mean	31.3	24.7	15.6	3.2	46.0	36.4
	Variance	168.6	226.5	825.8	3.9	628.7	657.4
	Skewness	0.2	0.6	1.7	0.5	0.3	0.6
	Kurtosis	-2.6	-1.7	5.8	-0.7	-1.9	-1.8
Political	Mean	22.8	18.5	23.3	14.0	31.0	26.3
	Variance	220.4	242.9	652.3	373.0	723.7	747.3
	Skewness	1.4	1.2	1.1	1.2	1.6	1.6
	Kurtosis	3.7	2.8	0.5	0.5	4.3	4.2

(c) Statistical Moments of MAE, MAPE and RMSE of Puts of Financial and Political Events with Gold Proxy  
 These tables contain gold proxy statistical moments of ex-ante and ex-post densities, ex-ante and ex-post density loss functions for calls, and ex-ante and ex-post density loss functions for puts.

### 2.5.3 Evolution of the Tail shape Parameter

To compute the option-implied measures, we follow the parallel interpolation process of Vilkov & Xiao (2013) for computing the TLM, using standardised volatilities with a maturity of 30 days from the Volatility Surface Data, which contains a smoothed implied volatility surface for a range of standard maturities and option delta points. We focus on using 15% of the lowest puts for this portion of the analysis. We select Out-of-the-Money implied volatilities for calls and puts: calls with deltas smaller than or equal to 0.5, and puts with deltas greater than  $-0.5$ . We smooth the delta 15% tail shape parameter using the Savitzky–Golay filter and compute the evolution of the tail shape parameter up until the end of 2019 to observe how the dynamics of the tail shape evolve over this period.

Figure 2.6 shows the full distribution of news events. From this Figure, it is possible to see that the spikes correspond to major news events. We specifically filter for events with a tail shape value above 34.6 which is less than 5% of the highest values to reduce the number of events for illustrative purposes. Here, financial news populates the figure more than political events, and it is observed that there is low volatility before these spikes. This figure implies that significant events result in a jump in the tail shape parameter. Around the year 2008, the figure shows the most volatility and uncertainty, with a section of the tail shape parameter reporting values close to zero.

Now, we no longer focus on the 5% highest tail shape values but all events. Figure 2.7a shows the subset of the tail shape parameter from Figure 2.6. In this time series, it is shown that

the LTCM crisis is the most significant event. This event occurs ex-post a spike in the tail shape parameter. From Figure 2.7b, it is shown that 9/11 and the Worldcom bankruptcy occurred ex-ante to jumps in the tail shape parameter. This could be because these events were unanticipated. The Iraq war occurs ex-post a significant spike, likely because the war was anticipated by pre-war events like 9/11 and the "Iraq Resolution," which authorised the U.S. President to "use any means necessary" against Iraq. The figures show that, for major events, low volatility is a strong predictor of financial crises.

### 2.5.3.1 News Events Selection

It is worth identifying some news events that are not included in Bloom (2009) selection criteria. Figure 2.7a shows a jump before the stock market crash on November 11th - 13th, 1997. This jump could be explained by several events, such as the "U.N. Imposes New Sanctions on Iraq", "Congress Rejects Funds for U.N. and I.M.F.", and "Eastman Kodak Laying Off 10,000 Employees."

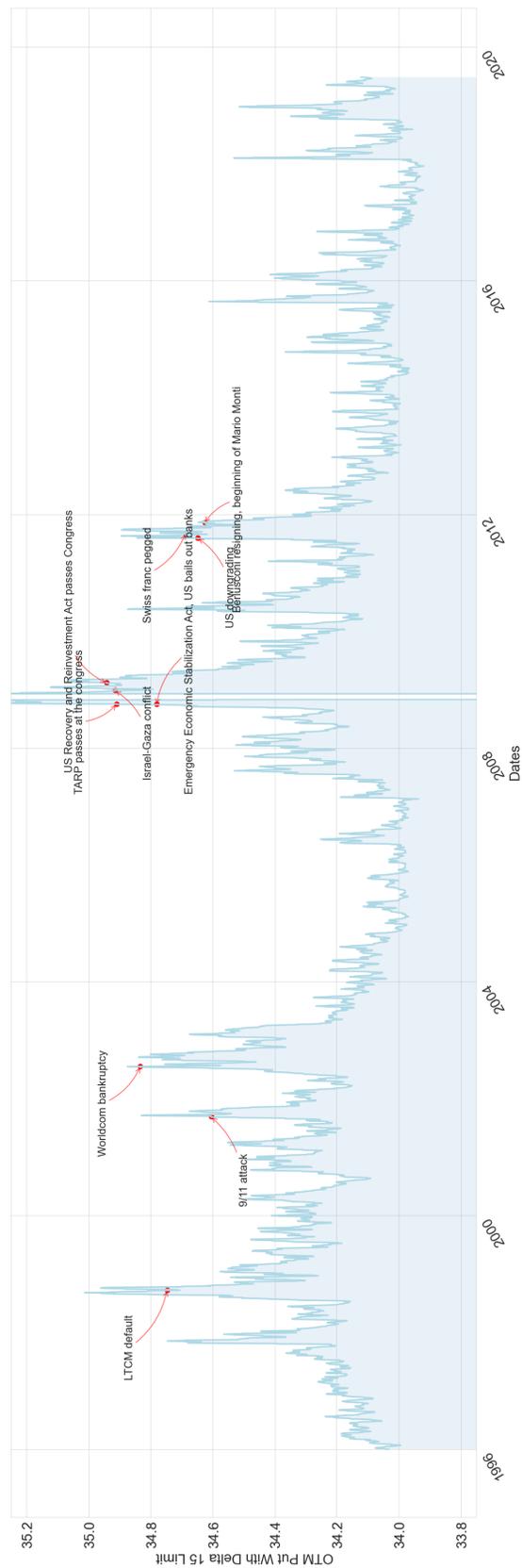
Another instance is the jump after the "LTCM default" on October 9th, 1998, when Germany raised interest rates. Thereafter, the Central Bank alarmed world economic centres with its first increase in five years, which led to European stock markets suffering.

Figure 2.7b shows a jump shortly after the 9/11 attack. In September, the stock markets plummeted 14%. Prices plunged in the second-largest decline in history as investors sought safe havens following the terrorist attacks on the U.S.

Figures 2.7, 2.8, and 2.9 show the TLM under the microscope, along with the corresponding

market events from 1996 up until 2015, and these figures explain the evolution of the tail shape parameter from years 1996-2007, 2007-2013, and 2013-2017, respectively. Here we can see the list of events specified in Table 2.20, and whether the uncertainty shock associated with the events were anticipated or not. This observation can be seen with a shock in the tail shape occurring before, during or after the event.

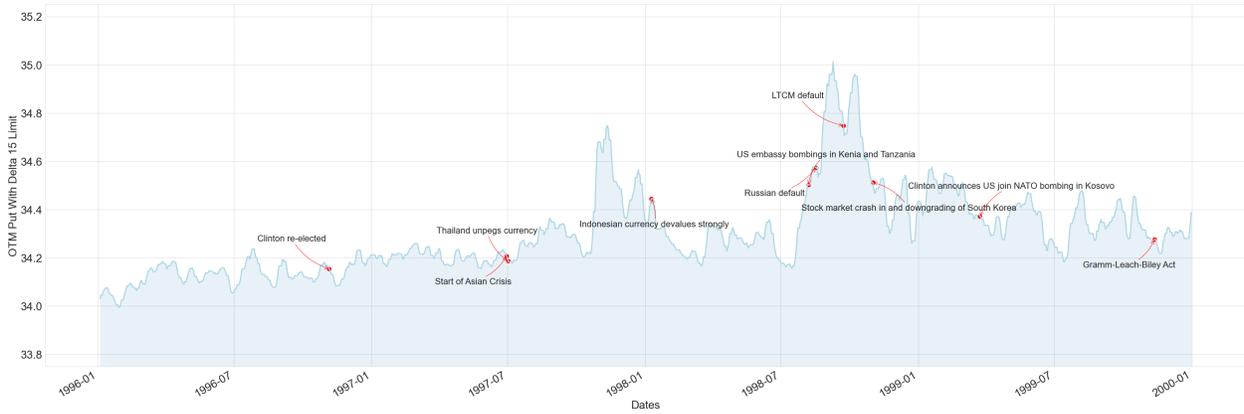
Figure 2.6: Tail Shape Parameter from 1996-2019



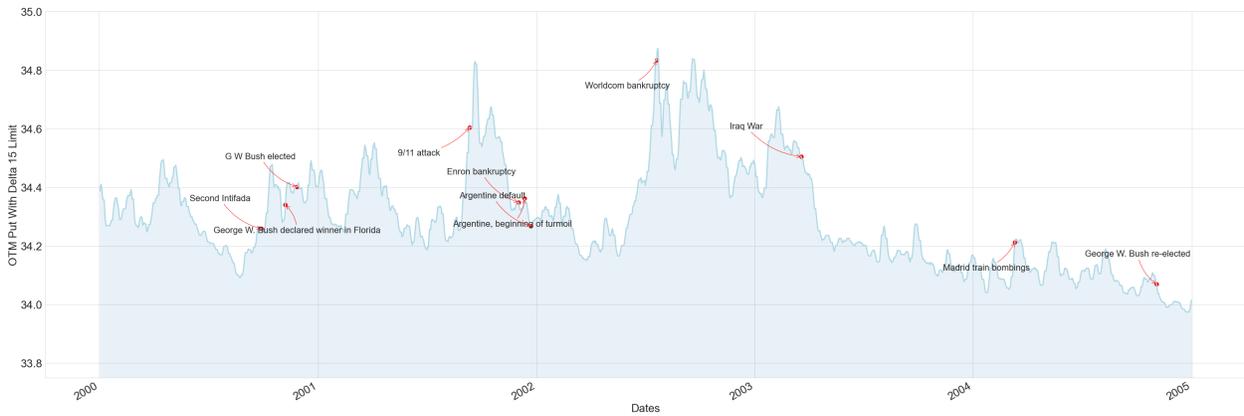
This chart explains the evolution of the tail shape parameter from years 1996-2019. The red arrows point to the financial and political events.

Figure 2.7: Tail Shape Parameter for Various Dates Spanning from 1996-2007

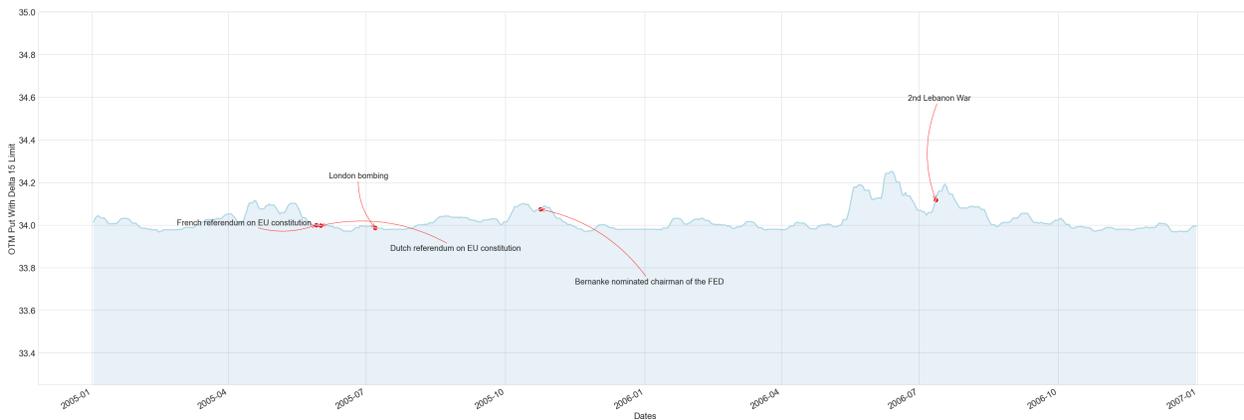
(a) Tail shape Parameter from 1996-2000



(b) Tail shape Parameter from 2000-2005



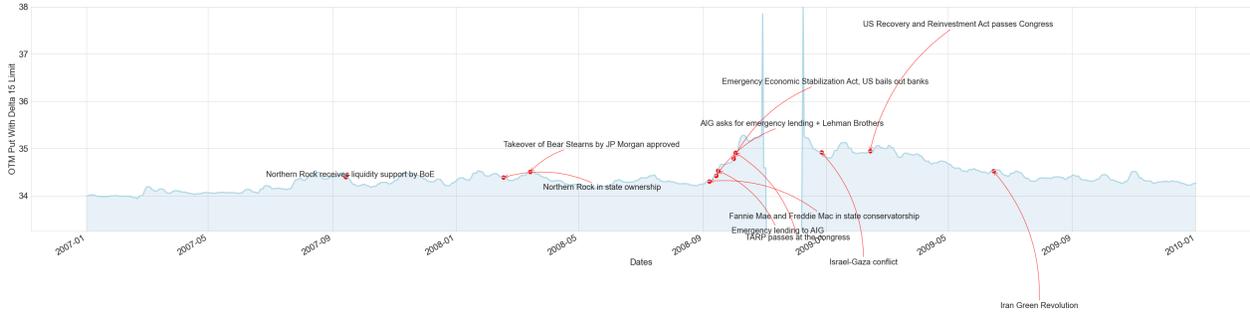
(c) Tail shape Parameter from 2005-2007



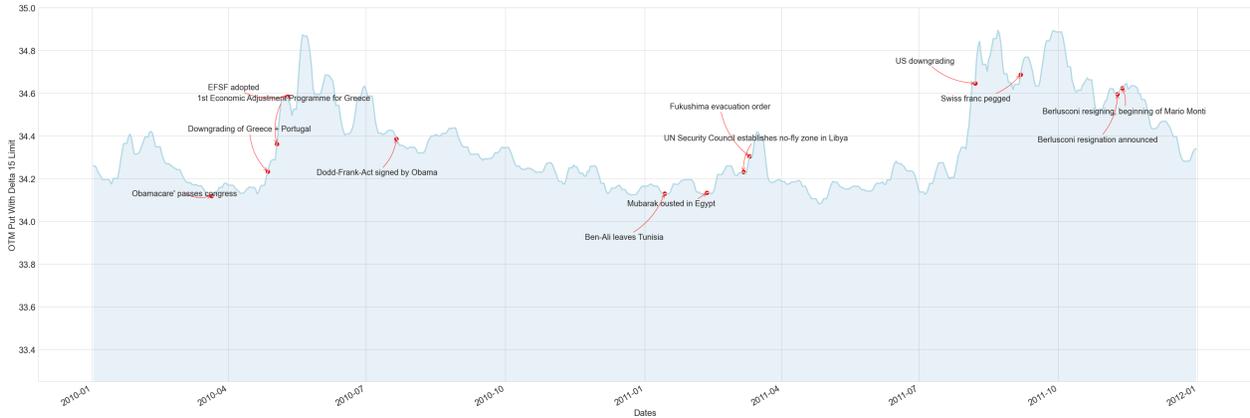
These charts explain the evolution of the tail shape parameter from years 1996-2007. The red arrows point to the financial and political events.

Figure 2.8: Tail Shape Parameter for Various Dates Spanning from 2007-2013

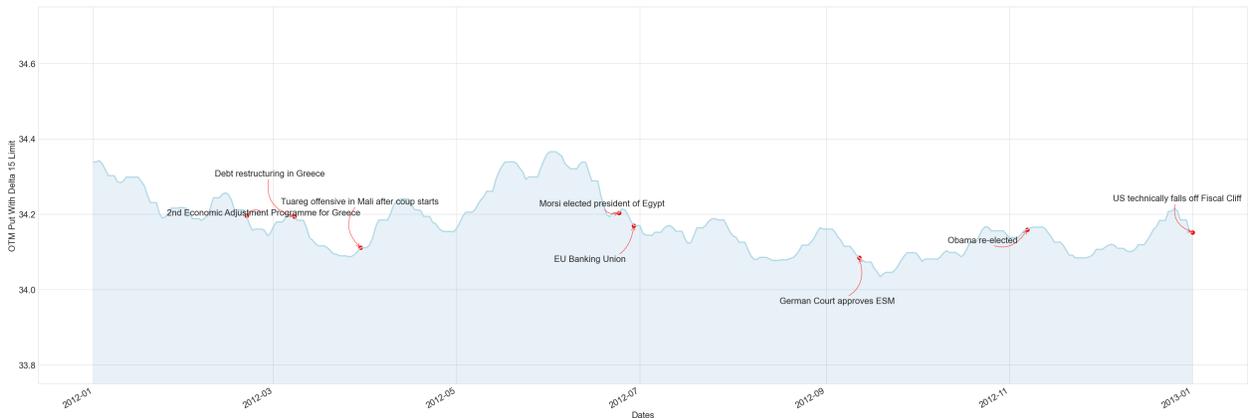
(a) Tail shape Parameter from 2007-2010



(b) Tail shape Parameter from 2010-2012



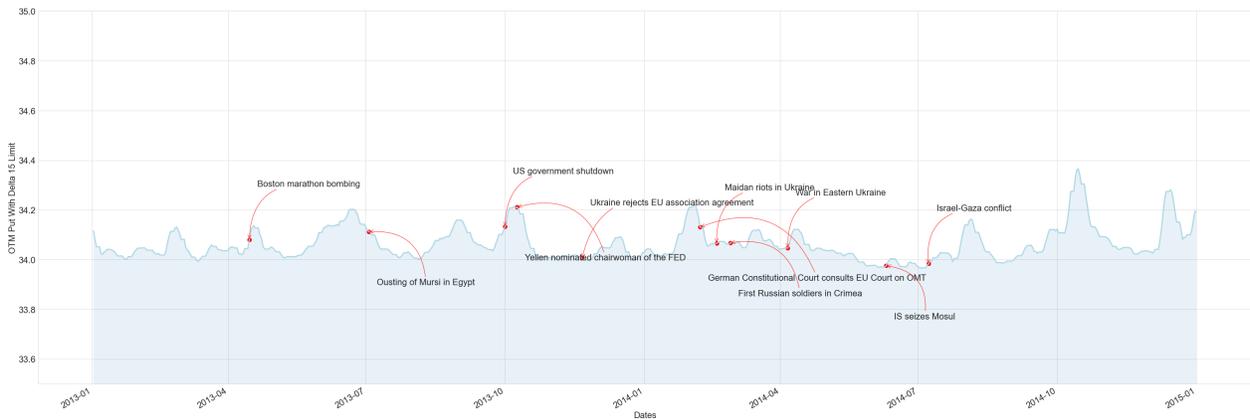
(c) Tail shape Parameter from 2012-2013



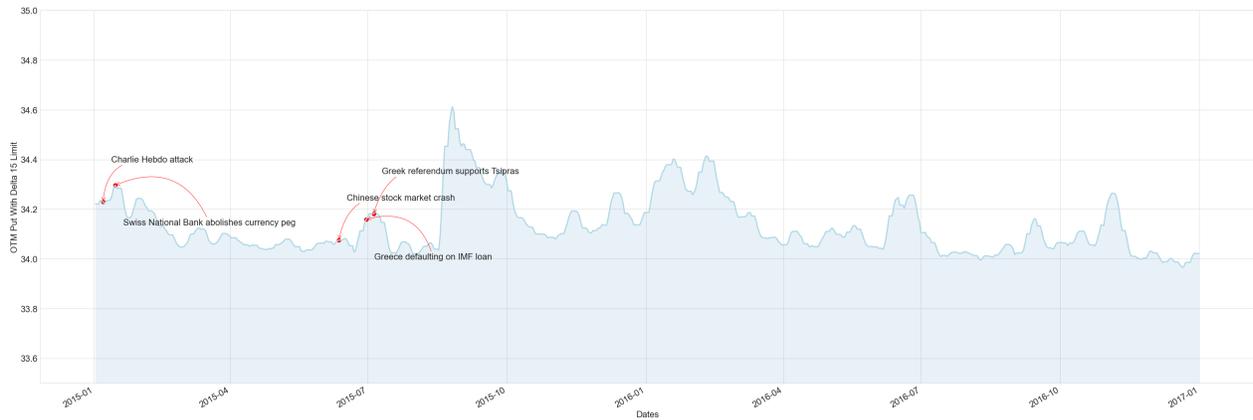
These charts explain the evolution of the tail shape parameter from years 2007-2013. The red arrows point to the financial and political events.

Figure 2.9: Tail Shape Parameter for Various Dates from 2013-2017

(a) Tail shape Parameter from 2013-2015

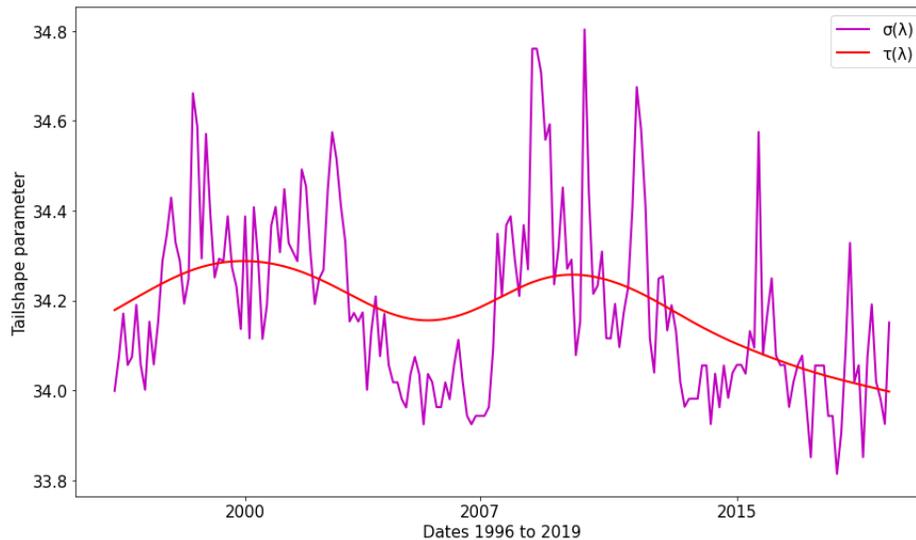


(b) Tail shape Parameter from 2015-2017



These charts explain the evolution of the tail shape parameter from years 2013-2017. The red arrows point to the financial and political events.

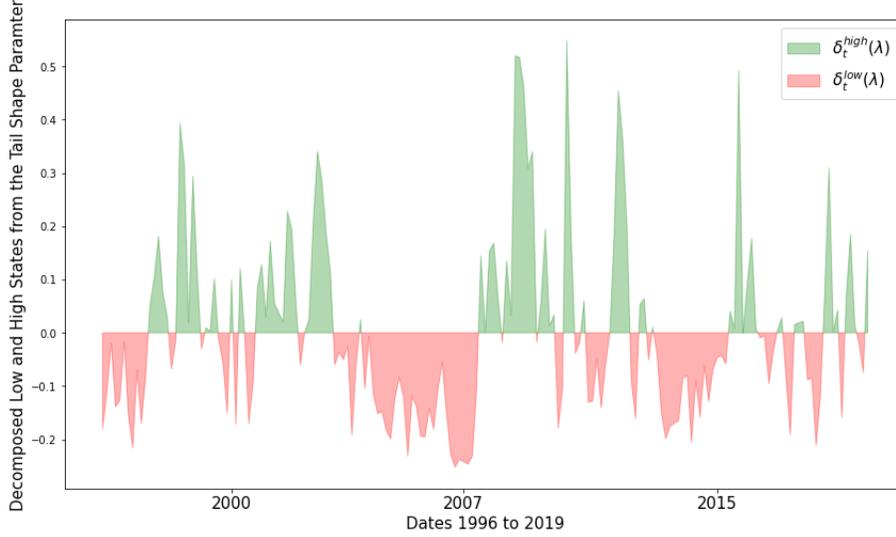
Figure 2.10: Monthly Time-Series of Tail Shape Parameter from 1996-2021



This chart illustrates the monthly time-series of the trend and cycle of the tail shape parameter from 1996 to 2021.

Figure 2.10 illustrates the tail shape parameter plotted with the trend, which is derived from the Hodrick and Prescott filter from Hodrick & Prescott (1997). The red line  $\tau_t$  represents the trend component, and  $\xi\sigma_t$  is the cyclical component (short-term deviations). Figure 2.11 depicts the decomposed cyclical behaviour of the tail shape parameter. From Equation 2.44, we can decompose  $\xi\sigma_t - \tau_t(\lambda)$  into  $\delta^{\text{high}}$  and  $\delta^{\text{low}}$  time series. In Table 2.5, we compute descriptive statistics for  $\xi\sigma$ ,  $|\xi\sigma - \tau|$ ,  $\delta^{\text{high}}$ , and  $\delta^{\text{low}}$ . This table shows that for  $\xi\sigma$ , the midpoint of the data is 34.183. The average value for  $\delta^{\text{high}}$  is 0.043, and for  $\delta^{\text{low}}$  it is -0.035. The descriptive statistics imply that the market is in a low-volatility state 55% of the time within this time horizon because the total count of  $\delta^{\text{low}}$  divided by the sum count of  $\delta^{\text{low}}$  and  $\delta^{\text{high}}$  is 55%.

Figure 2.11:  $\xi\sigma_t - \tau_t$  is Decomposed in to  $\delta_t^{\text{high}}(\lambda)$  and  $\delta_t^{\text{low}}(\lambda)$  Volatility for the Dates 1996-2021



This chart illustrates the monthly time-series of the decomposed high and low volatility states of the tail shape parameter from 1996-2021.

Table 2.5: Descriptive Analysis

Variable	$\xi\sigma$	$ \xi\sigma - \tau $	$\delta^{\text{high}}$	$\delta^{\text{low}}$
Count	8577	8577	3853	4724
Mean	34.225	0.039	0.043	-0.035
Std	0.211	0.046	0.052	0.041
Min	33.943	0.000	0.000	-0.475
25%	34.063	0.012	0.012	-0.046
50%	34.183	0.026	0.027	-0.025
75%	34.329	0.049	0.055	-0.012
Max	34.954	0.501	0.501	-0.00

This table shows time-series statistics for each variable indicated by the column headers for the period from 1996 to 2020. We present the mean, median, standard deviation, and the 25th, 50th, and 75th percentiles.  $\sigma$  is the volatility (tail shape parameter with a delta of 15%),  $\delta^{\text{high}}$  and  $\delta^{\text{low}}$  represent high and low volatility, respectively.

#### 2.5.4 Probit Model Results

Table 2.6 presents the results for the Probit regression model detailed in Equation 2.45, and Table 2.7 presents the results for the lagged Probit regression model detailed in Equation 2.46, where the last 5 years information ( $L = 10$ ) are considered. Here, each  $L$  represents a half-year. The first relationship considered is how volatility (as a measure based on the tail shape parameter) relates to the probability of a future crisis. From Table 2.6, the first two columns 1 and 2 explains the results as  $\beta\xi\sigma > 0$ . When considered on its own as an independent variable, volatility is not statistically significant; however, when control variables are considered volatility is highly statistically significant. This result implies that the likelihood of future crisis is completely captured by the control variable public debt-to-GDP ratio. By contrast, the cycle which is mathematically defined as the absolute value of the deviation of volatility from its trend  $|\xi\sigma - \tau|$ , is not significant and this can be seen in column 3. However, with the selected control variable  $|\xi\sigma - \tau|$ , and the control variable both gain statistical significance. The coefficient of the cycle is  $\beta|\xi\sigma - \tau| < 0$ . The cycle decreases as the probability of a crisis increases.

Following Danielsson & Valenzuela & Zer (2018), we include high and low volatilities defined in Equation 2.44 as separate regressors, reported in columns 5 and 6. When control variables are excluded, the coefficients  $\delta^{\text{high}}$  and  $\delta^{\text{low}}$  are not significant, with expected signs  $\beta\delta^{\text{high}} > 0$  and  $\beta\delta^{\text{low}} < 0$ . However, with the inclusion of the control variable, only  $\delta^{\text{low}}$  gains significance, with expected signs  $\beta\delta^{\text{high}} > 0$  and  $\beta\delta^{\text{low}} < 0$ . Danielsson & Valenzuela & Zer (2018)

assert that this result economically implies that low volatility affects agents' decision-making, induces excessive risk-taking, which leads to future credit problems and eventually difficulties for banks. Marginal effects reported in table show the change in the probability of political or financial crises following an instantaneous change in high or low volatility. Marginal effects in the Probit model shown in Equation 2.45, measures how a small change in an explanatory variable in  $\Gamma$  affects the probability that  $\mathbb{P}(J_t = 1)$ . The estimated Marginal Effects (ME) from the table show that the impact of low volatility on the probability of a crisis is economically meaningful: a unit decrease in volatility, when it is below its trend, translates into a 23.052% increase in the probability of market uncertainty.

Table 2.6: Volatility and Crisis

Dep Variable: $J_t$	1	2	3	4	5	6
$\xi\sigma_{t-1 \text{ to } t-10}$	0.006 (0.014)	0.656*** (0.198)				
$ \xi\sigma - \tau _{t-1 \text{ to } t-10}$			-17.857 (11.002)	-48.175** (16.552)		
$\delta_{t-1 \text{ to } t-10}^{\text{high}}$					-10.248 (13.866)	4.756 (30.603)
$\delta_{t-1 \text{ to } t-10}^{\text{low}}$					-16.309 (11.188)	-103.467** (37.726)
$J_{t-1 \text{ to } t-10}$	-3.138 (1.846)	53.922*** (16.399)	-2.254*** (0.585)	10.727* (4.359)	-3.895* (1.897)	24.310** (9.065)
$\Delta PD/GDP_{t-1 \text{ to } t-10}$		-0.478*** (0.143)		-0.042*** (0.014)		-0.194* (0.080)
Number of Observations	8501	8501	8501	8501	8501	8501
Pseudo $R^2$	0.029	0.320	0.057	0.158	0.066	0.332
Marginal Effects (%)						
$\xi\sigma_{t-1 \text{ to } t-10}$	0.002	0.148				
$ \xi\sigma - \tau _{t-1 \text{ to } t-10}$			-5.572	-13.356		
$\delta_{t-1 \text{ to } t-10}^{\text{high}}$					-3.167	1.060
$\delta_{t-1 \text{ to } t-10}^{\text{low}}$					-5.039	-23.052
$\Delta PD/GDP_{t-1 \text{ to } t-10}$	-1.02	-0.107	-0.703	-0.012		-0.043

This table presents the results for the regression equation introduced in Equation 2.45.  $\delta^{\text{high}}$  and  $\delta^{\text{low}}$  are high and low volatility introduced in Equation 2.44,  $\xi\sigma$  represents the volatility, and  $|\xi\sigma - \tau|$  is the absolute value of the deviation of volatility from its trend.  $\Delta PD/GDP$  denotes the public debt-to-GDP ratio. The past five year averages of the explanatory variables are used in the regressions and the data spans from 1996-2019. Here, stars are only intended to flag levels of significance. If a p-value is  $\leq 0.05$  it is flagged with one star (\*), if a p-value is  $\leq 0.01$  it is flagged with two stars (\*\*), subsequently, if a p-value is  $\leq 0.001$  it is flagged with three stars (\*\*\*).

Table 2.7: Volatility and Crisis Using Different Lag Lengths

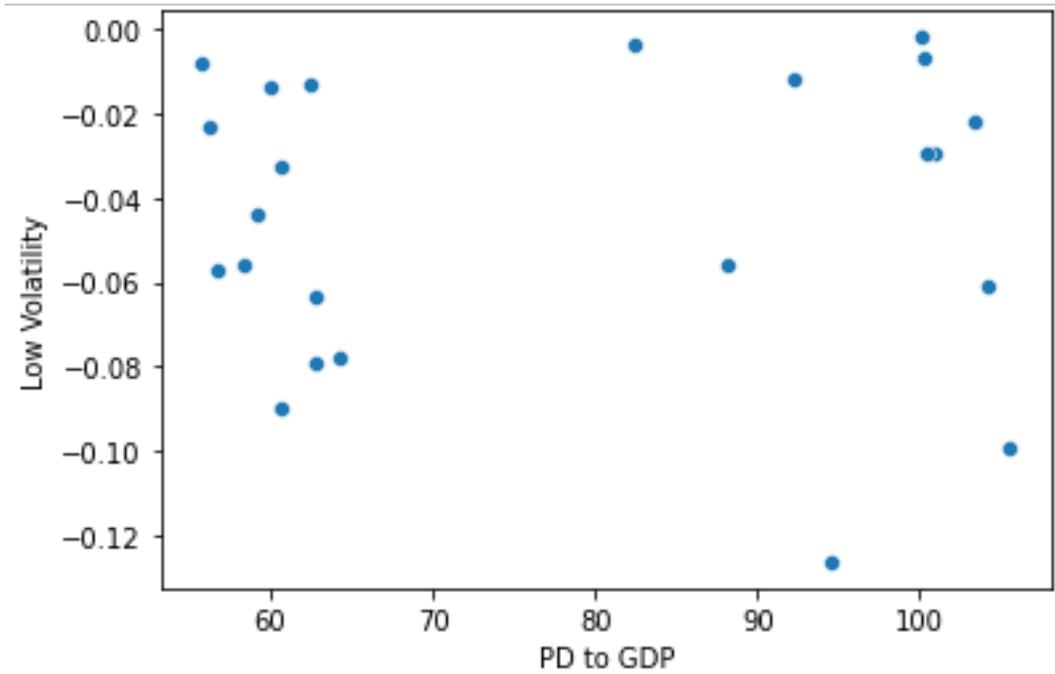
Dep Variable: $J_t$ ; $t=$ Half-yearly	$L = 1$	$L = 2$	$L = 3$	.	$L = 8$	$L = 9$	$L = 10$
$\delta_{t-1 \text{ to } t-L}^{\text{high}}$	7.257*	8.819	11.990*	.	21.276	2.913	4.756
	(2.962)	(4.524)	(5.636)	.	(20.378)	(26.153)	(30.603)
$\delta_{t-1 \text{ to } t-L}^{\text{low}}$	-20.953**	-17.126	-19.704	.	-81.223*	-68.144**	-103.467**
	(7.047)	(9.658)	(10.342)	.	(34.217)	(23.612)	(37.726)
$J_{t-1 \text{ to } t-L}$		9.390**	9.922**	.	29.266*	20.867**	24.310**
		(3.081)	(3.516)	.	(12.278)	(7.461)	(9.065)
$\Delta PD/GDP_{t-1 \text{ to } t-L}$	-0.025***	-0.064**	-0.071**	.	-0.205**	-0.144**	-0.194*
	(0.007)	(0.021)	(0.026)	.	(0.080)	(0.049)	(0.080)
Number of Observations	8501	8501	8501	.	8501	8501	8501
Pseudo $R^2$	0.122	0.314	0.278	.	0.282	0.275	0.332
Marginal Effects (%)							
$\delta_{t-1 \text{ to } t-L}^{\text{high}}$	2.085	2.016	2.879	.	5.096	0.702	1.060
$\delta_{t-1 \text{ to } t-L}^{\text{low}}$	-6.019	-3.915	-4.731	.	-19.455	-16.429	-23.052
$\Delta PD/GDP_{t-1 \text{ to } t-L}$	-0.007	-0.015	-0.017	.	-0.049	-0.035	-0.043

This table presents the results for the regression equation introduced in Equation 2.45 for  $L = 1, \dots, 10$ . These lags are half-yearly lags, therefore, when  $L = 10$ ,  $L$  represents five years.  $\delta^{\text{high}}$  and  $\delta^{\text{low}}$  are high and low volatility introduced in Equation 2.44,  $\xi\sigma$  represents the volatility, and  $|\xi\sigma - \tau|$  is the absolute value of the deviation of volatility from its trend.  $\Delta PD/GDP$  denotes the public debt-to-GDP ratio. Here, stars are only intended to flag levels of significance. If a p-value is  $\leq 0.05$  it is flagged with one star (\*), if a p-value is  $\leq 0.01$  it is flagged with two stars (\*\*), subsequently, if a p-value is  $\leq 0.001$  it is flagged with three stars (\*\*\*)

In Table 2.7, we examine the predictive power of low volatility using different lag lengths. The results show a negative relationship between low volatility and future financial crises when information up to 5 years is taken into account. Marginal effects reported at the bottom of the table show the change in the probability of market uncertainty following an instantaneous change in low volatility. We find that a decrease in volatility below its trend translates into a 6.019%, 4.731%, 19.455%, and 23.052% increase in the probability of market uncertainty when the last 1, 3, 8, and 10 half-years' information is used, respectively. The marginal effects for low volatility increase monotonically and reach a maximum when the last 5 years are considered, indicating that the economic impact is highest if the SPX remains in a low-volatility environment for 5 years. This result corresponds with Danielsson & Valenzuela & Zer (2018), who found that low volatility reaches a maximum at 5 years when analysing yearly lags up to 10 years. As established in Danielsson & Valenzuela & Zer (2018), we also obtain results that imply an increase in the debt-to-GDP ratio is negatively associated with the probability of future market uncertainty as we have negative and significant test statistics in both Probit model results in Table 2.6 and Table 2.7. We conduct additional tests to assess multicollinearity between low volatility and debt-to-GDP. The correlation coefficient between these variables is low, at 0.020937, indicating minimal linear relation. We further calculate the Variance Inflation Factor (VIF) to assess multicollinearity for both variables and obtain a value of 1.000439 which is very close to 1, suggesting no multicollinearity. Additionally, as shown in Figure 2.12, the scatter plot visually confirms the lack of correlation between the two variables.

Danielsson & Valenzuela & Zer (2018) included  $\text{INFLATION}_{t-1 \text{ to } t-L}$ , however, we do not

Figure 2.12: Low Volatility vs. Debt-To-GDP



This illustration shows the relationship between low volatility and debt-to-GDP.

include  $\text{INFLATION}_{t-1 \text{ to } t-L}$  and  $10 \text{ YEAR TREASURY YIELD}_{t-1 \text{ to } t-L}$  control variables due to multicollinearity observed in a correlation matrix.

## 2.6 Conclusion

My main contribution shows that low volatility is a strong predictor of financial crises. The evolution of the tail shape parameter and the PDF conforms to Keynes (1937) and Minsky (1977) theories, which suggest that economic agents observing low financial risk tend to increase risk-taking, leading to a crisis. We focused on various modelling approaches to estimate the risk-neutral option-implied PDF from market-traded options. The choice of methodology for estimating the RND to examine the forecastability of crashes is appropriate.

Using the DM test, the mixture of the log-normal method is the most accurate compared to the other estimated methods proposed in this study when forecasting S&P 500 European option call prices. The statistical analysis of the comprehensive dataset explains that both the mean and the variance are higher for ex-post densities compared to ex-ante densities. This observation shows how low volatility can paradoxically increase the probability of a systemic event. However, for ex-ante political density means, the variance is higher than the ex-post variance of density means. This dynamic could indicate that, in the case of political events, uncertainty is usually anticipated. Therefore, uncertainty is priced ex-ante or during the event. For the complete dataset, the mean of the ex-ante MAE, MAPE, and RMSE for both puts and calls is higher than the ex-post results, implying that forecasting errors decrease after the uncertain event, the results are shown in Table 2.8 to Table 2.19. Decomposing  $\delta^{\text{high}}$  and  $\delta^{\text{low}}$  volatilities allows for a focus on low volatility separately. My findings reveal that we gain higher marginal effects over an extended period of low volatility; this result implies that the economic impact is the highest if the economy remains in a low-volatility environment. These findings are useful to risk managers, hedgers, and monetary policymakers because low volatility could be used as an early warning indicator.

Table 2.8: Ex-ante and Ex-post MAE of Calls of Financial and Political Events

Type of Event	Event	MAE call t-1	MAE call t	MAE call t+1
Financial	Greece defaulting on IMF loan	8.886	9.369	8.730
Financial	Chinese stock market crash	9.996	9.903	8.334
Financial	Swiss National Bank abolishes currency peg	10.185	8.309	8.468
Political	Charlie Hebdo attack	9.329	9.405	9.696
Political	Israel-Gaza conflict	25.165	23.104	21.594
Political	IS seizes Mosul	17.271	16.813	16.406
Political	First Russian soldiers in Crimea	8.278	8.113	8.137
Political	9/11 Attack	6.659	23.837	18.321
Political	Maidan riots in Ukraine	28.683	10.868	9.176
Political	German Constitutional Court consults EU Court on OMT	27.550	26.154	24.918
Political	Ukraine rejects EU association agreement	11.998	11.837	12.876
Financial	Yellen nominated chairwoman of the FED	18.692	15.433	16.919
Political	US government shutdown	28.964	14.843	14.655
Political	Ousting of Mursi in Egypt	13.524	13.364	13.812
Political	Boston marathon bombing	14.364	13.035	1.175
Financial	US technically falls off Fiscal Cliff	8.669	8.611	8.826
Political	Obama re-elected	20.318	18.898	18.928
Financial	German Court approves ESM	19.276	18.818	17.235
Financial	EU Banking Union	10.679	11.077	11.757
Political	Tuareg offensive in Mali after coup starts	6.525	6.350	6.163
Financial	Debt restructuring in Greece	6.976	5.497	6.180
Financial	2nd Economic Adjustment Programme for Greece	12.878	10.474	10.196
Political	Berlusconi resignation announced	34.472	32.216	22.920
Financial	Swiss frank pegged	11.486	10.747	9.670
Political	Fukushima evacuation order	14.760	14.680	12.298
Political	UN Security Council establishes no-fly zone in Libya	23.554	19.927	19.883
Political	Mubarak ousted in Egypt	17.549	20.309	15.449
Political	Ben-Ali leaves Tunisia	13.973	13.941	12.421
Financial	Dodd-Frank-Act signed by Obama	12.587	11.034	13.312
Financial	EFSF adopted	13.071	15.361	13.926

(a) Panel A

Table 2.9: Ex-ante and Ex-post MAE of Calls of Financial and Political Events

Type of Event	Event	MAE call t-1	MAE call t	MAE call t+1
Financial	1st Economic Adjustment Programme for Greece	6.942	8.303	5.962
Financial	Downgrading of Greece + Portugal	9.444	8.878	9.067
Political	Obamacare' passes congress	12.419	10.605	8.843
Political	Iran Green Revolution	57.556	48.976	55.249
Financial	US Recovery and Reinvestment Act passes Congress	17.852	20.529	18.147
Political	Obama elected	21.592	16.073	14.269
Financial	TARP passes at the congress	27.334	25.733	22.376
Financial	Emergency Economic Stabilisation Act, US bails out banks	30.328	29.928	27.334
Financial	Emergency lending to AIG	20.028	8.131	11.625
Financial	AIG asks for emergency lending + Lehman Brothers	36.337	24.780	20.028
Financial	Takeover of Bear Stearns by JP Morgan approved	33.740	28.291	22.009
Financial	Northern Rock receives liquidity support by BoE	105.201	101.843	103.826
Political	2nd Lebanon War	84.493	80.892	80.318
Financial	Bernanke nominated chairman of the FED	59.088	58.580	57.681
Political	London bombing	68.680	67.085	69.174
Political	Dutch referendum on EU constitution	57.079	57.175	57.048
Political	George W. Bush re-elected	71.522	70.311	71.484
Political	Madrid train bombings	34.496	33.827	32.954
Political	Iraq War	25.776	30.273	30.658
Political	Argentine, beginning of turmoil	40.639	37.918	40.408
Political	G W Bush elected	73.204	72.315	69.953
Political	George W. Bush declared winner in Florida	114.378	108.083	106.385
Political	Second Intifada	162.072	167.741	157.635
Financial	Gramm-Leach-Bliley Act	137.415	139.228	136.898
Political	Clinton announces US join NATO bombing in Kosovo	96.494	97.151	99.994
Financial	LTCM default	96.211	92.951	90.598
Financial	Russian default	127.506	129.350	133.510
Political	US embassy bombings in Kenia and Tanzania	134.955	135.151	136.649
Financial	Indonesian currency devalues strongly	69.531	65.148	62.994
Financial	Thailand unpegs currency	47.197	49.131	50.516
Financial	Start of Asian Crisis	47.753	46.980	47.197
Political	Clinton re-elected	49.749	50.672	50.439
	MEAN	39.945	38.205	37.300
	STDEV	37.799	38.333	38.412

(a) Panel B

Table 2.10: Ex-ante and Ex-post MAPE of Calls of Financial and Political Events

Type of Event	Event	MAPE put t-1	MAPE put t	MAPE put t+1
Financial	Greece defaulting on IMF loan	0.092	0.188	0.104
Financial	Chinese stock market crash	0.157	0.296	0.153
Financial	Swiss National Bank abolishes currency peg	0.608	0.159	0.125
Political	Charlie Hebdo attack	0.647	0.077	0.076
Political	Israel-Gaza conflict	0.181	0.190	3.298
Political	IS seizes Mosul	0.173	0.169	0.170
Political	First Russian soldiers in Crimea	0.110	0.100	0.092
Political	9/11 Attack	0.288	1.454	1.354
Political	Maidan riots in Ukraine	0.176	0.158	0.107
Political	German Constitutional Court consults EU Court on OMT	1.887	0.323	0.317
Political	Ukraine rejects EU association agreement	0.177	0.564	0.117
Financial	Yellen nominated chairwoman of the FED	0.406	0.269	0.242
Political	US government shutdown	0.297	0.266	0.256
Political	Ousting of Mursi in Egypt	0.211	0.163	0.176
Political	Boston marathon bombing	0.231	0.235	0.136
Financial	US technically falls off Fiscal Cliff	1.024	0.116	3.841
Political	Obama re-elected	0.354	0.333	5.355
Financial	German Court approves ESM	0.354	2.893	6.078
Financial	EU Banking Union	0.193	0.200	0.350
Political	Tuareg offensive in Mali after coup starts	0.131	0.128	0.134
Financial	Debt restructuring in Greece	1.077	0.157	0.205
Financial	2nd Economic Adjustment Programme for Greece	1.148	0.162	0.157
Political	Berlusconi resignation announced	1.564	0.779	0.229
Financial	Swiss frank pegged	0.202	0.324	0.149
Political	Fukushima evacuation order	0.252	0.240	0.200
Political	UN Security Council establishes no-fly zone in Libya	0.279	0.283	0.275
Political	Mubarak ousted in Egypt	3.092	0.221	1.001
Political	Ben-Ali leaves Tunisia	0.332	0.332	0.548
Financial	Dodd-Frank-Act signed by Obama	0.298	0.261	0.154
Financial	EFSF adopted	0.259	0.790	0.260

(a) Panel A

Table 2.11: Ex-ante and Ex-post MAPE of Calls of Financial and Political Events

Type of Event	Event	MAPE put t-1	MAPE put t	MAPE put t+1
Financial	1st Economic Adjustment Programme for Greece	0.119	4.126	0.179
Financial	Downgrading of Greece + Portugal	0.164	0.141	0.614
Political	Obamacare' passes congress	0.303	2.158	0.212
Political	Iran Green Revolution	0.759	0.745	0.772
Financial	US Recovery and Reinvestment Act passes Congress	0.486	0.678	0.995
Political	Obama elected	6.044	0.352	2.492
Financial	TARP passes at the congress	0.496	0.719	0.581
Financial	Emergency Economic Stabilisation Act, US bails out banks	0.570	0.592	0.496
Financial	Emergency lending to AIG	42.524	0.490	0.468
Financial	AIG asks for emergency lending + Lehman Brothers	62.975	28.405	42.524
Financial	Takeover of Bear Stearns by JP Morgan approved	14.177	6.190	19.410
Financial	Northern Rock receives liquidity support by BoE	26.614	29.182	26.209
Political	2nd Lebanon War	62.166	46.824	43.814
Financial	Bernanke nominated chairman of the FED	109.769	102.063	102.301
Political	London bombing	81.948	77.881	81.622
Political	Dutch referendum on EU constitution	25.950	14.385	13.429
Political	George W. Bush re-elected	35.007	41.728	33.698
Political	Madrid train bombings	17.529	9.954	18.628
Political	Iraq War	34.631	49.718	66.834
Political	Argentine, beginning of turmoil	42.806	43.345	45.683
Political	G W Bush elected	29.445	23.930	37.542
Political	George W. Bush declared winner in Florida	94.420	93.806	94.488
Political	Second Intifada	322.250	155.987	146.325
Financial	Gramm-Leach-Bliley Act	21.432	24.947	20.480
Political	Clinton announces US join NATO bombing in Kosovo	7.114	7.651	6.288
Financial	LTCM default	49.353	32.110	54.398
Financial	Russian default	43.719	55.240	47.430
Political	US embassy bombings in Kenia and Tanzania	51.119	53.599	55.475
Financial	Indonesian currency devalues strongly	5.643	8.785	6.858
Financial	Thailand unpegs currency	1.483	1.250	1.163
Financial	Start of Asian Crisis	1.572	1.867	1.483
Political	Clinton re-elected	1.405	1.120	1.062
	MEAN	19.835	15.252	16.365
	STDEV	46.474	29.764	29.876

(a) Panel B

Table 2.12: Ex-ante and Ex-post RMSE of Calls of Financial and Political Events

Type of Event	Event	RMSE call t-1	RMSE call t	RMSE call t+1
Financial	Greece defaulting on IMF loan	10.537	10.571	9.941
Financial	Chinese stock market crash	11.555	11.365	9.898
Financial	Swiss National Bank abolishes currency peg	11.972	10.250	10.295
Political	Charlie Hebdo attack	10.634	11.310	11.451
Political	Israel-Gaza conflict	28.452	25.549	23.127
Political	IS seizes Mosul	20.015	19.552	19.091
Political	First Russian soldiers in Crimea	9.538	9.200	9.263
Political	9/11 Attack	11.235	29.431	25.254
Political	Maidan riots in Ukraine	33.145	13.056	10.766
Political	German Constitutional Court consults EU Court on OMT	32.059	30.166	28.582
Political	Ukraine rejects EU association agreement	13.942	14.166	15.376
Financial	Yellen nominated chairwoman of the FED	22.733	18.797	20.998
Political	US government shutdown	33.966	16.563	16.352
Political	Ousting of Mursi in Egypt	15.184	15.443	15.936
Political	Boston marathon bombing	16.497	15.219	1.427
Financial	US technically falls off Fiscal Cliff	10.568	11.002	11.151
Political	Obama re-elected	25.243	24.156	23.899
Financial	German Court approves ESM	24.780	24.354	22.161
Financial	EU Banking Union	13.817	14.100	15.626
Political	Tuareg offensive in Mali after coup starts	8.695	8.657	8.421
Financial	Debt restructuring in Greece	8.734	6.677	8.192
Financial	2nd Economic Adjustment Programme for Greece	16.102	12.412	12.270
Political	Berlusconi resignation announced	39.069	37.475	25.018
Financial	Swiss frank pegged	14.254	13.763	12.036
Political	Fukushima evacuation order	18.523	18.442	16.013
Political	UN Security Council establishes no-fly zone in Libya	31.071	25.868	25.814
Political	Mubarak ousted in Egypt	21.332	25.522	18.581
Political	Ben-Ali leaves Tunisia	16.887	16.875	16.328
Financial	Dodd-Frank-Act signed by Obama	17.764	14.569	18.649
Financial	EFSF adopted	15.464	21.196	17.724

(a) Panel A

Table 2.13: Ex-ante and Ex-post RMSE of Calls of Financial and Political Events

Type of Event	Event	RMSE call t-1	RMSE call t	RMSE call t+1
Financial	1st Economic Adjustment Programme for Greece	9.235	10.003	7.691
Financial	Downgrading of Greece + Portugal	11.563	10.816	12.066
Political	Obamacare' passes congress	17.875	14.308	11.306
Political	Iran Green Revolution	86.500	74.697	86.787
Financial	US Recovery and Reinvestment Act passes Congress	24.021	29.123	27.941
Political	Obama elected	31.609	26.365	20.164
Financial	TARP passes at the congress	36.279	32.638	31.302
Financial	Emergency Economic Stabilisation Act, US bails out banks	39.078	37.959	36.279
Financial	Emergency lending to AIG	26.616	10.125	15.182
Financial	AIG asks for emergency lending + Lehman Brothers	50.206	32.623	26.616
Financial	Takeover of Bear Stearns by JP Morgan approved	37.455	31.180	26.305
Financial	Northern Rock receives liquidity support by BoE	134.259	130.707	132.941
Political	2nd Lebanon War	104.995	101.554	99.721
Financial	Bernanke nominated chairman of the FED	76.550	75.564	74.477
Political	London bombing	83.270	81.459	83.395
Political	Dutch referendum on EU constitution	65.901	65.963	65.818
Political	George W. Bush re-elected	86.985	85.873	87.147
Political	Madrid train bombings	47.073	47.024	45.437
Political	Iraq War	33.445	35.347	36.150
Political	Argentine, beginning of turmoil	49.585	46.929	49.089
Political	G W Bush elected	81.195	80.480	77.089
Political	George W. Bush declared winner in Florida	129.045	121.573	119.816
Political	Second Intifada	177.319	182.613	173.861
Financial	Gramm-Leach-Bliley Act	157.067	159.208	155.531
Political	Clinton announces US join NATO bombing in Kosovo	110.548	111.353	114.550
Financial	LTCM default	112.119	108.186	105.941
Financial	Russian default	144.012	146.938	151.788
Political	US embassy bombings in Kenia and Tanzania	152.742	153.389	154.887
Financial	Indonesian currency devalues strongly	79.781	74.887	72.848
Financial	Thailand unpegs currency	57.484	59.803	61.469
Financial	Start of Asian Crisis	58.087	57.335	57.484
Political	Clinton re-elected	58.330	60.077	59.839
	MEAN	47.914	45.777	44.841
	STDEV	43.199	43.738	44.066

(a) Panel B

Table 2.14: Ex-ante and Ex-post MAE of Puts of Financial and Political Events

Type of Event	Event	MAE put t-1	MAE put t	MAE put t+1
Financial	Greece defaulting on IMF loan	8.687	5.842	8.599
Financial	Chinese stock market crash	4.804	4.171	10.351
Financial	Swiss National Bank abolishes currency peg	5.071	9.307	9.555
Political	Charlie Hebdo attack	4.607	10.290	10.129
Political	Israel-Gaza conflict	7.425	6.816	14.987
Political	IS seizes Mosul	17.041	16.725	16.396
Political	First Russian soldiers in Crimea	9.873	10.073	9.779
Political	9/11 Attack	4.754	26.233	24.271
Political	Maidan riots in Ukraine	12.020	4.209	11.106
Political	German Constitutional Court consults EU Court on OMT	18.336	32.614	31.411
Political	Ukraine rejects EU association agreement	13.745	9.088	2.843
Financial	Yellen nominated chairwoman of the FED	7.978	18.460	12.912
Political	US government shutdown	13.131	19.468	19.165
Political	Ousting of Mursi in Egypt	14.027	12.568	6.056
Political	Boston marathon bombing	16.077	14.833	0.466
Financial	US technically falls off Fiscal Cliff	4.190	8.457	5.081
Political	Obama re-elected	24.746	24.110	18.505
Financial	German Court approves ESM	19.570	18.596	9.877
Financial	EU Banking Union	11.954	11.725	3.808
Political	Tuareg offensive in Mali after coup starts	7.022	6.448	6.321
Financial	Debt restructuring in Greece	2.998	6.355	1.812
Financial	2nd Economic Adjustment Programme for Greece	4.591	12.818	12.272
Political	Berlusconi resignation announced	30.675	30.969	27.977
Financial	Swiss frank pegged	16.619	10.136	9.666
Political	Fukushima evacuation order	14.473	14.196	11.991
Political	UN Security Council establishes no-fly zone in Libya	10.412	17.750	17.457
Political	Mubarak ousted in Egypt	10.657	7.628	19.428
Political	Ben-Ali leaves Tunisia	17.235	17.142	4.195
Financial	Dodd-Frank-Act signed by Obama	10.273	13.871	6.147
Financial	EFSF adopted	16.652	7.088	13.191

(a) Panel A

Table 2.15: Ex-ante and Ex-post MAE of Puts of Financial and Political Events

Type of Event	Event	MAE put t-1	MAE put t	MAE put t+1
Financial	1st Economic Adjustment Programme for Greece	6.719	3.159	6.696
Financial	Downgrading of Greece + Portugal	9.799	9.598	2.173
Political	Obamacare' passes congress	6.296	7.540	9.416
Political	Iran Green Revolution	85.708	79.435	66.575
Financial	US Recovery and Reinvestment Act passes Congress	21.892	7.772	10.020
Political	Obama elected	16.314	17.583	13.849
Financial	TARP passes at the congress	32.013	38.556	35.720
Financial	Emergency Economic Stabilisation Act, US bails out banks	30.410	30.750	32.013
Financial	Emergency lending to AIG	21.391	14.707	13.076
Financial	AIG asks for emergency lending + Lehman Brothers	37.059	30.258	21.391
Financial	Takeover of Bear Stearns by JP Morgan approved	28.911	22.766	22.137
Financial	Northern Rock receives liquidity support by BoE	45.689	51.051	42.435
Political	2nd Lebanon War	30.316	34.381	35.427
Financial	Bernanke nominated chairman of the FED	27.957	25.477	24.736
Political	London bombing	27.816	28.114	24.680
Political	Dutch referendum on EU constitution	19.420	17.904	17.132
Political	George W. Bush re-elected	31.741	29.906	28.628
Political	Madrid train bombings	21.719	15.593	16.928
Political	Iraq War	46.592	41.947	40.993
Political	Argentina, beginning of turmoil	35.281	37.555	37.015
Political	G W Bush elected	60.428	57.216	60.427
Political	George W. Bush declared winner in Florida	94.618	103.345	104.385
Political	Second Intifada	90.310	86.281	97.835
Financial	Gramm-Leach-Bliley Act	60.609	57.073	58.819
Political	Clinton announces US join NATO bombing in Kosovo	46.312	44.567	40.732
Financial	LTCM default	48.506	52.102	48.207
Financial	Russian default	64.753	60.621	53.668
Political	US embassy bombings in Kenia and Tanzania	59.333	57.087	55.489
Financial	Indonesian currency devalues strongly	31.342	29.124	33.260
Financial	Thailand unpegs currency	20.198	16.609	13.365
Financial	Start of Asian Crisis	21.941	22.026	20.198
Political	Clinton re-elected	13.185	10.030	10.089
	MEAN	25.401	24.949	23.426
	STDEV	21.258	21.154	21.466

(a) Panel B

Table 2.16: Ex-ante and Ex-post MAPE of Put of Financial and Political Events

Type of Event	Event	MAPE call t-1	MAPE call t	MAPE call t+1
Financial	Greece defaulting on IMF loan	0.684	0.303	0.787
Financial	Chinese stock market crash	0.566	0.498	7.239
Financial	Swiss National Bank abolishes currency peg	0.443	1.054	1.375
Political	Charlie Hebdo attack	0.428	1.489	2.306
Political	Israel-Gaza conflict	0.650	0.544	9.454
Political	IS seizes Mosul	54.642	52.673	47.553
Political	First Russian soldiers in Crimea	5.742	6.765	6.382
Political	9/11 Attack	0.325	1.531	1.233
Political	Maidan riots in Ukraine	0.715	0.482	4.904
Political	German Constitutional Court consults EU Court on OMT	2.003	20.367	21.454
Political	Ukraine rejects EU association agreement	11.372	5.018	0.534
Financial	Yellen nominated chairwoman of the FED	0.506	27.027	4.257
Political	US government shutdown	0.620	26.452	25.773
Political	Ousting of Mursi in Egypt	5.844	4.425	0.496
Political	Boston marathon bombing	36.320	26.773	0.664
Financial	US technically falls off Fiscal Cliff	0.470	42.927	8.863
Political	Obama re-elected	23.962	19.578	6.305
Financial	German Court approves ESM	79.536	72.858	4.647
Financial	EU Banking Union	19.101	15.708	0.423
Political	Tuareg offensive in Mali after coup starts	12.181	15.655	14.877
Financial	Debt restructuring in Greece	0.323	8.032	0.362
Financial	2nd Economic Adjustment Programme for Greece	0.398	12.761	12.588
Political	Berlusconi resignation announced	3.255	1.529	9.070
Financial	Swiss frank pegged	16.228	0.434	0.747
Political	Fukushima evacuation order	77.357	68.640	51.462
Political	UN Security Council establishes no-fly zone in Libya	4.637	104.152	102.508
Political	Mubarak ousted in Egypt	0.783	0.484	10.272
Political	Ben-Ali leaves Tunisia	14.339	19.179	0.561
Financial	Dodd-Frank-Act signed by Obama	1.036	8.849	0.308
Financial	EFSF adopted	1.173	0.292	1.463

(a) Panel A

Table 2.17: Ex-ante and Ex-post MAPE of Put of Financial and Political Events

Type of Event	Event	MAPE call t-1	MAPE call t	MAPE call t+1
Financial	1st Economic Adjustment Programme for Greece	4.020	1.363	4.083
Financial	Downgrading of Greece + Portugal	9.764	6.791	0.407
Political	Obamacare' passes congress	0.358	1.905	14.338
Political	Iran Green Revolution	3.202	2.631	0.659
Financial	US Recovery and Reinvestment Act passes Congress	4.192	0.291	0.286
Political	Obama elected	4.006	3.615	2.888
Financial	TARP passes at the congress	0.738	0.815	0.547
Financial	Emergency Economic Stabilisation Act, US bails out banks	1.901	1.686	0.738
Financial	Emergency lending to AIG	6.565	0.194	0.908
Financial	AIG asks for emergency lending + Lehman Brothers	9.523	7.357	6.565
Financial	Takeover of Bear Stearns by JP Morgan approved	0.171	0.103	0.116
Financial	Northern Rock receives liquidity support by BoE	0.985	0.943	1.832
Political	2nd Lebanon War	0.837	0.919	0.745
Financial	Bernanke nominated chairman of the FED	8.961	3.402	12.026
Political	London bombing	7.772	0.864	0.787
Political	Dutch referendum on EU constitution	0.954	0.956	0.966
Political	George W. Bush re-elected	2.591	3.276	3.263
Political	Madrid train bombings	15.583	3.568	10.856
Political	Iraq War	4.816	5.832	8.704
Political	Argentine, beginning of turmoil	4.602	3.679	4.163
Political	G W Bush elected	0.822	0.990	0.892
Political	George W. Bush declared winner in Florida	2.425	2.378	2.083
Political	Second Intifada	31.775	30.092	41.518
Financial	Gramm-Leach-Bliley Act	0.962	0.989	0.990
Political	Clinton announces US join NATO bombing in Kosovo	0.864	0.869	0.939
Financial	LTCM default	1.634	4.182	3.514
Financial	Russian default	3.079	1.380	1.893
Political	US embassy bombings in Kenia and Tanzania	2.583	3.225	2.888
Financial	Indonesian currency devalues strongly	0.970	3.771	0.942
Financial	Thailand unpegs currency	0.999	1.000	1.000
Financial	Start of Asian Crisis	0.997	0.980	0.999
Political	Clinton re-elected	1.000	1.000	1.000
	MEAN	8.442	10.918	8.052
	STDEV	16.111	19.668	16.164

(a) Panel B

Table 2.18: Ex-ante and Ex-post RMSE of Puts of Financial and Political Events

Type of Event	Event	RMSE put t-1	RMSE put t	RMSE put t+1
Financial	Greece defaulting on IMF loan	9.830	7.369	9.817
Financial	Chinese stock market crash	6.534	5.899	11.615
Financial	Swiss National Bank abolishes currency peg	5.911	10.481	10.684
Political	Charlie Hebdo attack	5.321	11.310	11.137
Political	Israel-Gaza conflict	12.527	10.913	18.716
Political	IS seizes Mosul	17.982	17.642	17.303
Political	First Russian soldiers in Crimea	10.487	10.711	10.346
Political	9/11 Attack	12.639	31.235	27.444
Political	Maidan riots in Ukraine	19.074	5.372	11.756
Political	German Constitutional Court consults EU Court on OMT	23.011	33.669	32.524
Political	Ukraine rejects EU association agreement	14.155	10.007	4.105
Financial	Yellen nominated chairwoman of the FED	11.039	19.560	14.703
Political	US government shutdown	20.655	20.532	20.236
Political	Ousting of Mursi in Egypt	15.196	13.558	8.642
Political	Boston marathon bombing	16.707	15.577	0.766
Financial	US technically falls off Fiscal Cliff	5.527	8.974	5.998
Political	Obama re-elected	25.642	24.954	20.242
Financial	German Court approves ESM	20.657	19.731	11.585
Financial	EU Banking Union	12.731	12.410	5.063
Political	Tuareg offensive in Mali after coup starts	7.596	7.057	6.958
Financial	Debt restructuring in Greece	3.896	6.922	2.331
Financial	2nd Economic Adjustment Programme for Greece	6.173	13.363	12.808
Political	Berlusconi resignation announced	36.379	36.388	29.584
Financial	Swiss frank pegged	17.939	12.148	11.197
Political	Fukushima evacuation order	15.397	15.103	12.859
Political	UN Security Council establishes no-fly zone in Libya	12.958	18.945	18.639
Political	Mubarak ousted in Egypt	13.007	10.069	19.900
Political	Ben-Ali leaves Tunisia	17.696	17.605	5.474
Financial	Dodd-Frank-Act signed by Obama	11.264	14.606	7.195
Financial	EFSF adopted	17.282	8.616	13.934

(a) Panel A

Table 2.19: Ex-ante and Ex-post RMSE of Puts of Financial and Political Events

Type of Event	Event	RMSE put t-1	RMSE put t	RMSE put t+1
Financial	1st Economic Adjustment Programme for Greece	7.567	3.656	7.669
Financial	Downgrading of Greece + Portugal	10.369	10.270	2.877
Political	Obamacare' passes congress	7.758	8.381	10.044
Political	Iran Green Revolution	92.550	85.275	78.253
Financial	US Recovery and Reinvestment Act passes Congress	23.100	9.842	12.977
Political	Obama elected	20.376	20.753	17.050
Financial	TARP passes at the congress	33.867	39.412	36.406
Financial	Emergency Economic Stabilisation Act, US bails out banks	37.058	36.687	33.867
Financial	Emergency lending to AIG	27.398	15.862	14.138
Financial	AIG asks for emergency lending + Lehman Brothers	51.616	38.315	27.398
Financial	Takeover of Bear Stearns by JP Morgan approved	37.814	31.598	28.460
Financial	Northern Rock receives liquidity support by BoE	71.826	73.039	69.251
Political	2nd Lebanon War	48.772	53.657	59.824
Financial	Bernanke nominated chairman of the FED	47.380	44.341	43.400
Political	London bombing	48.138	48.795	44.276
Political	Dutch referendum on EU constitution	33.180	31.664	31.069
Political	George W. Bush re-elected	56.863	53.216	51.631
Political	Madrid train bombings	25.814	23.550	21.638
Political	Iraq War	58.984	51.079	49.944
Political	Argentine, beginning of turmoil	43.845	46.913	45.946
Political	G W Bush elected	85.611	79.876	84.791
Political	George W. Bush declared winner in Florida	135.326	143.342	143.180
Political	Second Intifada	127.811	126.732	138.920
Financial	Gramm-Leach-Bliley Act	98.663	94.542	95.723
Political	Clinton announces US join NATO bombing in Kosovo	69.719	66.767	62.435
Financial	LTCM default	83.303	83.432	74.478
Financial	Russian default	100.955	100.271	90.866
Political	US embassy bombings in Kenia and Tanzania	98.848	95.741	95.106
Financial	Indonesian currency devalues strongly	42.048	37.871	44.306
Financial	Thailand unpegs currency	31.331	26.379	21.785
Financial	Start of Asian Crisis	33.404	32.874	31.331
Political	Clinton re-elected	18.437	14.223	14.213
	MEAN	34.792	33.571	31.957
	STDEV	31.555	31.425	31.953

(a) Panel B

Table 2.20: List of Financial and Political Events

Type of Event	Event	Type of Event	Event
Financial	Greece defaulting on IMF loan	Financial	1st Economic Adjustment Programme for Greece
Financial	Chinese stock market crash	Financial	Downgrading of Greece + Portugal
Financial	Swiss National Bank abolishes currency peg	Political	Obamacare' passes congress
Political	Charlie Hebdo attack*	Political	Iran Green Revolution
Political	Israel-Gaza conflict	Financial	US Recovery and Reinvestment Act passes Congress
Political	IS seizes Mosul*	Political	Obama elected
Political	First Russian soldiers in Crimea	Financial	TARP passes at the congress
Political	Maidan riots in Ukraine	Financial	Emergency Economic Stabilisation Act, US bails out banks
Political	German Constitutional Court consults EU Court on OMT	Financial	Emergency lending to AIG
Political	Ukraine rejects EU association agreement*	Financial	AIG asks for emergency lending + Lehman Brothers*
Financial	Yellen nominated chairwoman of the FED	Financial	Takeover of Bear Stearns by JP Morgan approved
Political	US government shutdown	Financial	Northern Rock receives liquidity support by BoE*
Political	Ousting of Mursi in Egypt	Political	2nd Lebanon War
Political	Boston marathon bombing*	Financial	Bernanke nominated chairman of the FED
Financial	US technically falls off Fiscal Cliff	Political	London bombing*
Political	Obama re-elected	Political	Dutch referendum on EU constitution
Financial	German Court approves ESM*	Political	George W. Bush re-elected
Financial	EU Banking Union	Political	Madrid train bombings*
Political	Tuareg offensive in Mali after coup starts	Political	Iraq War
Financial	Debt restructuring in Greece	Political	Argentine, beginning of turmoil
Financial	2nd Economic Adjustment Programme for Greece	Political	G W Bush elected
Political	Berlusconi resignation announced*	Political	George W. Bush declared winner in Florida
Financial	Swiss frank pegged	Political	Second Intifada
Political	Fukushima evacuation order*	Financial	Gramm-Leach-Bliley Act
Political	UN Security Council establishes no-fly zone in Libya	Political	Clinton announces US join NATO bombing in Kosovo
Political	Mubarak ousted in Egypt	Financial	LTCM default*
Political	Ben-Ali leaves Tunisia	Financial	Russian default
Financial	Dodd-Frank-Act signed by Obama	Political	US embassy bombings in Kenia and Tanzania*
Financial	EFSF adopted*	Financial	Indonesian currency devalues strongly
Financial	Thailand unpegs currency*	Financial	Start of Asian Crisis
Political	9/11 Attack*	Political	Clinton re-elected

This table represents events sourced from (Bloom 2009), where comprises 38 these events for the identification of the VAR model estimated on about 400 monthly observations. Our table consists of 62 events, and our baseline analysis from the dataset horizon consists of 16 baseline events which are identified with an asterisk(\*).

## Chapter 3

### EXTREME VALUE THEORY-MANAGED PORTFOLIOS

## Abstract

This chapter proposes, analyses, and implements a trading strategy that leverages volatility to reduce portfolio tail risk while improving the Sharpe ratio. The portfolio is constructed using the Fama-French Five-Factor Model, which includes the original three factors, along with profitability and investment, assuming a linear relationship between these factors and stock returns. The analysis compares this strategy against benchmark portfolios: buy-and-hold and the Volatility-Managed (VM) strategy from Moreira & Muir (2017). The motivation for this chapter lies in the advantages of Extreme Value Theory (EVT), EVT is more effective at assessing the risk of extreme losses by focusing on the distribution's extreme tails, whereas VaR often underestimates this risk due to its assumption of normally distributed returns. EVT's flexible assumptions about tail distributions make it more accurate in capturing the asymmetry and fat tails of return distributions. This Chapter makes a valuable contribution by promoting the integration of Extreme Value Theory (EVT) in Volatility-Managed portfolio literature. It highlights the benefits of EVT in improving risk management and volatility-timing, leading to superior financial performance, especially in the context of market volatility and extreme events.

## Abbreviation

**VaR** Value-at-Risk

**CVaR** Conditional Value-at-Risk

**EVA** Extreme Value Analysis

**EVT** Extreme Value Theory

**EVTM** Extreme Value Theory-Managed

**VM** Volatility-Managed

**TRM** Tail Risk-Managed

**ARCH** Autoregressive Conditional Heteroscedasticity

**GARCH** Generalised Autoregressive Conditional Heteroscedasticity

**POT** Peaks-Over-Threshold

**RSJV** Relative Signed Jump Variation

**ADF** Augmented Dickey–Fuller Test

**ME** Mean Excess

**MVE** Mean-Variance Efficient

## Nomenclature

$t$	Current time
$F_u$	Distribution of the excess value over a threshold $u$
$\beta$	Scale of the generalised Pareto
$\xi$	Shape parameter of the generalised Pareto
$\alpha$	Specific probability quantile
$\mathcal{L}$	Log-likelihood function
$\tilde{\alpha}$	Set of discrete values of $\alpha$
$RV_t$	Realised variance
$SJ_t$	Signed jump variation at time $t$
$s_t$	Relative Signed Jump Variation (RSJV)
$R_t$	Return of an arbitrary asset on day $t$
$S_t$	Price of an arbitrary asset on day $t$
$f_t$	Return of a portfolio or Fama-French factor on day $t$
$\kappa_t(\alpha)$	A probability threshold which measures the extent of the mis-classification risk associated with the forecast of the probability $\alpha$ of an EVT violation

- $\gamma_t$  Denotes exposure of the strategy, calculated by scaling the variance of  $f_{t+1}^\zeta$  to the variance of the Fama-factor computed on an expanded window starting from  $t_0$
- $w_i$  Window size where  $i$  specifies which window. i.e.  $w_1 = 1000$ , and  $w_2 = 250$
- $\Omega_t$  Weights of a portfolio

### 3.1 Introduction

In this chapter, we propose, analyse, and implement a trading strategy that exploits directional predictability, aiming to reduce a portfolio's exposure to tail risk while improving its Sharpe ratio through mean-variance analysis.

Our analysis is motivated by the perspective of a mean-variance investor, who dynamically adjusts their portfolio allocation based on the attractiveness of the mean-variance trade-off, represented by the ratio of expected return to variance. Existing portfolio management strategies often fail to adequately address tail-risk, which can lead to significant losses during financial crises or market downturns. To address this gap, our Extreme Value Theory Managed (EVTM) portfolios explicitly incorporate Extreme Value Theory (EVT) to model and manage tail-risk, offering a more robust framework for mitigating extreme losses compared to traditional risk management approaches like Value-at-Risk (VaR) or standard deviation-based methods. While traditional models assume returns are normally distributed, financial returns often exhibit fat tails and skewness, making extreme events more likely than predicted by conventional models. EVTM portfolios leverage the predictability of extreme events through EVT, which is better suited for modelling tail behaviour, enabling more accurate estimation of tail-risk probabilities and improved portfolio adjustments during volatile periods.

Many existing strategies either focus on maximising returns or minimising risk, but few achieve an optimal balance, especially during periods of market stress. EVTM portfolios aim to deliver superior risk-adjusted returns by dynamically adjusting risk exposure based on

extreme value tail-risk forecasts, ensuring that the portfolio is better positioned to handle adverse market conditions without sacrificing long-term performance. Traditional portfolio strategies often advocate increasing or maintaining risk exposure during downturns, assuming mean reversion or market recovery. In contrast, EVTm portfolios challenge this conventional wisdom by reducing risk-taking during adverse times, thereby protecting capital and avoiding large drawdowns. This contrarian approach can lead to better long-term performance.

The EVTm framework is highly flexible, allowing for customisation based on different modelling choices, such as threshold selection and distribution assumptions. This adaptability makes it applicable across various markets and asset classes, enhancing its practical utility. While EVT is widely used in risk management, its application in portfolio management remains under-explored. By empirically validating the effectiveness of EVT in portfolio management, our research bridges a gap in the literature and provides a new tool for investors and portfolio managers. Additionally, EVTm portfolios address behavioural finance biases, such as overconfidence or loss aversion, by providing a systematic, rules-based approach to managing tail-risk, thereby improving decision-making during periods of market stress.

Existing strategies may not adequately prepare portfolios for extreme market events or prolonged periods of market turbulence. EVTm portfolios enhance portfolio resilience by explicitly accounting for extreme events, ensuring that the portfolio can withstand severe market shocks without catastrophic losses. Finally, as the volatility-managed literature continues to evolve, there is a need for innovative approaches that address the limitations of existing models. By introducing EVTm portfolios, our research contributes to the broader literature on portfolio management, and tail-risk modelling, offering a novel perspective and

practical insights for both academics and practitioners.

Our portfolio construction is based on the Fama-French Five-Factor Model from Fama & French (2015), which extends the original Three-Factor Model (Fama & French (1992)) by including two additional factors: profitability and investment. We chose to use the Fama-French five-factor model due to its strong theoretical foundations, empirical robustness, and alignment with our research objectives. The model builds on the well-established three-factor framework by incorporating profitability (RMW) and investment (CMA) factors. The model's proven explanatory power across various markets make it a reliable choice, as it captures the most significant anomalies for example, size, value, profitability, and investment. While other studies may include additional factors to address specific anomalies or explore niche areas, our research focuses on the core drivers of returns, ensuring clarity, interpretability, and comparability with a vast body of existing literature. Furthermore, limiting the analysis to five factors avoids potential issues related to data availability, ensuring the robustness of my findings.

We compare these portfolios against benchmarks, including buy-and-hold portfolios and Volatility-Managed (VM) portfolios from Moreira & Muir (2017). EVT is specifically designed to assess the risk of extreme losses by focusing on the tail of the distribution, unlike VaR, which may underestimate extreme risks when the underlying data deviates from normal distribution assumptions. EVT is particularly effective because it relies on flexible assumptions about tail distributions, whereas VaR assumes a normal distribution of returns, a limitation given that financial returns often exhibit fat tails and asymmetry.

This study explores various portfolio optimisation strategies, with a specific focus on the application of the Sortino ratio, Sharpe ratio, and EVT-VaR. We aim to provide compelling evidence of the superior performance of portfolios managed using the Sharpe ratio and Sortino ratio methods. Our results show that EVTMM portfolios optimised by the Sharpe ratio consistently exhibit higher Sharpe ratios across all factors compared to VM and unmanaged portfolios. Sortino-optimised EVTMM portfolios also deliver exceptional results, with significantly higher Sortino-optimised EVTMM Sharpe ratios than VM portfolios, highlighting the importance of incorporating downside risk into portfolio management.

Additionally, the Conditional Value-at-Risk (CVaR) analysis demonstrates that EVTMM portfolios manage extreme downside risks more effectively than VM portfolios, particularly at the 99% and 95% confidence levels. The presence of positive skewness in all EVTMM portfolios, contrasted with the negative skewness in VM portfolios, indicates a more favourable risk-return profile for EVTMM portfolios. The kurtosis analysis further suggests that EVTMM portfolios, optimised with Sharpe and Sortino ratios, exhibit a more moderate distribution of returns compared to the heavier-tailed VM portfolios.

The time-series spanning regression analysis supports these findings, with real-time single-factor and multi-factor EVTMM portfolios delivering universally positive and significant alphas. Sharpe-optimised EVTMM portfolios, in particular, exhibit alphas more than three times higher than those of VM portfolios. The higher  $R^2$  values observed in EVTMM portfolios indicate that their returns are sufficiently explained by the factors considered, reinforcing the effectiveness of EVTMM as a portfolio management strategy.

The study also examines break-even transaction costs, revealing that despite the higher costs associated with EVT<sub>M</sub> portfolios ranging from 60 to 140 basis points for single-factor portfolios, and 130 to 160 basis points for multi-factor portfolios, these portfolios still demonstrate superior performance in terms of Sharpe ratios and alphas. The positive break-even points across all EVT<sub>M</sub> portfolios highlight their resilience to transaction costs.

According to Prospect Theory, developed by Kai-Ineman & Tversky (1979), investors are more sensitive to losses than to gains (loss aversion). During volatile or turbulent market conditions, they tend to prefer safer, less volatile assets, while in calmer periods, they are more willing to pursue riskier assets in search of higher returns. By adopting EVT<sub>M</sub> portfolios and dynamically adjusting portfolio allocations in response to market fluctuations, investors can optimise risk-adjusted returns and mitigate the adverse effects of volatility. This adaptive approach enables more resilient portfolio performance across varying market environments, strategically balancing risk and reward.

The remainder of the Chapter is organised as follows: Section 3.2 provides an overview of relevant research on EVT and EVT-VaR estimation methods, as well as volatility portfolio management. Section 3.3 describes the daily data of the Fama-French factors Mkt-RF, SMB, HML, RMW, and CMA, including descriptive statistics sourced from Wharton Research Data Services (WRDS). Section 3.4 details the specific steps used to construct both single-EVT<sub>M</sub> and multi-EVT<sub>M</sub> portfolios, along with a discussion on break-even transaction costs. Section 3.5 presents the empirical analysis, which investigates the selection of optimal thresholds, the implementation of Logit regression analysis, the evaluation of risk using Conditional Value-at-Risk (CVaR), and portfolio performance through Sharpe ratio metrics. We also

explore spanning regressions and break-even transaction costs. Finally, Section 3.6 focuses on the empirical conclusions of our results.

### 3.2 Literature Review

Extreme Value Analysis (EVA) focuses on the tail behaviour of distributions, which is crucial in financial markets for understanding the risk of extreme losses or gains. The seminal contributions by Fréchet (1927), Fisher & Tippett (1928), and Gnedenko (1943) laid the groundwork for modern EVA. Their work introduced the Fisher-Tippett-Gnedenko theorem, which characterises the convergence of normalised maxima to a Generalised Extreme Value (GEV) distribution. The GEV distribution encompasses the Gumbel, Fréchet, and Weibull distributions, each representing different types of extreme value behaviour.

One common approach in EVA is the Peak Over Threshold (POT) method, which models exceedances over a high threshold using the Generalised Pareto Distribution (GPD). This method has been extensively used to model extreme returns in stock markets. For instance, McNeil & Frey (2000) demonstrated the effectiveness of the POT method in estimating Value-at-Risk (VaR) and Expected Shortfall (ES) for financial returns. Another approach, the Block Maxima method, involves dividing the data into blocks (e.g., months or years) and modelling the maximum value within each block using the GEV distribution. While less frequently used compared to POT, this method has been applied in stock markets to model annual maximum returns, as seen in the work of Longin (2000).

Despite its advantages, EVA has limitations. The selection of appropriate thresholds in the POT method and block sizes in the Block Maxima method can be challenging and subjective. Additionally, the assumption of stationarity in financial returns is often violated, requiring

advanced techniques to account for non-stationary and dynamic behaviour.

Combining EVT with VaR (EVT-VaR) provides a robust framework for estimating risk measures, particularly in the tails of the distribution where traditional methods might fail. This combination is useful for capturing the risk of rare but severe market movements. McNeil & Frey (2000) demonstrated that EVT provides better tail risk estimates than traditional methods, particularly in financial markets characterised by fat tails and volatility clustering. They proposed a method for estimating VaR that combines pseudo-maximum-likelihood fitting of GARCH models to estimate current volatility with EVT to estimate the tail of the innovation distribution in the GARCH model. This method estimates conditional quantiles (VaR) and conditional expected shortfalls, and backtesting of historical daily return series shows that it provides better 1-day estimates than methods that ignore the heavy tails of the stochastic nature of volatility.

Longin (2000) applied EVT to stock index returns and showed that EVT-based VaR estimates were more accurate in predicting extreme losses compared to standard techniques. Danielsson & Vries (1997) examined the performance of EVT in estimating market risk for daily returns of financial assets, concluding that EVT significantly enhances the precision of risk estimates. They showed that conditional parametric methods, such as GARCH with normal innovations, under-predict VaR for a sample of U.S. stock returns.

Despite its advantages, the application of EVT-VaR faces challenges, particularly with regard to threshold selection. The choice of threshold in the POT method is crucial and subjective, and it impacts the accuracy of the GPD fit.

Baran & Witzany (2011) demonstrate that EVT-VaR is an improvement over previous VaR methodologies, particularly because EVT focuses on the tails of the distribution. They apply EVT to estimate low quantiles of the profit and loss (P/L) distribution and compare these results with common Value-at-Risk (VaR) methodologies. Building on the fundamental principles of EVT, they use the Peaks-Over-Threshold method to model the tail of the loss distribution with the Generalised Pareto Distribution (GPD).

Gençay & Selçuk & Ulugülyağci (2003) adopted EVT to address the problem of underestimation and overestimation in computing VaR. They assert that most financial return series are asymmetric, and EVT is advantageous over models that assume symmetric distributions, such as t-distributions and normal distributions, which are typically used in ARCH and GARCH-like models, with the exception of the EGARCH model by Nelson (1992), which accounts for asymmetric distributions.

The foundation of portfolio management theory was first developed by Markowitz (1952) with the introduction of Modern Portfolio Theory (MPT) on portfolio selection, which emphasised the importance of diversification and the relationship between risk and return. This theory introduced the concept of the efficient frontier, illustrating a set of optimal portfolios that offer the highest expected return for a given level of risk. Building on Markowitz (1952) work, the Capital Asset Pricing Model (CAPM) developed by Sharpe (1964) introduced the notion of systematic and idiosyncratic risk, asserting that only systematic risk is rewarded with higher expected returns. The model introduced the concept of beta, a measure of an asset's sensitivity to market movements, and the security market line, which describes the expected return of an asset as a function of its beta.

As financial markets evolved, the development of multi-factor models, such as the Fama-French three-factor model by Fama & French (1996), included the market risk premium (Mkt-RF), the excess returns of small-cap stocks over large-cap stocks (SMB), and the excess returns of value stocks over growth stocks (HML), providing a more comprehensive explanation of asset returns. Subsequent research by Fama & French (2015) expanded these models to include factors that account for returns on robust operating profitability over weak operating profitability (RMW), and the returns of portfolios long on stocks with conservative investments and short on stocks with aggressive investments (CMA).

A significant advancement is the concept of risk parity. E. Qian (2011) focuses on allocating risk rather than capital equally across the portfolio components. This approach aims to construct portfolios where each asset class contributes equally to the overall portfolio risk, potentially offering more stable returns in diverse market conditions.

Empirical studies in portfolio management often focus on the performance of different asset allocation strategies, the impact of diversification, and the effectiveness of various risk management techniques. For example, studies have shown that while diversification can reduce risk, the benefits diminish beyond a certain number of assets Statman (1987).

Moreira & Muir (2017) investigate the performance and benefits of Volatility-Managed (VM) portfolios in the context of asset pricing and risk management. The authors explore how dynamic allocation strategies based on volatility can improve risk-adjusted returns and enhance portfolio diversification. The principal focus of the paper is on developing and evaluating Volatility-Managed portfolios compared to traditional buy-and-hold strategies.

Moreira & Muir (2017) provide an economic understanding of the effectiveness of Volatility-Managed portfolios. They argue that the strategy is grounded in the theory that investors prefer less volatile assets during turbulent market conditions and seek riskier assets when conditions are calmer. By dynamically adjusting the portfolio's allocation, investors can achieve better risk-adjusted returns and mitigate the impact of market fluctuations.

In addition, Barroso & Detzel (2021) assess how VM strategies perform poorly in the presence of trading costs. DeMiguel & Martin-Utrera & Uppal (2021) show that the superior performance of a VM strategy is maintained in real-time when multi-factor portfolios are considered. Cederburg et al. (2020) evaluate equity strategies to determine the effectiveness of Volatility-Managed portfolios for real-time investors. Their portfolios tend to show significantly positive alphas in spanning regressions, aligning with Moreira & Muir (2017). Barroso & Detzel (2021) also examine whether limits to arbitrage explain the abnormal returns of Volatility-Managed portfolios, finding that these abnormal returns are negligible in long-only portfolios of hard-to-arbitrage stocks. The utility function from volatility managed portfolios are significantly higher for out-of-sample mean-variance-efficient portfolios composed of easy-to-arbitrage stocks than for hard-to-arbitrage stocks. This contradicts the typical finding that anomalies are stronger where arbitrage is difficult. Additionally, the abnormal returns of Volatility-Managed portfolios are significant only during times of high liquidity and sentiment, aligning with models suggesting that unsophisticated traders under-react to informed order flow during such periods.

F. Wang & X. S. Yan (2021) investigate the performance of Volatility-Managed portfolios and find that portfolios scaled by downside volatility perform significantly better than those scaled

by total volatility. This improved performance is demonstrated through spanning regressions, direct Sharpe ratio comparisons, and real-time trading strategies. The enhanced performance of downside Volatility-Managed portfolios is primarily attributed to return timing, where downside volatility negatively predicts future returns. Additionally, the study shows that using fixed-weight strategies can significantly enhance the performance of Volatility-Managed portfolios for real-time investors. Qiao & S. Yan & B. Deng (2020) find that downside volatility and overall volatility generally move together but are not highly correlated during periods of high volatility.

Urom et al. (2022) develop a modified Sortino ratio, which presents a criterion for minimising downside risk for any chosen expected return and loss benchmark. Their method allows investors to delineate the optimal mix of risky and risk-free components in the portfolio. They select a threshold  $T(\gamma)$ , which mimics the portfolio composition in the sense that it equals the risk-free rate plus  $\gamma$  times the portfolio's equity risk premium. Portfolios scaled by downside volatility enhance the mean-variance frontiers beyond those of original and Volatility-Managed portfolios, improving the Sharpe ratios of tangency portfolios.<sup>1</sup> This is because downside Volatility-Managed portfolios are not encompassed by original or Volatility-Managed portfolios, thereby expanding the investment opportunity set. In contrast, portfolios managed by upside volatility do not provide this benefit.

De Nicolo (2023) contributes to the literature by focusing on tail risk and developing the Tail Risk Managed (TRM) portfolio. A key difference between TRM portfolios and VM portfolios

---

<sup>1</sup>A tangency portfolio lies at the point where the efficient frontier is tangent to the highest possible Capital Market Line (CML) in the risk-return space.

is that the TRM strategy exploits a predictive model of targeted levels of tail risks. Therefore, a TRM strategy is more flexible than a VM strategy, as it is based on an expanded set of tail risk predictors, whereas the VM strategy relies on a function of the conditional volatility of returns. Consequently, certain elements of the construction of the EVTMM portfolio method are motivated by De Nicolo (2023).

We take this further by deploying Extreme Value Analysis (EVA) to focus on a framework that deals with the analysis of extreme events in probability distributions by assessing statistical models for tail-related risk measures. Ultimately, we are able to model, capture, and understand the behaviour of rare or extreme events. Moreira & Muir (2017) found that Volatility-Managed (VM) portfolios, which reduce risk during high-volatility periods, generate significant excess returns, improve Sharpe ratios, and yield substantial utility gains for mean-variance investors. These findings are significant for portfolio management, suggesting that adjusting portfolio risk in response to market volatility can lead to better risk-adjusted returns.

Similarly, De Nicolo (2023) assesses single-factor and multi-factor portfolios and uses the Manipulation-Proof Performance Metric (MPPM) proposed by Goetzmann et al. (2007). This metric, derived from the expected utility function of an investor with a constant relative risk aversion coefficient, captures the potential benefits of tail risk management. The analysis shows that, relative to the MPPMs of multi-factor VM portfolios, multi-factor TRM portfolios exhibit higher MPPMs for all the international TRM multi-factor portfolios.

Božović (2024) applies a volatility timing strategy using an implied volatility index and

evaluates the proposed VIX-based strategy against a broad set of realistic transaction costs, demonstrating its robustness. The strategy provides a mechanism for simultaneous timing of volatility and tail risk. Božović (2024) proposes using four methods to evaluate single-factor and multi-factor portfolios, which include a VIX-managed strategy, a downside VIX-managed strategy, a realised variance-managed strategy, and a downside realised variance-managed strategy. He finds that VIX-managed portfolios require the least rebalancing among the tested strategies and are, therefore, the least burdened by frequent trading costs. Additionally, implied volatility smiles add an extra layer in timing extreme returns on top of volatility.

Our motivation stems from the many advantages EVT possesses. The potential problems that arise with the  $\text{VaR}_\alpha$  method are discussed by Odening & Hinrichs (2003), who explain that  $\text{VaR}_\alpha$  can fail when the return distribution is fat-tailed. EVT overcomes these problems. This Chapter contributes to the literature as the first study applying a risk and volatility timing strategy using a Extreme Value Theory.

### 3.3 Data Description

Economic data needed for the simulation and analysis of EVT<sub>M</sub> portfolios is obtained from Wharton Research Data Services ([WRDS](#)). The US sample includes daily data of the Fama-French factors Mkt-RF, SMB, HML, RMW, and CMA from July 1st, 1963 to January 26th, 2023. Mkt-RF represents the excess returns for the US equity market, SMB is small-minus-big, HML is high-minus-low, RMW is robust-minus-weak, and CMA is conservative-minus-aggressive. These factors are understood as the excess market return (Mkt-RF), size factor (SMB), value factor (HML), profitability factor (RMW), and investment factor (CMA). The Fama-French factors are part of the Fama-French five-factor model, developed by Fama & French ([2015](#)). The results cover the period from June 20th, 1968 to February 28th, 2023, yielding 13,769 forecasts.

For descriptive statistics and higher moments, we consider the mean, standard deviation, skewness, kurtosis, ADF, and Jarque-Bera statistics. From [Table 3.1](#) and [Figure 3.1](#), we observe negative skewness for the factors Mkt-RF, SMB, and CMA. However, SMB exhibits the highest kurtosis, indicating that a large proportion of the data resides in the tails. Consequently, this brings the tails closer to the mean. The ADF test implies we reject the null hypothesis that a unit root exists. Therefore, given that all Fama-French factors are negative and statistically significant, we assume all are stationary.

The Jarque-Bera test is a goodness-of-fit test to determine whether the sample data has skewness and kurtosis matching a normal distribution. The test statistic is always nonnegative,

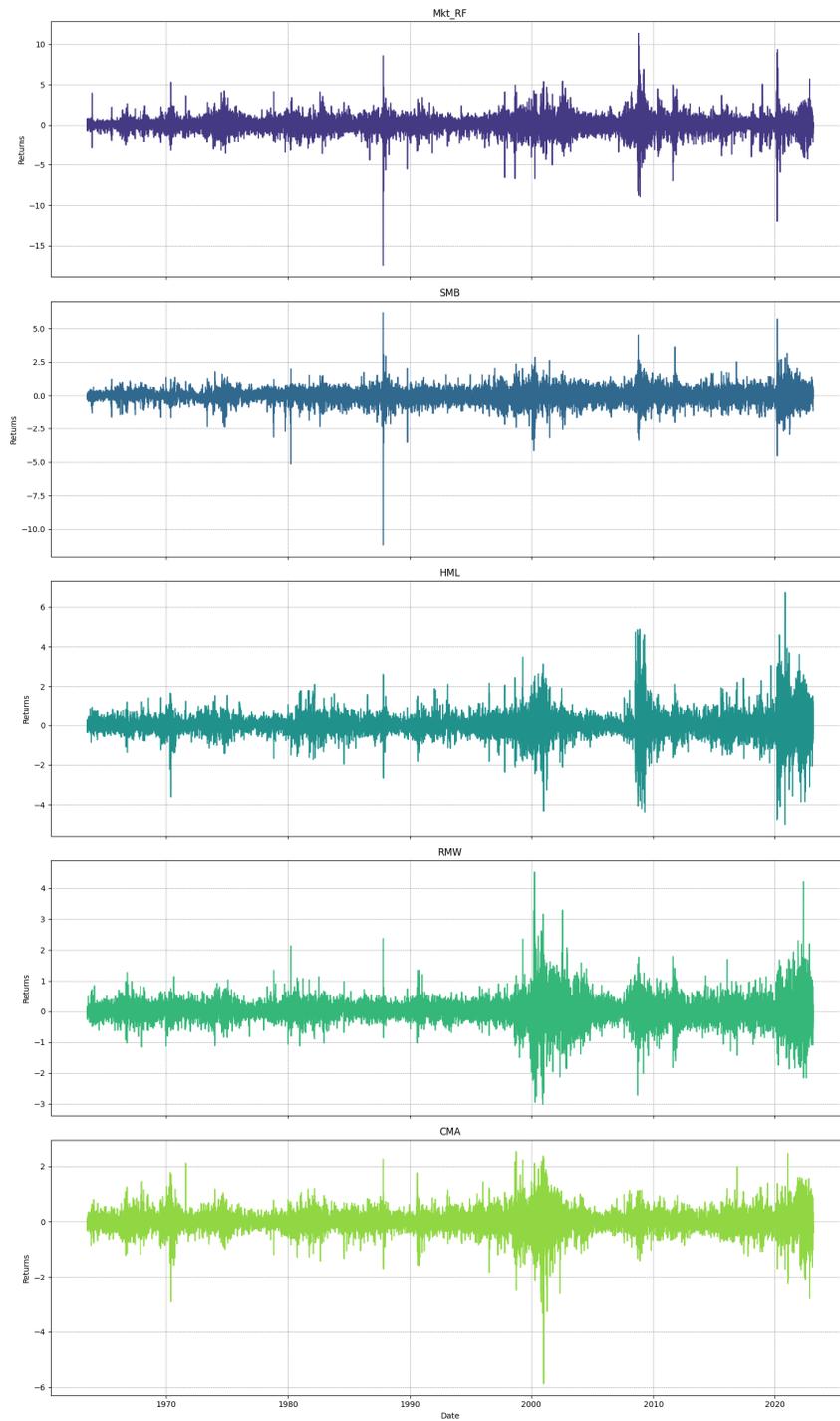
Table 3.1: Summary Statistics for Data

Fama-French Factor	Mean	Standard Deviation	Skewness	Kurtosis	ADF	Jarque-Bera Statistic
Mkt-RF	0.0003	0.0102	-0.5263	15.5464	-21.3814***	153651.3996***
SMB	0.0001	0.0054	-0.6862	17.8989	-21.9672***	203933.6077***
HML	0.0001	0.0058	0.27990	11.04473	-17.0097***	77395.1034***
RMW	0.0001	0.0040	0.3362	9.0501	-17.4335***	52117.7280***
CMA	0.0001	0.0038	-0.3294	9.2625	-18.5368***	54568.0854***

In this Table, we report statistics, including the ADF and Jarque-Bera Statistic of the Fama-French factors from July 1<sup>st</sup>, 1963 to February 28<sup>th</sup>, 2023, consisting of 15018 data points. Stars are only intended to flag levels of significance. If a p-value is  $\leq 0.05$  it is flagged with one star (\*), if a p-value is  $\leq 0.01$  it is flagged with two stars (\*\*), subsequently, if a p-value is  $\leq 0.001$  it is flagged with three stars (\*\*\*).

and if it is far from zero, it indicates that the data does not follow a normal distribution. The Jarque-Bera test for all Fama-French factors is significant, with SMB presenting the largest test statistic. Given that all Jarque-Bera test statistics are far from zero, this confirms that the data does not follow a normal distribution.

Figure 3.1: Fama-French charts



These plots are an illustration of the Mkt-RF, SMB, HML, RMW and CMA Fama-French factor time series. This plot consists of data within the range July 1<sup>st</sup>, 1963 to February 28<sup>th</sup>, 2023.

## 3.4 Methodology

The methodology section outlines the detailed steps and approaches employed in constructing both the single-factor and multi-factor EVT<sub>M</sub> portfolios, while also providing a discussion on break-even transaction costs.

### 3.4.1 EVT-VaR Peaks-Over-Threshold

The Peaks-Over-Threshold (POT) method is used to model extreme values by focusing on the data that exceeds a certain threshold. The core principle of this method is to identify a threshold that separates extreme values from the rest of the data, and then fit the extreme values to a model by examining the tail of the distribution for values exceeding this threshold. According to the Pickands–Balkema–De Haan theorem, as discussed in Balkema & Haan (1974) and Pickands (1975), it can be demonstrated that for a sufficiently large threshold  $u$ , the distribution of the values exceeding this threshold approximates to a Generalised Pareto Distribution.

We assume that market losses represent the realisation  $x$  of a random variable  $X$  that exceeds a high threshold  $u$ , and that  $X$  follows a cumulative distribution function  $F_u(x)$ , which represents the distribution of exceedances over the threshold  $u$ . When observed values surpass this threshold, they are referred to as exceedances, denoted as  $X(x_1, x_2, \dots, x_n)$ . The

conditional excess distribution is defined as:

$$F_u(x) = \mathbb{P} \left( X - u \leq x \mid X > u \right) = \frac{F(x + u) - Fu}{1 - F(u)}, \quad (3.1)$$

for  $0 \leq x \leq x_F - u$ , where  $x_F$  is the finite or infinite right endpoint of the cumulative distribution  $F$ . The function  $F_u$  characterises the distribution of values exceeding the threshold  $u$ , given that this threshold has been surpassed. According to Balkema & Haan (1974) and Pickands (1975), for a broad class of underlying distributions  $F$  and sufficiently large  $u$ ,  $F_u$  can be approximated by the Generalised Pareto Distribution (GPD), defined as:

$$G_{\xi, \beta(u)}(x) = \begin{cases} 1 - \left(1 + \frac{\xi x}{\beta(u)}\right)^{-\frac{1}{\xi}}, & \text{if } \xi \neq 0, \\ 1 - e^{-\frac{x}{\beta(u)}}, & \text{if } \xi = 0, \end{cases} \quad (3.2)$$

where the parameter  $\beta$  represents the scale and  $\xi$  the shape. We estimate the parameters using the Maximum Likelihood Estimation (MLE) method, by deriving the log-likelihood function and maximising this function to find suitable estimators. Other approaches include Markov Chain Monte Carlo (MCMC), Maximum Product of Spacing (MPS), and Method of Moments (MoM). From Equation 3.2, for  $\xi \neq 0$ , the log-likelihood function for an i.i.d. sample is given by,

$$\mathcal{L}(\xi, \beta) = -k \ln(\beta) - \left(1 + \frac{1}{\xi}\right) \sum_{i=1}^k \ln \left(1 + \frac{\xi x_i}{\beta}\right), \quad (3.3)$$

where  $x_i \geq 0$  for  $\xi > 0$  and  $0 \leq x_i \leq -\beta/\xi$ . When  $\xi = 0$ , the log-likelihood function simplifies

to:

$$\mathcal{L}(\xi, \beta) = -k \ln(\beta) - \left(\frac{1}{\beta}\right) \sum_{i=1}^k x_i. \quad (3.4)$$

As explained by Fernandez (2003), an estimate for the Extreme Value Theory Value-at-Risk (EVT-VaR) is defined as:

$$\widehat{EVT}^\alpha = u + \frac{\widehat{\xi}}{\widehat{\beta}} \left( \left( \frac{1-\alpha}{k/n} \right)^{-\widehat{\xi}} - 1 \right), \quad (3.5)$$

where  $k$  represents the number of exceedances over the threshold  $u$ , and  $n$  is the total number of observations (positive and negative returns).

The choice of the method, whether MLE, MPS, MCMC, or MoM, depends on the complexity of the problem, the size and nature of the dataset, and the available computational resources. Abdulali et al. (2022) reviewed and compared these estimation methods, particularly for parameter estimation from extreme value distributions. The study found that while MPS yielded better results in terms of Mean Squared Errors (MSE), its goodness-of-fit statistic was comparable to that of MLE. Thus, we employ the MLE approach for the GPD equation in Equation 3.2.

Evaluating threshold excesses, as described in Equation 3.1, often involves plotting the Mean Excess (ME), which is a drawback for systematic computation. Ghosh & Resnick (2010) suggest that the ME plot can be useful as a diagnostic tool for tail or quantile estimation in risk management and other extreme value analysis problems.

### 3.4.1.1 Summary Steps to Compute EVT-VaR

In simple terms, we implement the following steps for calculating each factor. For a rolling window  $w_1 = 1000$ , we calculate  $EVT_t^\alpha$ , the EVT-VaR for all  $\alpha$ . The calculation process is as follows:

1. **Defining the Threshold:** We determine the threshold  $u$  for the left tail by selecting the 5th percentile of the factor's distribution over every  $w_1$  window. This threshold  $u$  represents a threshold value below which the worst 5% of observations lie, thus identifying extreme negative returns.
2. **Extracting Exceedances:** After defining the threshold, we extract the exceedances by subtracting the values below the threshold from the threshold itself. This yields the conditional excess distribution  $F_u$ , which describes the distribution of returns exceeding the threshold.
3. **Fitting the Generalised Pareto Distribution (GPD):** We then fit a GPD to the exceedances, as expressed in Equation 3.2. Maximum Likelihood Estimation (MLE) is employed to estimate the GPD parameters for the specified quantiles  $\tilde{\alpha}$ .
4. **Computing Quantiles:** Using the GPD with the estimated parameters from the previous step, we compute the quantiles corresponding to the probability  $1 - \alpha$  for  $\alpha \in \tilde{\alpha}$ , where  $\tilde{\alpha} = \{0.01, 0.05, 0.10, 0.25\}$ .

5. **Estimating EVT-VaR:** Finally, the estimated quantile is subtracted from the threshold  $u$  to estimate the potential loss at each probability level  $\alpha$ , referred to as the EVT-VaR. This approach accounts for extreme tail risk, producing an extreme estimate of potential losses as  $\alpha$  increases.

### 3.4.2 EVTM Methodology

EVTM portfolios are constructed using probability forecasts obtained through Logit regressions. To compute  $EVT_t^\alpha$ , we determine the EVT-VaR by calculating the 5th percentile of the returns for the Fama-factor within the data window  $w_1$ . This approach eliminates the need for more subjective procedures, such as the mean excess plot, which requires visual inspection to determine the threshold. The 5th percentile is the choice for the threshold in EVT as it identifies the threshold for the most extreme 5% of the data points, representing tail events. Allen & Singh & Powell (2013) also utilise the lower 5% and 10% quantile thresholds ( $u$ ) to compute the Peaks Over Threshold (POT) EVT-VaR.

We compute the exceedances, which are the differences between the threshold and the values of the returns that are below the threshold. Essentially, we measure how far below the threshold each return is. We fit the Generalised Pareto Distribution (GPD) to the exceedances using MLE which is specified in Equations 3.3 and 3.4. This fitting process yields the GPD parameters, shape and scale, that best describe the tail behaviour of the data. Finally, we compute the quantiles for extreme events using the fitted GPD. By subtracting this quantile from the threshold, we obtain the EVT-based VaR for  $\alpha \in \tilde{\alpha}$ , where

$\tilde{\alpha} = \{0.01, 0.05, 0.10, 0.25\}$ .

### 3.4.2.1 Constructing EVT<sub>M</sub> Portfolios

We closely follow De Nicolo (2023) methodology to construct the EVT<sub>M</sub> portfolios. We compute the predicted conditional probability  $EVT_t^\alpha$  violation of the return  $R_{t+1}$  of a factor, denoted by,

$$\mathbb{P}(R_{t+1} < -\widehat{EVT}_t^\alpha | X_t), \quad (3.6)$$

where  $\widehat{EVT}_t^\alpha$  is the estimate of  $EVT_t^\alpha$  at date  $t$  and  $X_t$  is a set of predictors.

An EVT<sub>M</sub> portfolio scales up or down exposures to the original factor dynamically as a function of the difference between the predicted conditional probability of  $EVT_t^\alpha$  violations  $\mathbb{P}(R_{t+1} < -EVT_t^\alpha | X_t)$ , and its unconditional probability  $\alpha$ . The difference,

$$\mathbb{P}(R_{t+1} < -\widehat{EVT}_t^\alpha | X_t) - \alpha, \quad (3.7)$$

serves as a signal for trading decisions. If,

$$\mathbb{P}(R_{t+1} < -\widehat{EVT}_t^\alpha | X_t) - \alpha \geq 0, \quad (3.8)$$

the exposure to the factor is increased or decreased, with the scaling factor being proportional to  $\mathbb{P}(R_{t+1} < -\widehat{EVT}_t^\alpha | X_t) - \alpha$ . A positive difference implies increasing tail risk, while a negative difference indicates decreasing tail risk as measured by EVT. The probability  $\mathbb{P}(R_{t+1} < -\widehat{EVT}_t^\alpha | X_t)$  at the forecasting date is assumed to equal the true probability plus a

bias term:

$$P_{t+1}(\alpha) = \mathbb{P}(R_{t+1} < -EVT_t^\alpha) + \eta_t, \quad (3.9)$$

where  $\eta_t$  is the bias term or error term. In other words,  $P_{t+1}(\alpha)$  represents the probability that the factor's return is less than the negative EVT estimate for a given probability  $\alpha$ , inclusive of the bias term  $\eta_t$ . The term  $\eta_t$  reflects the estimation error associated with the probability estimate from the logit model.

We construct real-time single-factor EVTMM portfolios and multi-factor EVTMM portfolios for a set of US Fama-French factors<sup>2</sup> widely used in the literature. We evaluate EVTMM performance relative to the buy-hold portfolios and VM portfolios, which are also based on US Fama-French factors. The Sharpe ratio of each  $f_t^\zeta(\alpha)$  is computed over a rolling data window prior to time  $t$ , and the scaling factor  $\gamma_t^\alpha$  corresponding to the  $\alpha^*$  that yields the highest Sharpe ratio is selected.  $f_t^\zeta(\alpha)$  denotes the EVTMM portfolio with a specified  $\alpha$ .

We introduce a threshold  $\kappa_t(\alpha)$  defined as:

$$\kappa_t(\alpha) \equiv \alpha + \eta_t, \quad (3.10)$$

which acts as a probability threshold. The difference between the estimated probability  $\widehat{P}_{t+1}(\alpha)$  and the threshold  $\kappa_t(\alpha)$  is used to determine the direction of the tail risk prediction at the forecasting date. Specifically, this difference serves as the signal triggering trading

---

<sup>2</sup><https://wrds-www.wharton.upenn.edu/pages/get-data/Fama-French-portfolios-and-factors/Fama-French-portfolios/>

Table 3.2:  $\widehat{P}_{t+1}(\alpha) - \kappa_t(\alpha)$  Confusion Matrix

	$R_{t+1} < -\widehat{EVT}_t^\alpha$	$R_{t+1} \geq -\widehat{EVT}_t^\alpha$
$\widehat{P}_{t+1}(\alpha) - \kappa_t(\alpha) \geq 0$	$a_{11}(\kappa_t(\alpha))$	$a_{10}(\kappa_t(\alpha))$
$\widehat{P}_{t+1}(\alpha) - \kappa_t(\alpha) < 0$	$a_{01}(\kappa_t(\alpha))$	$a_{00}(\kappa_t(\alpha))$

This table illustrates a confusion matrix of the trading action signal relative to the EVT-VaR violation, where  $a_{11}(\kappa_t(\alpha))$  and  $a_{00}(\kappa_t(\alpha))$  are the frequencies of correct classification, and  $a_{01}(\kappa_t(\alpha))$  and  $a_{10}(\kappa_t(\alpha))$  are false negatives and false positives, respectively.

actions. If:

$$\widehat{P}_{t+1}(\alpha) - \kappa_t(\alpha) > 0, \quad (3.11)$$

or:

$$\widehat{P}_{t+1}(\alpha) - \kappa_t(\alpha) < 0, \quad (3.12)$$

this signals either a high or low probability of an EVT violation, respectively, as indicated by the Equations 3.11 and 3.12. The estimate of  $\kappa_t(\alpha) \in (0, 1)$  determines the degree of misclassification risk associated with the forecasted EVT violation probability. The confusion matrix in Table 3.2 explains the trading signal relative to EVT-VaR violations, where  $a_{11}(\kappa_t(\alpha))$  and  $a_{00}(\kappa_t(\alpha))$  are the frequencies of correct classification, while the other frequencies are false positives and false negatives.

The size of the bias term  $\eta_t$  directly influences the magnitude of misclassification risk. This risk can be minimised by selecting, at each date, the optimal threshold  $\kappa_t^*(\alpha)$  that corresponds to the minimum sum of Type 1 and Type 2 errors. These errors are represented by the terms  $a_{10}(\kappa_t(\alpha))$  and  $a_{01}(\kappa_t(\alpha))$ , and are derived from the confusion matrix estimated at each date. The optimisation process follows this logic: we calculate the sum of Type 1 and Type 2 errors from the confusion matrix (as shown in Table 3.2), and then iterate over a set of evenly

spaced threshold values between 0 and 1. Our goal is to minimise the sum of these errors for cases where our predictions exceed the threshold. Initially, the minimum error sum is set to infinity. If, during the loop, the current error sum is less than the minimum error sum, the current threshold becomes our optimal threshold. This process ensures that the selected  $\kappa_t^*(\alpha)$  is the threshold that best balances the trade-off between Type 1 and Type 2 errors, thus minimising misclassification risk in the predictions.

### 3.4.2.2 Single-Factor EVT<sub>M</sub> Portfolios

Logit models for different values of  $\alpha$  may entail different predictive power. This variation is influenced by the underlying time-varying nature of returns and probability predictors, as well as the size of the set of possible EVT-VaR violations. The predictors of the probability of  $EVT_t^\alpha$  violations in the returns of EVT<sub>M</sub> portfolios are the realised squared variances and semi-variances of the Fama-French factor. Realised variance is valuable because it provides an accurate measure of volatility, making it useful for volatility forecasting, while semi-variances capture the potential downside risk of an investment portfolio. Consequently, these predictors enhance the model's ability to predict returns.

To understand the distribution and the size of potential EVT violations, we assess how different values of  $\alpha$  affect this predictability. We focus on the left tail of the return distribution by constructing single-factor EVT<sub>M</sub> portfolios for a set of discrete values of  $\alpha$ , where we define  $\tilde{\alpha} = \{0.01, 0.05, 0.10, 0.25\}$ . As explained below, we choose  $\alpha$  at each date based on an optimised criterion that balances risk and return to construct the EVT<sub>M</sub> portfolio.

To estimate  $EVT_t^\alpha$ , we set  $EVT_t^\alpha$  equal to the empirical quantile of the original factor,

estimated over a data window prior to the forecasting date. A Logit model is then employed to predict  $\widehat{P}_{t+1}(\alpha)$ . Here,  $RV_t$  is the realised volatility, and  $s_t$  is the Relative Signed Jump Variation (RSJV) measure introduced by Bollerslev & Li & Zhao (2020), where  $RV_t^+$  and  $RV_t^-$  are the positive and negative realised semi-variances, respectively. Therefore,

$$s_t = \frac{RV_t^+ - RV_t^-}{RV_t}. \quad (3.13)$$

$RV_t^+$  and  $RV_t^-$  are interpreted as capturing ‘good’ and ‘bad’ volatility by focusing on positive and negative semi-variances, as well as embedded jumps. The Relative Signed Jump Variation (RSJV), is extensively theorised from Bollerslev & Li & Zhao (2020). They assume,

$$S_t = \int_0^t \mu_\tau d\tau + \int_0^t \sigma_\tau dB_\tau + J_t. \quad (3.14)$$

Where  $S_t$  denotes the natural logarithmic price of an arbitrary asset on day  $t$ ,  $\mu$  and  $\sigma$  denotes the drift and volatility process, respectively,  $B_\tau$  is the Brownian motion and  $J_t$  is the pure jump process. Assuming prices over trading day  $[t, t + 1]$ , the return on day  $t + 1$  is given by  $R_{t+1} = S_{t+1} - S_t$ . Thus, the daily Realised Variance (RV) is the sum of squared returns:

$$RV_t \rightarrow \int_{t-1}^t \sigma_\tau^2 ds + \sum_{t-1 \leq \tau \leq t} J_\tau^2, \quad (3.15)$$

with the positive and negative realised semi-variance measures defined as:

$$RV_t^+ = \sum_{t=1}^n R_t^2 1\{R_t > 0\}, \quad (3.16)$$

$$RV_t^- = \sum_{t=1}^n R_t^2 1\{R_t < 0\}, \quad (3.17)$$

where 1 is the indicator function taking the value 1 if the argument is true. With  $RV_t$  simplified to:

$$RV_t = \sum_{t=1}^n R_t^2, \quad (3.18)$$

the positive and negative realised semi-variances sum up to the total daily realised variance:

$$RV_t = RV_t^+ + RV_t^-. \quad (3.19)$$

We will refer the Signed Jump (SJ) variation to the positive minus negative volatility measure,

$RV_t^+ - RV_t^-$ :

$$SJ_t = RV_t^+ - RV_t^- \rightarrow \sum_{t-1 \leq \tau \leq t} J_\tau^2 1_{(J_\tau > 0)} - J_\tau^2 1_{(J_\tau < 0)}, \quad (3.20)$$

We normalise the Signed Jump (SJ) variation by the total realised variance, defining the Relative Signed Jump Variation (RSJV) as:

$$RSJV_t = \frac{SJ_t}{RV_t} = s_t, \quad (3.21)$$

This normalisation effectively removes the overall volatility level from the SJ measure, constraining the RSJV to lie between -1 and 1, making it easier to interpret as a measure of relative volatility asymmetry. We then define an indicator function for the probability of

EVT-VaR violations as:

$$I_{t+1}^\alpha = \begin{cases} 1, & \text{if } f_{t+1} < -\widehat{EVT}_t^\alpha, \\ 0, & \text{otherwise,} \end{cases} \quad (3.22)$$

where we compute the Logit regression using the binary endogenous response variable  $I_{t+1}^\alpha$  at time  $t + 1$  for a specified  $\alpha$ , and  $RV_t$  and  $s_t$  are exogenous variables. The Logit model to predict  $\widehat{P}_{t+1}(\alpha)$  takes the form:

$$\mathbb{P}(I_{t+1}^\alpha | X_t) = \Lambda(g(X_{t+1}\beta, p)) = \Lambda(\beta_{0\alpha} + \beta_{1\alpha}RV_t + \beta_{2\alpha}s_t), \quad (3.23)$$

where  $\Lambda(\cdot)$  represents the logistic function. Therefore, the probability forecast of the model at each time  $t$  is:

$$\widehat{P}_{t+1}(p) \equiv \mathbb{E}_t \Lambda(g(X_{t+1}\widehat{\beta}, p)) \quad (3.24)$$

The return of a single-factor EVTm portfolio for each  $\alpha \in \widetilde{\alpha}$  is given by:

$$f_{t+1}^{\zeta^*(\alpha)} = \gamma_t(\alpha) \frac{\kappa^*(\alpha)}{\widehat{P}_{t+1}(\alpha)} f_{t+1} \text{ for } \alpha \in \widetilde{\alpha}, \quad (3.25)$$

where  $\gamma_t(\alpha)$  is set to ensure the variance of  $f_{t+1}^{\zeta^*(\alpha)}$  is equal to the variance of the Fama-French factor computed on an expanded window starting from  $t_0$ .  $\gamma_t(\alpha)$  is the exposure to the strategy and is scaled down if  $\widehat{P}_{t+1}(\alpha) > \kappa_t(\alpha)$  and scaled up if the reverse is true. The scaling factor corresponding to the  $\alpha \in \widetilde{\alpha}$  that delivers the maximum Sharpe ratio is then chosen for the single-factor EVTm portfolio, calculated over a moving data window  $[t-w_2, t]$ .

At each date, the optimal  $\alpha$  is selected by solving the following optimisation problem:

$$p^* = \arg \max_{\alpha \in \tilde{\alpha}} \frac{w_2^{-1} \sum_{i=t-w_2}^t f_i^{\zeta^*(\alpha)}}{\sigma(f^{\zeta^*(\alpha)})}. \quad (3.26)$$

The return of the optimised single-factor EVTm portfolio is then computed using the following equation:

$$f_{t+1}^{\zeta^*} = \zeta_t^* f_{t+1}, \quad \text{where } \zeta_t^* \equiv \gamma_t(\alpha^*) \frac{\kappa_t^*(\alpha^*)}{\widehat{P}_{t+1}(\alpha^*)}. \quad (3.27)$$

This formulation dynamically adjusts the portfolio exposure based on the predicted risk, optimising returns while managing tail risk.

### 3.4.2.3 Multi-Factor EVTm Portfolios

Where we have a set of  $M > 1$  factors, the mean-variance efficient multi-factor portfolios includes the original factors and their EVTm factors. By incorporating both the original and EVTm-computed factors, we aim to capture diversification benefits across these two types of portfolios. The return of a multi-factor EVTm portfolio is given by:

$$f_{M,t+1}^{\zeta^*} = \sum_{m=1}^M \widehat{\Omega}_{j,t}^* f_{m,t+1} + \sum \bar{\Omega}_{j,t}^* f_{m,t+1}^{\zeta^*}, \quad (3.28)$$

where  $\widehat{\Omega}_{m,t}^*$  are the optimal weights assigned to each original factor, and  $\bar{\Omega}_{m,t}^*$  are the weights assigned to each single-factor EVTm portfolio. In Equation 3.29 below, the vector  $\Omega_t$  of optimal weights of the multi-factor portfolio is computed on an expanding data window,

starting with 1,000 trading days at  $t_0$ , by solving the following optimisation problem:

$$\max_{\Omega_t} = \mu_t^T \Omega_t - \frac{\delta}{2} \Omega_t^T \widehat{\Sigma}_t \Omega_t. \quad (3.29)$$

Where  $\mu_t$  as the sample conditional mean vector of factor returns,  $\widehat{\Sigma}_t$  is the covariance matrix of the factors. The parameter  $\delta$  is used to penalise portfolio variance, and is set to  $\delta = 2$ , ensuring that the portfolio is balanced between maximising expected returns and minimising risk.

This formulation enables the construction of a multi-factor EVTm portfolio that optimally combines both traditional and extreme value tail-risk-adjusted factors, improving risk-adjusted returns by accounting for extreme market events.

#### 3.4.2.4 Summary Steps to Compute EVTm portfolios

Let  $t_0$  and  $t$  denote the starting date and the forecasting date respectively. The construction of baseline EVTm single-factor portfolios is implemented with daily data according to the following steps:

1. **Data Window and Empirical Quantiles:** Using a moving data window  $[t - w_1, t]$ , where  $w_1 = 1,000$  trading days, we compute empirical quantiles  $EVT_t^\alpha$  for all  $\alpha \in \tilde{\alpha}$ . We choose  $w_1 = 1,000$  trading days. The choice of  $w_1 = 1000$  trading days ensures that the data window captures significant extreme events spanning approximately four years. This allows us to compute the estimated EVT-VaR which we express as  $\widehat{EVT}_t^\alpha$ .

2. **Volatility and RSJV Computation:** We calculate the realised volatility  $RV_t$  and the RSJV measure  $s_t$  from Equation 3.21 using data of the 22 days preceding and including the forecasting date.

3. **Logit Regression:** We estimate Logit regressions, as specified in Equation 3.23 with  $s_t$  and  $RV_t$  as input variables over the moving data window  $[t - w_1, t]$ , obtaining  $\widehat{P}_{t+1}(\alpha)$ ,  $\kappa_t^*(\alpha)$ , and the single-factor EVTm portfolio:

$$f_{t+1}^\zeta(\alpha) = \gamma_t(\alpha) \frac{\kappa_t^*(\alpha)}{\widehat{P}_{t+1}(\alpha)} f_{t+1}, \quad (3.30)$$

for each  $\alpha \in \tilde{\alpha}$ . Here  $\gamma_t(\alpha)$  is chosen to ensure the variance of  $f_{t+1}^\zeta(\alpha)$  equals to the variance of the Fama-French factor ( $f_{t+1}(\alpha)$ ) computed on an expanded window starting from  $t_0$ . The scaling factor  $\gamma_t(\alpha)$  adjusts exposure to the strategy, increasing it when  $\widehat{P}_{t+1}(\alpha) < \kappa_t^*(\alpha)$  and dynamically scaled down if the opposite is true.

4. **Predicted Probabilities:** From  $I_{t+1}^\alpha$ , we obtain predicted probabilities  $\widehat{P}_{t+1}$ , along with beta coefficients  $\beta_{1\alpha}$  and  $\beta_{2\alpha}$ .

5. **Threshold Estimation:** We estimate  $\eta_t$ , and subsequently compute the threshold  $\kappa_t(\alpha)$  from Equation 3.10.

6. **Confusion Matrix and Misclassification Analysis:** Use the confusion matrix to assess misclassification, as shown in Table 3.2. The threshold  $\kappa^*(\alpha)$  is optimised by minimising the error sum.

7. **Scaling Factor Calculation:** We compute the scaling factor  $\frac{\kappa^*(\alpha)}{\widehat{P}_{t+1}(\alpha)}$  and initialise

$\gamma_t(\alpha)$  to 1 for  $\gamma_t(\alpha) \frac{\kappa_t^*(\alpha)}{\widehat{R}_{t+1}(\alpha)}$ .

8. **Volatility Adjustment:** We perform a volatility adjustment loop for each  $\alpha \in \tilde{\alpha}$ , adjusting  $\gamma_t(\alpha)$  to ensure the variance of  $f_{t+1}^\zeta(\alpha)$  equal to the variance of the Fama-French factor ( $f_{t+1}(\alpha)$ ) computed on an expanded window starting from  $t_0$ .

9. **Optimal  $\alpha^*$  Selection:** We select the scaling factor  $\gamma_t(\alpha)$  that corresponds to the value of  $\alpha^* \in \tilde{\alpha}$  yielding the maximum Sharpe ratio for the single-factor EVTm portfolio  $f_{t+1}^{\zeta^*(\alpha)}(\alpha)$  over the moving window  $[t-w_2, t]$ , where  $w_2 = 250$  trading days (approximately one year).  $\alpha$  is chosen based on the past out-of-sample performance of the single-factor EVTm portfolios over the preceding  $w_2$  days. We select the scaling factor  $\gamma_t(\alpha)$  that corresponds to the value of  $\alpha^* \in \tilde{\alpha}$  yielding the maximum Sharpe ratio for the single-factor EVTm portfolio over the moving window  $[t-w_2, t]$ .

10. **Final Portfolio Construction:** Once the optimal parameters are updated, we calculate the optimal single-factor portfolio as:

$$f_{t+1}^{\zeta^*} = \zeta_t^* f_{t+1}, \quad \text{where } \zeta_t^* \equiv \gamma_t(\alpha^*) \frac{\kappa_t^*(\alpha^*)}{\widehat{R}_{t+1}(\alpha^*)}. \quad (3.31)$$

11. **Performance Comparison:** We also compare portfolio performance using the Sortino ratio, which, unlike the Sharpe ratio, focuses on downside risk by dividing expected returns by the standard deviation of downside returns. The Sortino ratio lies between the Sharpe ratio and CVaR, providing a more targeted risk metric.

### 3.4.3 Volatility Managed Portfolios

VM portfolios, as introduced by Moreira & Muir (2017), represent a class of dynamic investment strategies designed to adjust a portfolio's exposure to market risk based on market volatility. The central concept is to increase exposure to risky assets during periods when market volatility is low, under the assumption that the risk-adjusted return is more favourable, and to scale back that exposure when volatility is high, thereby reducing downside risk. This adaptive approach seeks to enhance returns by volatility-timing, effectively exploiting the inverse relationship often observed between volatility and expected returns across various asset classes and risk factors.

#### 3.4.3.1 Single-Factor Volatility Managed Portfolios

For the benchmark portfolio, we construct the Volatility-Managed (VM) portfolios from Moreira & Muir (2017) by scaling an excess return by the inverse of its conditional variance, where using the  $RV_t$  of previous 22 days the strategy increases or decreases risk exposure to the portfolio according to variation of conditional variance. The VM portfolio is given by:

$$f_{t+1}^{\zeta} = \frac{c}{\widehat{\sigma}_t^2(f)} f_{t+1}, \quad (3.32)$$

where  $f_{t+1}$  is the buy-and-hold portfolio excess return,  $\widehat{\sigma}_t^2(f)$  is a proxy for the portfolio's conditional variance, and the constant  $c$  which controls the average exposure of the strategy.

### 3.4.3.2 Multi-Factor VM Portfolios

Moreira & Muir (2017) extends their analysis to a multi-factor framework. They construct a portfolio by combining multiple factors, selecting weights that ensure the portfolio is Mean-Variance Efficient (MVE) for the set of factors. They refer to this as the multi-factor MVE portfolio. Consequently, the MVE alpha ( $a$ ) becomes the appropriate measure of expansion on the mean-variance frontier. A positive MVE alpha indicates that the Volatility-Managed strategy enhances Sharpe ratios compared to the highest buy-and-hold Sharpe ratio obtainable by someone using multiple factors.

The MVE portfolio is constructed with  $F_{t+1}$  representing a vector of factor returns, and  $b$  be the static weights that maximise the in-sample Sharpe ratio. The MVE portfolio is then defined as  $f_{t+1}^{MVE,\zeta} = b'F_{t+1}$ . We further define the Volatility-Managed portfolio as,

$$f_{t+1}^{MVE,\zeta} = \frac{c}{\widehat{\sigma}_t^2(f_{t+1}^{MVE})} f_{t+1}^{MVE}, \quad (3.33)$$

where  $c$  is a constant that normalises the variance of the Volatility-Managed portfolio to match that of the MVE portfolio. This strategy adjusts the conditional beta on the MVE portfolio but maintains the relative weights across the individual factors, thereby focusing solely on the time-series aspect of volatility timing.

### 3.4.3.3 Summary Steps to Compute VM portfolios

1. For simplicity, we construct the portfolio by using a proxy for the conditional variance which is the previous month's realised variance:

$$\widehat{\sigma}_t^2(f) = RV_{t_{prev}}(f), \quad (3.34)$$

2. The previous month's realised variance is expressed as,

$$RV_{t_{prev}}(f) = \sum_{d=\frac{1}{22}}^1 (f_{t+d} - \frac{1}{22} \sum_{d=\frac{1}{22}}^1 f_{t+d})^2, \quad (3.35)$$

3. The formation of the VM portfolio as shown in Equation 3.32, is expressed as:

$$f_{t+1}^\zeta = \frac{c}{\widehat{\sigma}_t^2(f)} f_{t+1}, \quad (3.36)$$

4. We choose  $c$  so that the managed portfolio has the same unconditional standard deviation as the buy-and-hold portfolio,  $f_{t+1}$ .
5. We compute the optimal portfolio weight  $\Omega_t^*$  to be proportional to the risk-return trade-off:

$$\Omega_t^* \propto \sigma_t^2(f) E_t[f_{t+1}]. \quad (3.37)$$

### 3.4.4 Transaction Costs

To extend the analysis, it is crucial to evaluate whether the EVTm or VM portfolios survive trading costs. In essence, we follow approaches similar to De Nicolo (2023) and Moreira & Muir (2017). Their method analyses the potential impact of trading costs, this is achieved by computing the break-even level of these costs which is the level that equates the Sharpe ratio of single-factor and multi-factor managed portfolios to the Sharpe ratio of the unmanaged portfolios.

The methodology for determining the break-even level of trading costs is as follows. The return of the portfolio net of trading costs at time  $t$  is given by:

$$f_t^{\zeta^{*net}} = f_t^{\zeta^*} - TC_t |\Delta\Omega|, \quad (3.38)$$

where,  $f_t^{\zeta^{*net}}$  is the portfolio return after accounting for trading costs,  $|\Delta\Omega|$  is the rebalancing of the portfolio weights at each date, and  $TC_t$  is the trading cost per unit of rebalance. The Sharpe ratio of the original factor is  $E[f_t]/\sigma^f$ , while the Sharpe ratio of the EVTm portfolio net of trading costs is  $E[f_t^{\zeta^{*net}}]/\sigma^{\zeta^{*net}}$ . Rearranging Equation 3.38, the trading costs that equate these two Sharpe ratios at each date are therefore:

$$TC_t |\Delta\Omega| = f_t^{\zeta^*} - \frac{\sigma^{\zeta^*}}{\sigma^f} E[f_t] \quad (3.39)$$

The Equation 3.39 sets up the condition where the Sharpe ratio of the original portfolio

(before trading costs) is equal to the EVTM Sharpe ratio after accounting for trading costs. It shows how much trading cost is required to reduce the Sharpe ratio of the portfolio to the level of the original Sharpe ratio. These computations are applied to both single-factor and multi-factor EVTM portfolios in a similar manner, ensuring a consistent comparison of performance after accounting for trading costs.

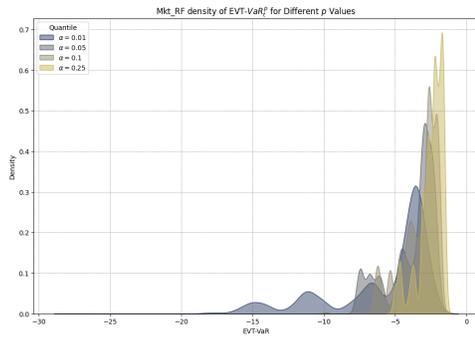
### 3.5 Empirical Analysis

The empirical analysis section encompasses a comprehensive investigation into several key aspects, including the selection of optimal thresholds, the application of Logit regression analysis, the evaluation of risk using Conditional Value-at-Risk (CVaR), and the assessment of portfolio performance through Sharpe ratio metrics. Additionally, it explores the estimation of portfolio alphas and examines break-even transaction costs, providing detailed insights into the financial efficiency and risk management strategies within the portfolio.

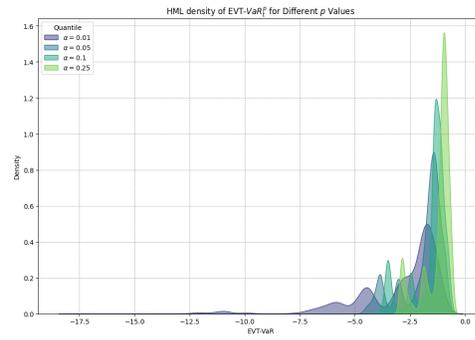
We compare the performance of ETM portfolios with buy-and-hold factors and VM portfolios. The first metric we assess is the Conditional Value-at-Risk, measured at 1% and 5% VaRs, denoted CVaR1 and CVaR5. The second set of metrics we assess are the Sharpe ratio and the ‘modified’ Sharpe ratio, where the ‘modified’ Sharpe ratio is the ratio of excess returns divided by its respective CVaR, as detailed in Xiong & Idzorek (2018). The relationship between CVaR and the Sharpe ratio lies in their shared focus on risk assessment, though they serve distinct purposes and offer complementary insights. CVaR focuses on downside risk and provides insight into potential losses beyond a certain threshold, quantifying the extreme tail risk of an investment. The Sharpe ratio evaluates risk-adjusted performance by considering both risk and return. In some cases, investors may use both measures together to gain a more comprehensive understanding of an investment’s risk and return profile. Investors with different risk appetites will prioritise one measure over the other. A risk-averse investor will pay more attention to CVaR to assess downside risk, while a risk-seeking investor will prioritise higher returns for a given level of risk by focusing on the Sharpe ratio.

The density of EVT-VaR estimations for  $\alpha \in \tilde{\alpha}$  is displayed in Figure 3.2. We can observe the density range of EVT estimations for every factor. This serves as an additional indicator that explains the distribution shape of the set of  $\tilde{\alpha}$ , and we notice similar patterns across all factors. As expected, a lower  $\alpha$ , such as 0.01, yields a fatter tail in comparison to higher orders of  $\tilde{\alpha}$ . We can observe that CMA has the highest densities for  $\alpha = \{0.25\}$ , which indicates that most of the data points are around -1.

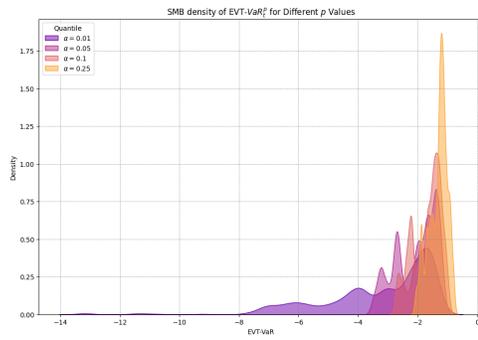
Figure 3.2:  $EVT^\alpha$  Density Plots of Fama-French Factors



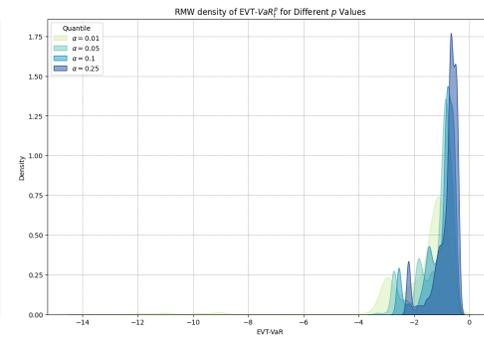
(a) Mkt-RF Density Plot



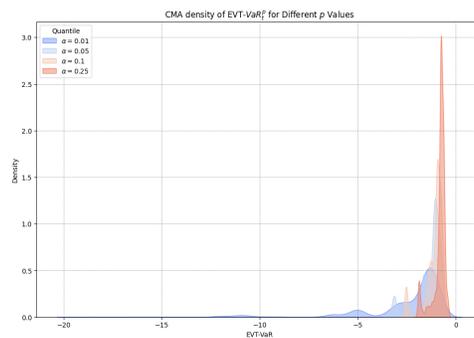
(b) HML Density Plot



(c) SMB Density Plot



(d) RMW Density Plot



(e) CMA Density Plot

This figure illustrates the densities of  $\hat{\alpha}$  for each factor.

Table 3.3: EVT M Optimal Threshold  $\kappa^*(\alpha)$  Estimates

$\alpha$	Mean	Std.dev	Min	Max	Mean	Std.dev	Min	Max
Mkt-RF								
1%	0.0845	0.0624	$3.6199 \times 10^{-5}$	0.3755	0.0829	0.0591	$6.424 \times 10^{-5}$	0.3766
5%	0.0981	0.0692	0.0002	0.4538	0.1016	0.0702	$2.6120 \times 10^{-5}$	0.4082
10%	0.1302	0.0813	0.0002	0.4696	0.1292	0.0810	$1.4167 \times 10^{-5}$	0.3943
25%	0.2470	0.0978	0.0121	0.5639	0.2515	0.0974	0.0097	0.5870
SMB								
1%	0.0851	0.0623	$7.9044 \times 10^{-5}$	0.3410	0.0836	0.0627	0.0002	0.3539
5%	0.1001	0.0693	0.0002	0.3347	0.1033	0.0691	0.0009	0.4062
10%	0.1286	0.0807	0.0010	0.4289	0.1328	0.0787	$1.7002 \times 10^{-7}$	0.4497
25%	0.2576	0.0986	0.0186	0.6198	0.2529	0.1001	0.0001	0.5787
CMA								
1%	0.0805	0.0590	0.0001	0.3325				
5%	0.1014	0.0707	0.0002	0.4354				
10%	0.1297	0.0780	$5.429 \times 10^{-5}$	0.4200				
25%	0.2511	0.0971	0.0001	0.5582				

In this Table, we produce statistics of the optimal thresholds  $\kappa^*(\alpha)$  where we minimise the sum of Type 1 and Type 2 errors in the out-of-sample 13,769 Logit regressions for the five factors Mkt-RF, SMB, HML, RMW, and CMA.

### 3.5.1 Single-Factor EVT M Portfolios

We observe in Table 3.3 that the means of  $\alpha = \{0.01, 0.05\}$  for all factors are closely matched. However, for  $\alpha = \{0.25\}$ , we see a larger average, with SMB having the highest mean of the optimal thresholds. There is variation across the standard deviation, minimum, and maximum statistics. Also, for each factor as  $\alpha$  increases, there is an increase in the standard deviation and maximum metrics.

We compute the out-of-sample estimates of the Logit models for each  $\alpha \in \tilde{\alpha}$ , obtained using data from a moving window of  $w_1 = 1000$  trading days. Table 3.4 reports statistics of the

estimated probabilities and statistics of the estimated coefficients of  $RV_t$  and the  $RSJV_t$  from Equations 3.18 and 3.21, respectively. The mean probability forecasts in Table 3.4 are a fraction of the nominal values of  $\alpha$ , but show similar results across each factor. The standard deviations of both estimated coefficients  $\beta_{1\alpha}$  and  $\beta_{2\alpha}$  decrease as  $\alpha$  increases, indicating lower sensitivity of the probability of EVT-VaR violations to changes in the predictors across all samples. The absolute mean of  $\beta_{1\alpha}$  for all factors decreases as the quantile  $\alpha$  increases, and this also applies to  $\beta_{2\alpha}$ , except for factors CMA and SMB, which contain both negative and positive mean values. In all Logit regressions, the coefficient  $\beta_{1\alpha}$  is negative, capturing a negative predictive relationship between  $RV_t$  and tail risk.

Table 3.4: EVTm Logit Estimates

$\alpha$	$P_{t+1}$ (mean)	$p_{t+1}$ (perc5)	$P_{t+1}$ (perc95)	$\beta_{1\alpha}$ (mean $RV_t$ )	$\beta_{1\alpha}$ (std.dev $RV_t$ )	$\beta_{2\alpha}$ (mean RSJV)	$\beta_{2\alpha}$ (std.dev RSJV)
Mkt-RF							
1	0.0010	0.0009	0.0010	-2.0978	0.7419	-0.0823	0.2879
5	0.0028	0.0019	0.0040	-1.6825	0.6592	-0.0620	0.2256
10	0.0050	0.0039	0.0060	-1.5095	0.60111	-0.0554	0.2036
25	0.0117	0.0090	0.0149	-1.2649	0.4977	-0.0458	0.1695
HML							
1	0.0010	0.0009	0.0010	-4.2394	1.6004	-0.0061	1.3650
5	0.0028	0.0019	0.0040	-3.3141	1.5522	-0.0492	1.0382
10	0.0050	0.0029	0.0070	-2.9914	1.3948	-0.0481	0.9288
25	0.0120	0.0090	0.0150	-2.4687	1.1022	-0.0376	0.7506
CMA							
1	0.0010	0.0009	0.0010	-4.7916	2.0703	0.0647	1.5630
5	0.0027	0.0009	0.0040	-4.3292	1.6834	-0.0744	1.4058
10	0.0050	0.0020	0.0070	-3.8944	1.5250	-0.0814	1.2600
25	0.0121	0.0090	0.0150	-3.2223	1.2520	-0.0701	1.0445
RMW							
1	0.0010	0.0009	0.0010	-5.4718	2.4322	-0.4297	1.9196
5	0.0028	0.0010	0.0040	-4.4430	1.9833	-0.2654	1.5414
10	0.0049	0.0029	0.0069	-4.0250	1.7971	-0.2346	1.3867
25	0.0115	0.0080	0.0149	-3.3589	1.4546	-0.1983	1.1521
SMB							
1	0.0010	0.0009	0.0010	-3.3382	1.2768	0.0413	0.7242
5	0.0026	0.0009	0.0040	-2.8923	1.0536	-0.0031	0.6655
10	0.0049	0.0020	0.0079	-2.5835	0.9611	0.0013	0.6056
25	0.0122	0.0099	0.0150	-2.1197	0.7602	-0.0038	0.4893

In this Table, we produce statistics of predicted probabilities of the Logit estimates  $P_{t+1}(\alpha) = E_t \Lambda(g(X_t \hat{\beta}, \alpha))$  together with parameters  $\beta_{1\alpha}$  and  $\beta_{2\alpha}$ , carried out over the 13,769 data points, with windows of length 1,000, for the Fama-French factors Mkt-RF, HML, CMA, RMW, and SMB.

### 3.5.1.1 Sharpe Ratio and Sortino Ratio Optimised Portfolios

Table 3.5 shows the first set of metrics for the original unmanaged factors, VM, and the EVTm portfolios. We optimise both EVTm portfolios using the Sharpe ratio and Sortino ratio. We compute all Sharpe ratio metrics (Sharpe1, Sharpe5, and Sharpe) for the VM and EVTm portfolios, and these metrics are reported as ratios relative to the annualised Sharpe ratio of their respective unmanaged factors. We also compute Sortino-optimised portfolios for EVTm. The results suggest that for all single-factor EVTm portfolios, both CVaR levels are higher than those obtained with the VM portfolios and unmanaged factors. Therefore, we find that the CVaR at the 99% and 95% confidence levels for both VM and unmanaged portfolios, on average, indicate that when the portfolio's return is worse than the 99th and 95th percentiles, the incurred losses will be greater than those of the EVTm portfolios. For example, in the VM portfolio Mkt-RF factor case, when the portfolio's return is worse than the 99th percentile CVaR, the Mkt-RF VM portfolio will incur losses that are worse than the Mkt-RF EVTm portfolios by 3.85% (4.65%-0.8%) or more. We also see that all EVTm Sharpe ratios are higher than those of the VM portfolios and unmanaged factors. Specifically, for the Sortino-optimised EVTm portfolio, the Sharpe ratio is 3.460 times higher than that of the VM portfolio. We obtain varied results when comparing EVTm portfolios optimised by the Sharpe ratio to Sortino-optimised EVTm portfolios. CVaR levels are either lower or the same for the Sortino-optimised EVTm portfolios compared to Sharpe-optimised EVTm portfolios.

For all EVTm portfolios, there is positive skewness; however, this is not the case for unmanaged

and VM portfolios. The kurtosis measures whether our values are heavy-tailed or light-tailed relative to a normal distribution. High kurtosis implies that the data tends to have heavy tails. The VM HML has the highest kurtosis, and regarding EVTm portfolios, the Sharpe-optimised EVTm HML and Sortino-optimised EVTm RMW factors are shown to have the highest kurtosis.

Table 3.5: Unmanaged, VM, Sharpe Optimised EVTm, and Sortino Optimised EVTm Portfolio Performance Measures for Factors Mkt-RF, SMB, CMA, HML, and RMW

		Mkt-RF				SMB			
Metrics	Factor	VM	EVTm (Sharpe)	EVTm (Sortino)	Factor	VM	EVTm (Sharpe)	EVTm (Sortino)	
CVaR1	-4.1317	-4.6516	-0.8251	-0.8251	-2.0121	-2.4162	-1.4093	-0.3205	
CVaR5	-2.4139	-2.5217	-1.6902	-1.5349	-1.2414	-1.3574	-0.9703	-0.8660	
Sharpe1	0.1026	0.2261	2.4462	2.4417	0.0613	0.1541	0.8009	3.6736	
Sharpe5	0.1756	0.4171	1.1941	1.3125	0.0994	0.2744	1.1632	1.3597	
Sharpe	0.4137	1.0262	2.1360	2.1285	0.2282	0.6883	2.2885	2.3816	
Mean	0.0267	0.0274	0.0526	0.0525	0.0077	0.0054	0.0162	0.0169	
St.Dev	1.0248	1.0248	0.9449	0.9465	0.5410	0.5411	0.4932	0.4944	
Skewness	-0.5291	-0.7446	0.3531	0.3465	-0.7048	-0.6537	0.0550	0.1275	
Kurtosis	15.6033	18.4121	19.4638	19.4206	18.2824	11.2903	18.3261	19.1964	

		CMA				HML			
Metrics	Factor	VM	EVTm (Sharpe)	EVTm (Sortino)	Factor	VM	EVTm (Sharpe)	EVTm (Sortino)	
CVaR1	-1.4782	-1.3941	-0.6751	-0.5268	-2.3745	-2.2471	-1.2164	-1.2164	
CVaR5	-0.8429	-0.8552	-0.4373	-0.3907	-1.3370	-1.3023	-0.6486	-0.6380	
Sharpe1	0.1466	0.2492	0.9514	1.1649	0.1886	0.2732	1.0490	0.9985	
Sharpe5	0.2572	0.4062	1.4687	1.5707	0.2934	0.4714	1.9674	1.9038	
Sharpe	0.5742	0.9201	1.9234	1.8699	0.4116	1.0577	2.2915	2.1973	
Mean	0.0136	0.0126	0.0232	0.0222	0.0151	0.0159	0.0331	0.0315	
St.Dev	0.3775	0.3776	0.3339	0.3282	0.5805	0.5805	0.5569	0.5528	
Skewness	-0.3426	0.7541	0.3861	0.3948	0.2977	1.6246	1.3192	1.0002	
Kurtosis	9.4688	14.1758	10.5351	10.9190	11.2649	41.9132	21.0673	18.6020	

		RMW			
Metrics	Factor	VM	EVTm (Sharpe)	EVTm (Sortino)	
CVaR1	-1.550	-1.4431	-0.3861	-0.3861	
CVaR5	-0.9086	-0.8796	-0.4156	-0.4045	
Sharpe1	0.1394	0.3562	1.7238	1.7032	
Sharpe5	0.2379	0.5844	1.6012	1.6258	
Sharpe	0.5430	1.2913	1.7958	1.7756	
Mean	0.0136	0.0176	0.0228	0.0225	
St.Dev	0.3980	0.3981	0.3706	0.3703	
Skewness	0.3427	0.3770	1.1549	1.1494	
Kurtosis	9.3067	14.6324	20.7605	20.8496	

In this Table, we report statistics of Sharpe and Sortino optimised EVTm, VM, and unmanaged portfolios, Conditional Value-at-Risk measures at 1% and 5% VaRs (CVaR1 and CVaR5), adapted Sharpe metrics (Sharpe1 and Sharpe5) given by the ratio of average returns to CVaR1 and CVaR5, the standard Sharpe ratio, and Moments (Mean, Standard Deviation, Skewness and Kurtosis).

### 3.5.1.2 Spanning Regressions

In Table 3.6, we conduct a time-series spanning regression analysis of the single-factor EVTMM and VM portfolios against the unmanaged buy-and-hold factors, employing the following regression model:

$$f_t^S = a + bf_t + \varepsilon_t. \quad (3.40)$$

A positive and significant estimated alpha( $a$ ) indicates that the ex-post Sharpe ratio of an investment in the original factor has improved through the inclusion of the EVTMM or VM factor in a portfolio. Specifically, a positive intercept suggests that volatility timing enhances Sharpe ratios relative to the original factors, further implying that the EVTMM or VM strategy effectively expands the mean-variance frontier. Our approach is grounded in the extensive empirical asset pricing literature, ensuring that these factors are accurately identified.

The single-factor alphas hold significant economic meaning, particularly when the individual factors precisely describe the opportunity set available to investors, or when these factors exhibit low correlations with one another—indicating that each factor captures a distinct dimension of risk. Moreover, the single-factor results are crucial in demonstrating that the empirical patterns we document are pervasive across different factors, reinforcing the notion that our findings are primarily driven by the time-series relationship between risk and return.

Table 3.6 presents the detailed results of the time-series spanning regressions. Our analysis is conducted on a factor-by-factor basis, with the alphas from the regressions being annualised. The findings reveal that the alphas associated with the real-time single-factor EVTMM portfolios

are universally positive and significant across all factors, with the Sharpe-optimised EVTm portfolios exhibiting the highest alphas, except in the case of the SMB factor, where the Sortino-optimised EVTm portfolio achieves the highest alpha. In contrast, the alphas from the spanning regressions using real-time VM portfolios are generally positive but not significant, with the notable exception of the RMW factor. Our VM alpha results are similar to De Nicolo (2023) Table 6 results where most VM factor alphas are not significant except for the MOM factor. Angelidis & Tessaromatis (2023) also show insignificant alphas from spanning regressions for the MKT, SMB, and CMA factors. They state the differences in the reported alphas for the VM market portfolio could be due to; a shorter period we use in the estimation of alpha (from 1967 to 2020), and/or the risk model used to adjust for risk. In our case, we our results are from 1968 to 2023, which differs from Moreira & Muir (2017) which is from 1926 to 2015. Despite the generally strong positive alphas, the largest is observed in the Sharpe-optimised EVTm Mkt-RF factor.

Single-factor EVTm portfolios have significantly higher  $R^2$  values compared to single-factor VM portfolios, implying that the returns of the single-factor EVTm portfolios are better explained by the factors considered in the regression model. Consequently, the EVTm portfolios consistently achieve significantly higher Sharpe ratios compared to the VM portfolios for all the factors analysed, underscoring the superior performance of the EVTm strategy under this metric.

Table 3.6: Single-Factor Portfolio Alphas

VM					
	Mkt-RF	SMB	CMA	HML	RMW
$b$	0.66*** (0.006)	0.7167*** (0.006)	0.667*** (0.006)	0.566*** (0.007)	0.593*** (0.007)
$a$	2.469 (1.512)	-0.076 (0.756)	0.8568 (0.504)	1.865 (1.008)	2.394*** (0.756)
$R^2$	0.44	0.52	0.45	0.32	0.35
EVTM (Sharpe)					
	Mkt-RF	SMB	CMA	HML	RMW
$b$	0.782*** (0.004)	0.777*** (0.004)	0.753*** (0.004)	0.810*** (0.004)	0.791*** (0.004)
$a$	8.110*** (1.008)	3.352*** (0.504)	2.923*** (0.252)	5.216*** (1.008)	2.848*** (0.504)
$R^2$	0.77	0.77	0.75	0.76	0.76
EVTM (Sortino)					
	Mkt-RF	SMB	CMA	HML	RMW
$b$	0.782*** (0.004)	0.778*** (0.004)	0.738*** (0.004)	0.802*** (0.004)	0.790 *** (0.004)
$a$	8.064*** (1.008)	3.528*** (0.504)	2.722*** (0.252)	4.838*** (0.504)	2.772*** (0.504)
$R^2$	0.76	0.77	0.74	0.76	0.76

We run time-series regressions of each managed portfolio factor on the unmanaged factor, using  $f_t^\zeta = a + bf_t + \varepsilon_t$ .  $f_t^\zeta$  represents managed factor portfolios for VM and EVTm portfolios. Stars are only intended to flag levels of significance. If a p-value is  $\leq 0.05$  it is flagged with one star (\*), if a p-value is  $\leq 0.01$  it is flagged with two stars (\*\*), subsequently, if a p-value is  $\leq 0.001$  it is flagged with three stars (\*\*\*).

### 3.5.2 Multi-Factor EVTm Portfolios

Regarding the multi-factor case, Table 3.7 presents a comprehensive set of performance metrics for the MVE, VM, and EVTm portfolios. In this analysis, we optimise the multi-factor EVTm portfolios using both the Sharpe ratio and the Sortino ratio as performance measures. Specifically, we compute a range of Sharpe ratio metrics (including Sharpe1, Sharpe5, and the overall Sharpe) for the MVE, multi-factor VM, and EVTm portfolios. These metrics are expressed as ratios relative to the annualised Sharpe ratio of their respective unmanaged MVE portfolio, providing a clear comparison of performance enhancements. The multi-factor EVTm portfolios are further optimised using weights derived from Equation 3.29, with the optimisation process conducted on an expanding data window that begins with 1,000 trading days at  $t_0$ .

The empirical results indicate that for all multi-factor EVTm portfolios, both the 99th and 95th percentile CVaR levels are lower than those observed in the VM portfolios. However, the MVE portfolio achieves the lowest CVaR metrics overall, highlighting its effectiveness in minimising extreme losses. Specifically, the CVaR for the MVE portfolio suggests that, at the 99th and 95th percentiles, the portfolio could incur losses worse than 0.81% and 0.47%, respectively.

Interestingly, the analysis reveals that most of the Sharpe ratio metrics for the multi-factor VM portfolios are the highest, except for the overall Sharpe ratio, where the multi-factor EVTm portfolios outperform both the VM portfolios and the MVE factor portfolio.

Table 3.7: EVTm Sharpe Ratio and Sortino Ratio Optimised Performance Measures

Metrics	MVE	VM	EVTm (Sharpe)	EVTm (Sortino)
CVaR1	-0.8116	-0.2852	-0.5918	-0.6003
CVaR5	-0.4731	-0.2375	-0.3836	-0.3838
Sharpe1	0.2943	0.7467	0.5577	0.5335
Sharpe5	0.5049	0.8964	0.8603	0.8344
Sharpe	1.1777	1.0528	1.7924	1.7437
Mean	0.0150	0.0155	0.0245	0.0238
St.Dev	0.2028	0.2023	0.1841	0.1837
Skewness	0.04751	-0.4304	0.3693	0.2749
Kurtosis	8.0345	11.8894	4.3906	4.5342

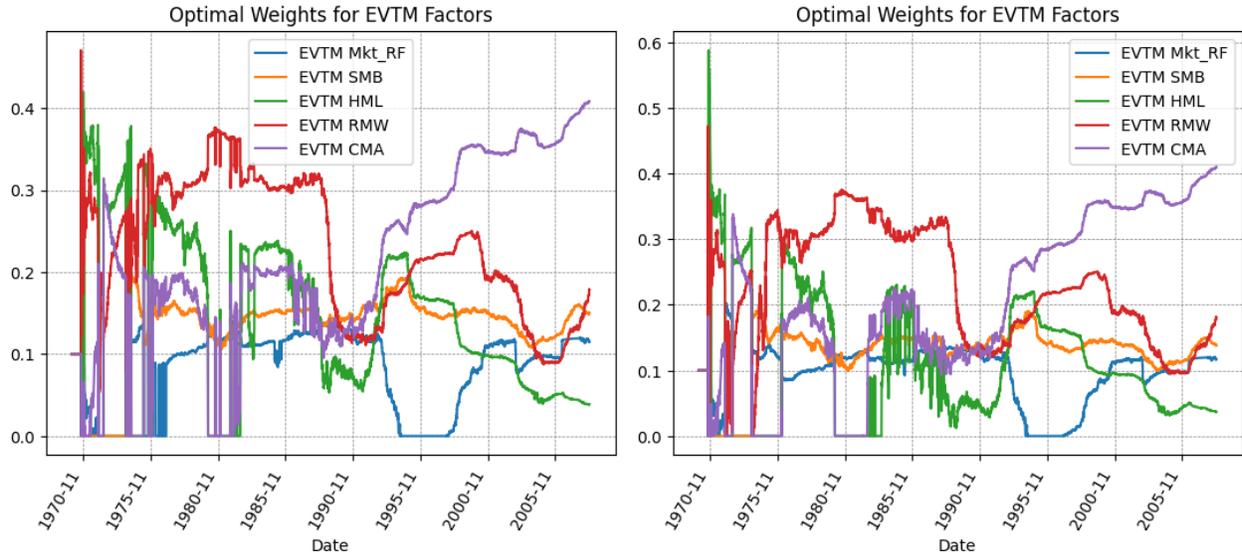
In this Table, we report statistics of Sharpe and Sortino optimised EVTm multi-factor portfolios, and VM and MVE portfolios, in relation to Sortino ratio, Conditional Value-at-Risk measures at 1% and 5% VaRs (CVaR1 and CVaR5), an adapted Sharpe metric given by the ratio of average returns to CVaR1 and CVaR5, as Sharpe1 and Sharpe5 respectively, the standard Sharpe ratio, and Moments (Mean, Standard Deviation, Skewness and Kurtosis).

Furthermore, the analysis of skewness and kurtosis shows a distinct pattern. The multi-factor EVTm and MVE portfolios display positive skewness, indicating a propensity for more frequent small gains and fewer large losses. In contrast, the multi-factor VM portfolios exhibit negative skewness, suggesting a higher likelihood of extreme negative returns. Additionally, the multi-factor VM portfolios followed by the MVE portfolios have the highest kurtosis, indicative of heavier tails and a greater likelihood of extreme outcomes.

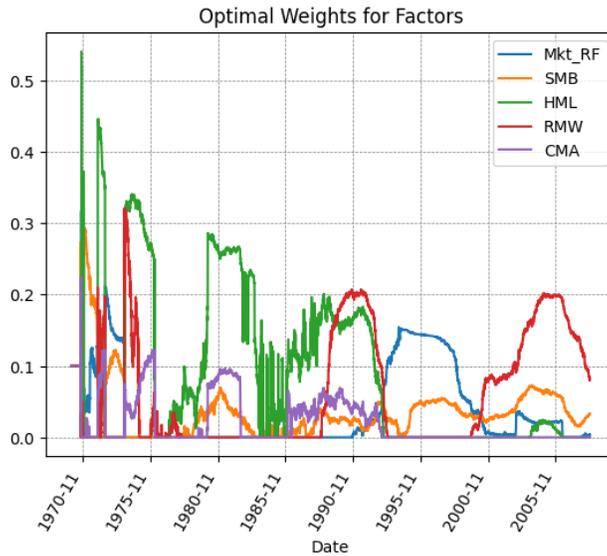
Figures 3.3 depict the time series of weights used for the original factor, as well as the Sortino-optimised and Sharpe-optimised multi-factor EVTm portfolios, which are optimised using weights generated from Equation 3.29. These plots provide a visual representation of how the allocation of weights changes over time for each optimisation strategy, offering insights into the dynamic adjustments made to enhance portfolio performance.

Figure 3.3: Plots of Weights of MVE, Sortino and Sharpe Optimised EVTM Portfolios

(a) Optimal Weights for Sharpe Optimised Multi-Factor EVTM (b) Optimal Weights for Sortino Optimised Multi-Factor EVTM



(c) Optimal Weights for MVE portfolio



This plot illustrates the weights of the unmanaged MVE portfolio, Sortino optimised multi-factor EVTM portfolio and Sharpe optimised multi-factor EVTM portfolio, optimised with the Equation 3.29.

### 3.5.2.1 Spanning Regressions

In Table 3.8, we perform a time-series spanning regression analysis of the multi-factor EVTm and VM portfolios relative to the Mean-Variance Efficient (MVE) portfolio, utilising the regression model outlined in Equation 3.40. The empirical results are striking. The alphas associated with the real-time multi-factor EVTm portfolios are both positive and statistically significant, indicating that these portfolios deliver superior returns after adjusting for the risks associated with the MVE benchmark. Notably, the Sharpe-optimised multi-factor EVTm portfolio stands out, with its alpha values being more than three times higher than those of the multi-factor VM portfolio. This substantial outperformance highlights the efficacy of the EVTm strategy in enhancing portfolio returns.

Moreover, the alpha derived from the spanning regressions of the real-time multi-factor VM portfolio is also positive and significant, demonstrating that the VM approach adds value, albeit to a lesser extent than the EVTm strategy. Additionally, the VM portfolio reports the highest  $R^2$ , suggesting that the returns of the VM portfolio are well-explained by the factors considered in the regression model.

Overall, these findings emphasise that the multi-factor EVTm portfolios consistently achieve significantly higher Sharpe ratios compared to the multi-factor VM portfolio. This underscores the superior performance of the EVTm strategy in optimising risk-adjusted returns, reinforcing its value as a robust portfolio management approach.

Table 3.8: Multi-Factor Portfolio Alphas

	VM	EVTM (Sharpe)	EVTM (Sortino)
$b$	0.603***	0.292***	0.315***
	(0.007)	(0.007)	(0.007)
$a$	1.663***	5.065***	4.788***
	(0.252)	(0.252)	(0.252)
$R^2$	0.37	0.10	0.12

We run time-series regressions of each multi-factor managed portfolio factor on the MVE portfolio, using  $f_t^\zeta = a + bf_t + \varepsilon_t$ .  $f_t^\zeta$  represents multi-factor managed factors for VM and EVT M portfolios. Stars are only intended to flag levels of significance. If a p-value is  $\leq 0.05$  it is flagged with one star (\*), if a p-value is  $\leq 0.01$  it is flagged with two stars (\*\*), subsequently, if a p-value is  $\leq 0.001$  it is flagged with three stars (\*\*\*).

### 3.5.3 Break-Even Transaction Costs

The results presented in Table 3.9 provide a detailed analysis of the mean level of break-even trading costs across both single-factor and multi-factor portfolios. The analysis reveals that the break-even costs for the EVT M portfolios are consistently higher than those for the VM portfolios across all single-factor and multi-factor portfolio cases. Importantly, these break-even costs for the EVT M portfolios are all positive, indicating that the portfolios can absorb relatively higher trading costs before their performance is adversely affected.

In contrast, the VM portfolios show some instances of negative break-even basis points, particularly in the SMB and CMA factors, suggesting that these portfolios may struggle to maintain performance once trading costs are factored in. The estimates for break-even costs in single-factor EVT M portfolios range from 60 to 140 basis points, while for multi-factor EVT M portfolios, the range is slightly higher, between 130 and 160 basis points. These higher break-even points indicate that the EVT M portfolios, despite incurring larger trading

Table 3.9: Break-Even Transaction Costs for Single-Factor and Multi-Factor Portfolios

Single-Factor Portfolio Break-Even Basis Points			
	VM	EVTM (Sharpe)	EVTM (Sortino)
Mkt-RF	107	72	74
HML	7	104	102
SMB	-55	66	72
RMW	71	117	118
CMA	-11	133	121
Multi-Factor Portfolio Break-Even Basis Points			
	VM	EVTM (Sharpe)	EVTM (Sortino)
	32	138	157

In this table, we report means of break-even trading costs for rebalancing for single-factor and multi-factor VM and EVTm portfolios.

costs, have a greater capacity to generate returns that justify these costs.

Moreover, although the break-even points for the EVTm portfolios are substantial, the superior performance of these portfolios is evident when considering other metrics such as the Sharpe ratios and alphas. These metrics consistently demonstrate significant improvements in risk-adjusted returns for EVTm portfolios compared to their VM counterparts.

### 3.6 Conclusion

In conclusion, the comprehensive analysis of various portfolio optimisation strategies, specifically focusing on the application of the Sortino ratio, and Sharpe ratio, combined with Extreme Value Theory (EVT)-VaR, provides compelling evidence for the superior performance of portfolios. The findings consistently demonstrate that EVTm portfolios optimised by the Sharpe ratio exhibit higher Sharpe ratios across all factors compared to their VM and unmanaged counterparts. Notably, the Sortino-optimised EVTm portfolios deliver exceptional results, with the Sortino-optimised Sharpe ratio being 3.460 times higher than that of the VM portfolio, further underscoring the effectiveness of incorporating downside risk into portfolio management.

Moreover, the analysis of Conditional Value-at-Risk (CVaR) reveals that EVTm portfolios manage extreme downside risks more effectively than VM portfolios, particularly at the 99% and 95% confidence levels. For instance, in the case of the Mkt-RF factor, the VM portfolio is expected to incur losses 3.85% worse than those of EVTm portfolios in extreme tail risk events. The presence of positive skewness in all EVTm portfolios, contrasted with the negative skewness in VM portfolios, indicates a more favourable risk-return profile for the EVTm strategy. Additionally, the kurtosis analysis suggests that EVTm portfolios, optimised using the Sharpe and Sortino ratios, exhibit a more moderate distribution of returns compared to the heavier-tailed VM portfolios.

The time-series spanning regression analysis further corroborates these findings. The alphas

associated with real-time single-factor and multi-factor EVTMM portfolios are universally positive and significant, with the Sharpe-optimised EVTMM portfolios showing alphas often more than three times higher than those of the VM portfolios. This robust performance is also reflected in the higher  $R^2$  values of the single-factor EVTMM portfolios, indicating that their returns are appropriately-explained by the factors considered, further solidifying EVTMM as a superior portfolio management strategy.

The analysis of break-even transaction costs adds another layer of insight. Despite the higher break-even costs associated with EVTMM portfolios ranging from 60 to 140 basis points for single-factor portfolios, and 130 to 160 basis points for multi-factor portfolios, the performance metrics, including Sharpe ratios and alphas strongly support the superior risk-adjusted returns of EVTMM portfolios compared to VM portfolios. The positive break-even points across all EVTMM portfolios highlight their resilience to transaction costs.

EVTMM and VM managed portfolios are crucial because, as Prospect Theory asserts, with loss aversion, economic agents are more sensitive to losses than to gains. Economic agents tend to be risk-seeking when they are in a position of gains (e.g., in calmer periods), where they are more willing to take on additional risk in pursuit of higher returns. Conversely, during periods of potential losses (e.g., market turbulence), they become more risk-averse and shift towards safer assets to mitigate risk. By adopting EVTMM portfolios and dynamically adjusting portfolio allocations in response to changing market conditions, investors can optimise risk-adjusted returns and better mitigate the adverse effects of market volatility. This adaptive approach allows for more resilient portfolio performance across varying market environments, strategically balancing risk and reward.

Overall, this study emphasises that the integration of EVT, alongside the optimisation of portfolios using both the Sharpe and Sortino ratios, significantly enhances portfolio performance by effectively managing downside risks and optimising risk-adjusted returns. This Chapter contributes to the expanding literature on Volatility-Managed portfolios by presenting the first study to fully integrate a volatility timing strategy based on Extreme Value Theory (EVT). By incorporating EVT, the study enhances the accuracy of extreme risk estimation, offering a more robust framework for managing portfolio volatility, particularly in periods of market stress. These findings are particularly relevant for portfolio managers and market risk managers when estimating the risk of extreme market movements. The results provide a strong argument for adopting the EVT strategy in both single-factor and multi-factor portfolio contexts, offering a robust approach to achieving superior financial outcomes in the face of market volatility and extreme events.

## Chapter 4

### A ROUGH GARCH-TYPE LSTM REALISED VOLATILITY MODEL

## Abstract

This study introduces novel rough hybrid Long Short-Term Memory Models (rGARCH-LSTM, rEGARCH-LSTM, and rGE-LSTM) designed for forecasting realised volatility by integrating roughness, the LSTM model, and different Generalised Autoregressive Conditional Heteroscedasticity (GARCH)-type models. Using intraday SPX index data, we explore hybrid models that combine LSTM with GARCH-type models and time-series analysis methods. These hybrid models are compared against standard models such as GARCH-LSTM and EGARCH-LSTM (exponential GARCH). Our findings reveal that roughness, when combined with GARCH and/or EGARCH models within an LSTM framework, significantly improves predictive accuracy in forecasting realised volatility. The primary contribution of this study lies in the emphasis on rough hybrid LSTM models, which effectively learn sequential patterns and enhance prediction accuracy in stock market realised volatility. This study evaluates the performance of our predictive models under different market volatility conditions using Markov switching regression models on a macroeconomic events dataset. Realised volatility is decomposed into three states, high, medium, and low volatility states to assess the efficacy of our models. Thus, we provide empirical evidence supporting the phenomenon of volatility clustering, demonstrating that heightened volatility contributes to increased forecasting errors. While some variability is observed in which model yields the lowest MAE and RMSE results

across the different volatility states, rough-based LSTM models consistently emerge as the preferred option for next-day forecasts. They demonstrate a decrease in loss function values when the input variables include the Hurst exponent. This suggests that rough LSTM models maintain their predictive efficacy across different volatility states, in contrast to non-rough LSTM models.

## Abbreviation

**LSTM** Long Short-Term Memory

**GARCH** Generalised Autoregressive Conditional Heteroscedasticity

**EGARCH** Exponential Generalised Autoregressive Conditional Heteroscedasticity

**MAE** Mean Absolute Error

**MSE** Mean Squared Error

**HAMAE** Heteroscedasticity Adjusted Mean Absolute Error

**HAMSE** Heteroscedasticity Adjusted Mean Squared Error

**MAPE** Mean Absolute Percentage Error

**AIC** Akaike Information Criterion

**BIC** Bayesian Information Criterion

**ADF** Augmented Dickey–Fuller

**EWMA** Exponentially Weighted Moving Average

**DM** Diebold-Mariano Test

**WS** Wilcoxon Signed-Rank Test

**AR** Autoregressive model

## Nomenclature

$t$	Current time
$\tau$	Current 5-minute time
$H$	Hurst exponent
$(R/S)_n$	Rescaled range for all sub-series of length $n$
$r_t$	Return at time $t$
$P_\tau$	Close price of the 5-minute time series at time $\tau$
$R_t$	Daily logarithmic return at time $t$
$R_\tau$	The 5-minute logarithmic return at time $\tau$
$p$	Lag order of asymmetric shocks (innovations) in GARCH-type models
$q$	Lag order of lagged volatility in GARCH-type models
$RV_t$	Realised volatility
$c_t$	Cell state of the LSTM model at time $t$
$f_t$	Forget gate of the LSTM model at time $t$
$h_t$	Hidden state of the LSTM model at time $t$
$x_t$	Current input of the LSTM model at time $t$

$i_t$	Input gate of the LSTM model at time $t$
$\tilde{c}_t$	Candidate cell state of the LSTM model at time $t$
$o_t$	Output gate of the LSTM model at time $t$
$s$	Sigmoid function

## 4.1 Introduction

Volatility in financial markets is a crucial indicator of both economic stability and investor confidence. Understanding and forecasting market volatility has become increasingly important, especially in light of major macroeconomic events such as financial crises, policy changes, and geopolitical tensions. Popular models used to forecast realised volatility based on historical data include GARCH-type models, Stochastic Volatility (SV) models, ARFIMA models, HAR-RV models, regime-switching models, and machine learning approaches. The ability to accurately forecast volatility aids in risk management, portfolio optimisation, and policy formulation.

In this study, the Hurst exponent, GARCH-type models, and the LSTM model are explained, computed, and empirically analysed. These statistical, econometric, and machine learning models are chosen for their ability to capture long-term dependencies (Hurst exponent), short-term volatility dynamics, volatility clustering, leverage effects (GARCH-type models), and complex non-linear relationships with long-term memory (LSTM models). LSTM models are particularly effective for modelling time series with long-term dependencies and patterns that traditional models might not capture. By combining these models, a more powerful and nuanced approach to forecasting realised volatility can be achieved.

Kim & Won (2018) demonstrate how combining a neural network model with multiple econometric models, rather than relying on a single econometric model, significantly enhances prediction performance compared to the existing literature. Additionally, the empirical results

of Hu & Ni & Wen (2020) show that GARCH forecasts can serve as informative features to substantially increase predictive power, and that integrating LSTM models is an effective approach to constructing deep neural network structures that further improve prediction performance. Ke et al. (2023) develop a hybrid model for predicting commodity prices using the Hurst parameter and GARCH model.

The Hurst exponent, introduced by Hurst (1951), characterises the long-term memory or the phenomenon of long-range dependence in time-series data, particularly the autocorrelation of data over time. Originally developed to analyse the flood and drought patterns of the Nile River, the Hurst exponent, denoted by  $H$ , has several key applications in various fields, especially in finance, geophysics, and hydrology. The nature of a time series can be determined based on its value of  $H$ , as explained in the Methodology section of this chapter. The Hurst exponent can indicate whether the fractional Brownian motion process of a financial instrument or market index exhibits negative autocorrelation, positive autocorrelation, or a random walk, as shown by B. B. Mandelbrot & Van Ness (1968). This information is valuable for developing trading strategies and risk management policies.

The GARCH (Generalised Autoregressive Conditional Heteroscedasticity) and EGARCH (Exponential GARCH) models are widely used in financial econometrics for modelling time-varying volatility. The GARCH model effectively captures the volatility clustering phenomenon in financial markets, where high-volatility events tend to follow other high-volatility events, and the same is true for low-volatility events. GARCH models are also useful for forecasting future volatility. The EGARCH model, on the other hand, allows for asymmetric volatility modelling by explicitly capturing the leverage effect, where negative

shocks to asset returns might impact volatility differently compared to positive shocks of the same magnitude. This feature is particularly useful in financial markets that exhibit asymmetric responses to news.

Unlike the GARCH model, which requires certain parameters to be positive to ensure a positive conditional variance, the EGARCH model does not impose such constraints due to its logarithmic specification. This allows for greater flexibility in modelling and simplifies parameter estimation. EGARCH can capture complex dynamics in the volatility process, including long memory and persistent effects of shocks, making it suitable for a wide range of financial series. By accurately modelling the skewness and kurtosis of financial returns, EGARCH models improve the valuation of financial derivatives and enhance risk assessment practices by accounting for the asymmetric and leptokurtic nature of asset returns. Both GARCH and EGARCH models are integral tools in quantitative finance due to their ability to model and forecast volatility accurately. For this reason, we combine specifications of both GARCH-type models to leverage the advantages of each model in our hybrid model.

We now move on to the Long Short-Term Memory (LSTM) model, which is a type of Recurrent Neural Network (RNN) that avoids the vanishing gradient problem. At the heart of LSTM networks are memory cells that retain information for extended periods. Each cell has mechanisms called gates that regulate the flow of information into and out of the cell, helping it remember or forget information. LSTMs have three types of gates: input gates, output gates, and forget gates. These gates determine whether to allow new information in (input gate), delete stored information (forget gate), or let it influence the output at the current time-step (output gate).

Next, we develop the rough GARCH-type LSTM models, which combine GARCH-type attributes with the Hurst exponent. Specifically, we incorporate GARCH conditional volatility, EGARCH conditional volatility, GARCH residuals, EGARCH residuals, and the Hurst exponent factors of the stock index at each time  $t$ , and feed these factors into the LSTM model. The input consists of the aforementioned factors of the stock index. Ultimately, the implicit connections among the stock index factors are modelled to provide valuable information for forecasting realised volatility.

Furthermore, a comparative analysis of all forecasted outcomes from the GARCH-type LSTM models is conducted using the Wilcoxon Signed-Rank (WS) test and the Diebold-Mariano (DM) test. The results show high statistical significance, particularly with rough-based GARCH-type LSTM models. Our empirical findings suggest that incorporating a rolling Hurst exponent along with GARCH-type models, which effectively capture volatility clustering, asymmetry, and the leverage effect, optimally enhances the predictive capabilities of forecasting  $RV_t$ . Additionally, in the context of the macroeconomic events dataset, various loss functions are evaluated across three volatility states generated by Markov switching models, covering low, medium, and high volatility states. The analysis confirms that rough GARCH-type models outperform their non-rough counterparts. This comparative analysis also highlights an increase in forecasting errors associated with macroeconomic events, providing empirical evidence for the phenomenon of volatility clustering. The increase in errors suggests that high volatility significantly impacts and amplifies forecasting errors.

The chapter is structured as follows: the literature review in Section 4.2 provides an overview of relevant research on the Hurst exponent, GARCH-type models, and the LSTM model.

In Section 4.3, the data is presented, and details of the SPX dataset and the interest rate (the 3-Month Treasury Constant Maturity Rate, DGS3MO) are discussed. The methodology in Section 4.4 introduces the Hurst exponent and explains how roughness can be captured in its estimation. Afterward, GARCH and EGARCH components explained, extracted and analysed. The LSTM model is then constructed and illustrated. Finally, in Section 4.5, the empirical analysis focusses on the model configuration, and covers a substantial number of macroeconomic events where  $RV_t$  is further decomposed into three volatility states. We present statistical tests and loss function results that quantify the error margin between a model's predictions and the actual target realised volatility to determine the best model. The conclusion, discussed in Section 4.6, offers insights and potential stakeholders.

## 4.2 Literature Review

Forecasting financial asset prices presents a formidable challenge in time-series analysis, primarily due to the inherent noise and pronounced volatility characteristic of the financial markets. From a practical standpoint, the ability to accurately predict the volatility of financial time series holds substantial significance. It equips policymakers, risk managers, and speculators—whether they rely on fundamental or technical analysis—with a potent tool to realise financial gains and manage or mitigate risks.

Numerous scholarly articles have extensively investigated the phenomenon of persistence across diverse financial assets, encompassing stocks as explained by Lo (1991) and Los & Yalamova (2004), exchange rates as assessed by Da Silva et al. (2007), commodity prices explored by Serletis & Rosenberg (2007), and cryptocurrencies as analysed by Bariviera

(2017). Various methodologies have been employed to assess long-range persistence in assets. Caporale & L. Gil-Alana & Plastun (2018) deploy the R/S analysis, Caporale & L. A. Gil-Alana (2013) utilise fractional integration, Barunik & Kristoufek (2010) explore the generalised Hurst exponent approach, Grech & Mazur (2005) use the detrended moving average, and Kantelhardt et al. (2002) implement the multifractal generalisation. Recent studies have found that the log-volatility of asset returns exhibits roughness. Gatheral & Jaisson & Rosenbaum (2018) examine the smoothness of the volatility process, concluding that log-volatility behaves as a fractional Brownian motion (fBm) with a Hurst exponent  $H$  of approximately 0.1 at any reasonable timescale, indicating that volatility exhibits roughness.

B. Qian & Rasheed (2004) show that the Hurst exponent provides a measure of predictability. They use the Hurst exponent to classify sets of financial data representing different time periods. Their findings suggest that periods with larger Hurst exponents can be predicted more accurately than those with  $H$  values closer to those of random series, indicating that stock markets are not entirely random in all periods. Most studies have predominantly utilised daily data, as shown by Zunino et al. (2009), with a smaller number focusing on weekly data, as depicted by MacDonald & Taylor (1992), or monthly frequencies which are studied by Caporale & L. Gil-Alana & Plastun (2019). Additionally, there is a limited body of literature examining long-memory properties in high-frequency data. For instance, Andersen & Bollerslev (1997) explored persistence in 5-minute returns in the FX and US stock markets, uncovering long-memory properties in volatility. Cotter (2005) observed long-memory properties in the UK futures market using 5-minute interval data. Thus, the Hurst exponent is a versatile tool for analysing time-series data, providing insights into the

data's autocorrelation structure and aiding in modelling and forecasting future behaviour based on historical patterns.

Engle (1982) introduced a novel framework for forecasting volatility in financial markets, effectively establishing an approach to financial time-series analysis and leading to several developments in the field. The GARCH model was first introduced in Bollerslev (1987) seminal paper, which proposed a process to capture the impact of past shocks (squared residuals) on current volatility while accounting for the influence of past volatility. Bollerslev (1987) implemented parameter estimation using maximum likelihood methods, highlighting the need for formal tests to detect GARCH effects. Through empirical examples, he demonstrated the superior performance of GARCH over ARCH in modelling inflation, particularly with the GARCH (1,1) specification.

While ARCH/GARCH models effectively capture volatility clustering in-sample, Nelson (1992) identified weaknesses in the symmetric GARCH model, highlighting its inability to account for the negative correlation between asset returns and changes in return volatility. To address these limitations, Nelson developed the Exponential GARCH (EGARCH) model, which incorporates the sign and magnitude of shocks without requiring non-negativity constraints. This development paved the way for variations like Threshold GARCH (TGARCH) from Zakoian (1994), which allow for skewness and leptokurtosis in the distribution of residuals, which have more capability in capturing extreme events. In summary, the EGARCH model introduced by Nelson (1992) captures the asymmetric phenomenon known as the leverage effect, meaning it accounts for both positive and negative shocks (good and bad news) in relation to asset returns.

Engle & Patton (2007) emphasised key features of predictive models, including volatility clustering, mean reversion, asymmetry, and the impact of exogenous variables. They favoured asymmetric GARCH models for their ability to capture these characteristics, though they acknowledged challenges in incorporating exogenous variables. Vilasuso (2002) and Zhou & Kang (2011) explored alternative models to address GARCH's limitations, such as insufficient memory and inability to capture long-lasting shocks. Vilasuso (2002) recommended the Fractionally Integrated GARCH (FIGARCH) model, while Zhou & Kang (2011) compared long-memory and short-memory models, finding that long-memory models outperformed in forecasting real estate volatility.

Kışınbay (2010) assessed nonlinear GARCH models and highlighted the superiority of asymmetrical models, particularly EGARCH, in predicting equity volatility. Kışınbay (2010) also emphasised the importance of out-of-sample forecasting and provided a comprehensive evaluation of model performance. Instances where authors used GARCH models to forecast realised volatility include Kim & Won (2018), who chose a constant GARCH (1,1) model for their financial time-series, assuming its superiority over models with higher orders. C. Deng et al. (2020) also used GARCH (1,1) with high-frequency data to model daily realised volatility. Kambouroudis & McMillan & Tsakou (2016) adapted a model that integrates an asymmetric GARCH model with both implied and realised volatility, using (asymmetric) ARMA models. Koopman & Jungbacker & Hol (2005) compared the forecasting effectiveness of long-memory models for realised volatility with Stochastic Volatility (SV) models and GARCH models, both based on daily return series.

Corsi (2009) proposed a new realised volatility model designed to directly model and forecast

the time series behaviour of volatility. Their aim was to develop a conditional volatility model based on realised volatility that could accurately reproduce the long memory persistence observed in empirical data. Through simulations, Corsi (2009) demonstrated that the HAR-RV (Heterogeneous Autoregressive Model of Realised Volatility) successfully captures key empirical features such as long memory and fat tails.

When applied to USD/CHF data, the HAR-RV model showed remarkable out-of-sample forecasting performance, consistently and significantly outperforming competing models such as RiskMetrics, AR (Autoregressive), and ARFIMA (Autoregressive Fractionally Integrated Moving Average). Specifically, for next-day-ahead forecasts, the results were as follows: RiskMetrics (3.5945), AR(1) (2.9404), AR(3) (2.9088), ARFIMA(5, 0.401, 0) (2.8916), and HAR-RV (2.8472). These results highlight the superior predictive accuracy of the HAR-RV model compared to traditional approaches.

Poon & Granger (2003) found interesting results in 66 studies that investigated the comparative performance of forecasting techniques; HISVOL (historical volatility), ISD (option-implied standard deviation, based on the Black–Scholes model and its generalisations), GARCH, and SV (stochastic volatility). Table 4.1 presents pairwise comparisons drawn from these 66 studies, where some compared only one pair of forecasting techniques, while others compared several. Among those involving both HISVOL and GARCH models, 22 studies (56%) found HISVOL superior to GARCH, and 17 studies (44%) found GARCH superior to HISVOL. 17 studies (94%) found ISD to be superior to GARCH and 26 studies (76%) found ISD to be superior to HISVOL.

Table 4.1: Pairwise Comparisons of Forecasting Models

	Number of Studies	Studies Percentage
HISVOL > GARCH	22	56%
GARCH > HISVOL	17	44%
HISVOL > ISD	8	24%
ISD > HISVOL	26	76%
GARCH > ISD	1	6%
ISD > GARCH	17	94%
SV > HISVOL	3	100%
SV > GARCH	3	100%
GARCH > SV	1	100%
ISD > SV	1	100%

This Table from Poon & Granger (2003) shows comparative performance of HISVOL (historical volatility), ISD (option implied standard deviation, based on the Black-Scholes mode and various generalisations), GARCH, and SV (stochastic volatility model forecasts that historical volatility models).

The evolution of machine learning techniques from traditional models to more advanced variants has marked a significant advancement in this field. Contemporary machine learning models adeptly simulate complex real-world data by extracting robust features that capture relevant information, demonstrating superior predictive performance compared to traditional linear models, as evidenced in the studies by Salakhutdinov & Hinton (2009) and Bengio & Courville & Vincent (2013). Given the complexities inherent in financial asset prices, the integration of deep learning techniques with financial market forecasting represents one of the most promising areas of research, as indicated by Cavalcante et al. (2016). This convergence offers a sophisticated analytical framework capable of decoding the nuanced dynamics of financial markets, holding great potential for advancing our understanding and predictive capabilities in this domain.

Hybrid models that combine traditional GARCH models with machine learning features

have also been explored to improve predictive performance. There is substantial literature examining the integration of neural network models with traditional econometric models, particularly in the domain of financial market volatility forecasting. For instance, Roh (2007) explores the combination of a neural network with a financial time-series model using the KOSPI 200 index. Roh (2007) demonstrates that a model that amalgamates a feedforward neural network with traditional models such as EWMA, GARCH, and EGARCH outperforms a standalone GARCH model. Further studies by Tseng et al. (2008) integrate an EGARCH model with a feedforward neural network to estimate the volatility of Taiwan's stock index option prices, achieving superior performance over the EGARCH model alone by reducing stochasticity and nonlinearity in the error term.

Expanding upon this, Kristjanpoller & Fadic & Minutolo (2014) investigate a hybrid neural network GARCH model for forecasting volatility in three Latin American markets. Their findings reveal that the hybrid model significantly lowered the Mean Absolute Percentage Error (MAPE) compared to the conventional GARCH model. In a related study, Kristjanpoller & Minutolo (2016) employ a hybrid ANN-GARCH model to predict oil price volatility, incorporating additional variables such as indices and exchange rates related to oil prices. They identify an optimal architecture for different time windows of volatility, which demonstrates a 30.6% improvement in prediction accuracy over the GARCH model. Furthermore, Kristjanpoller & Hernández (2017) compared the predictive capabilities of GARCH-type models with hybrid neural network models, finding that the latter were more effective in forecasting the volatility of major metals, underscoring the advantages of integrating neural networks with traditional econometric approaches in financial volatility forecasting.

Vidal & Kristjanpoller (2020) show that deep learning can be applied in financial forecasting, improving the results of classic models. Predicting volatility for different types of financial assets is one of the most complex tasks in time-series prediction, mainly due to its noisy, non-stationary, and heteroscedastic structure. Vidal & Kristjanpoller (2020) show their model significantly improves the forecast of gold volatility by combining two deep learning methodologies: Long Short-Term Memory (LSTM) and Convolutional Neural Networks (CNN). The CNN-LSTM hybrid model, which uses convolutional and LSTM layers to learn from the training data, shows substantial improvement over standalone GARCH and LSTM models, with a 37% reduction in Mean Squared Error (MSE) compared to the classic GARCH model and an 18% reduction compared to the LSTM model. Additionally, Hajizadeh et al. (2012) develop two hybrid models that combine EGARCH with a feedforward network. The results indicate significant enhancements in model accuracy, evidenced by reduced prediction errors compared to using either a single econometric model or a single Artificial Neural Network (ANN) model.

Kim & Won (2018) propose a new hybrid Long Short-Term Memory (LSTM) model to forecast stock price volatility, combining the LSTM model with various Generalised Autoregressive Conditional Heteroscedasticity (GARCH)-type models. Using KOSPI 200 index data, they demonstrate that hybrid models that combine LSTM with one to three GARCH-type models significantly reduce prediction errors. The LSTM model with three GARCH-type models showed the lowest prediction errors in terms of Mean Absolute Error (MAE), Mean Squared Error (MSE), Heteroscedasticity-Adjusted MAE (HAMAE), and Heteroscedasticity-Adjusted MSE (HAMSE). The GEW-LSTM model from Kim & Won (2018) showed 57.3%, 24.7%,

and 48% smaller MSE, HAMAЕ, and HAMSE, respectively, compared to the standalone GARCH model.

A hybrid model for predicting commodity prices using the Hurst parameter and GARCH model is introduced by Ke et al. (2023). They propose the EEMD-Hurst-LSTM prediction method, which is designed to forecast trends in China's commodity futures market. This method combines Ensemble Empirical Mode Decomposition (EEMD) with an adaptive fractal Hurst index derived from intraday data, integrating these into an LSTM model.

This body of research collectively illustrates the potential of hybrid models to enhance predictive accuracy in financial markets, paving the way for more robust and reliable econometric forecasting tools.

### 4.3 Data Description

This section discusses and describes the data used within this chapter.

The dataset consists of intraday 5-minute open, high, low, and close (OHLC) prices for the S&P 500 Index, sourced from [firstratedata](#). The data range spans from January 1st, 2008, to December 31st, 2022. We select the close price  $P_t$  for each timestamp  $t$  as our primary price. This data is used to compute the log-returns and, subsequently, the realised volatility.

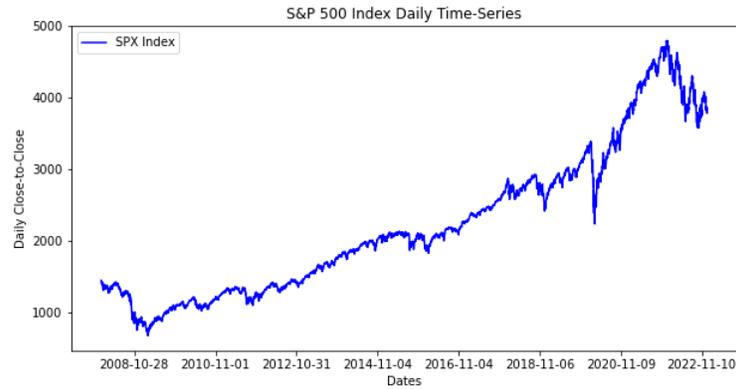
Table 4.2 presents descriptive statistics such as mean, standard deviation, skewness, and kurtosis for the log-return time-series data of the SPX index. Additionally, we conduct tests for stationarity and normality using the Augmented Dickey–Fuller (ADF) test and the Jarque-Bera test, respectively. To convert the indexes to a stationary time series, we transform the data by differencing the return value. This is done by subtracting the natural logarithm of the current price from the previous price in the time-series, resulting in  $R_\tau$ , where,

$$R_\tau = \ln \left( \frac{P_\tau}{P_{\tau-1}} \right), \quad (4.1)$$

and  $P_\tau$  is the price of the time-series at time  $\tau$ .

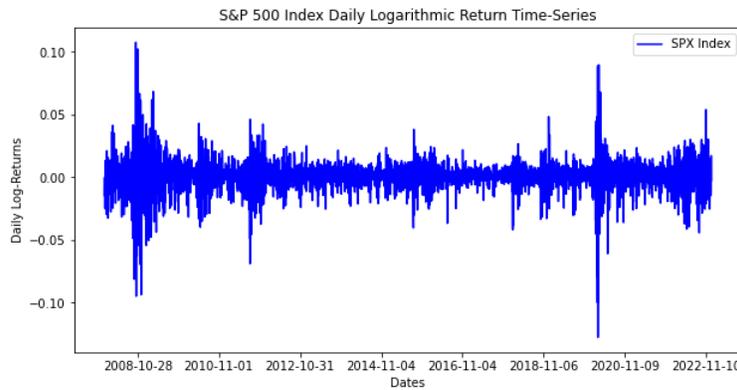
Figures 4.1, 4.2, and 4.3 show the daily close price, as well as the daily and intraday logarithmic returns for the S&P 500 index over the entire dataset, spanning from January 1st, 2008, to

Figure 4.1: S&P 500 index daily close price  $P_t$  time-series



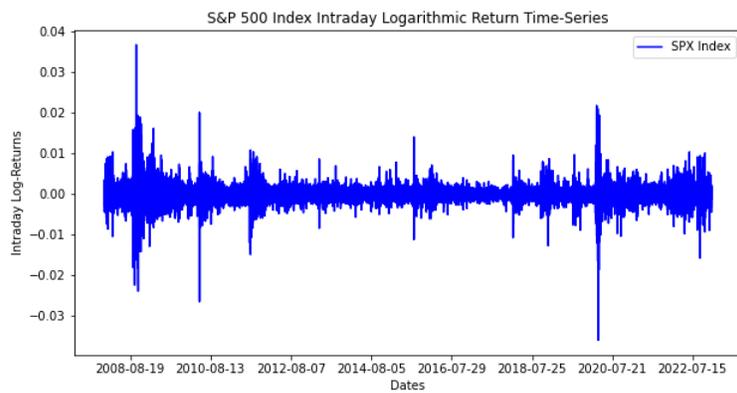
This chart describes the S&P 500 daily close-to-close time-series.

Figure 4.2: S&P 500 index daily logarithmic return  $R_t$  time-series



This chart describes the S&P 500 daily logarithmic returns of time-series.

Figure 4.3: S&P 500 index intraday logarithmic return  $R_\tau$  time-series



This chart describes the S&P 500 intraday logarithmic returns of time-series.

Table 4.2: Summary Statistics for Data

Index	Mean	Std	Skewness	Kurtosis	ADF	Jarque-Bera
SPX $R_t$ (daily)	0.000263	0.013185	-0.5015	11.6486	-10.90774***	21284***
SPX $RV_t$	0.000094	0.000248	9.8421	140.7872	-6.470815***	3146633***

Stars are only intended to flag levels of significance. If a p-value is  $\leq 0.05$  it is flagged with one star (\*), if a p-value is  $\leq 0.01$  it is flagged with two stars (\*\*), subsequently, if a p-value is  $\leq 0.001$  it is flagged with three stars (\*\*\*).

December 31st, 2022.

Table 4.2 shows that the distribution of daily SPX index logarithmic returns exhibits negative skewness. Additionally, the Jarque–Bera test rejects the normality of the return distribution. The ADF test also rejects stationarity with a value of -10.91, under the condition that the thresholds are -2.57, -2.86, and -3.43 at the 10%, 5%, and 1% significance levels, respectively. Thus, heterogeneity likely exists in the time series, as the distribution deviates from normality.

To compute the intraday returns, we divide each trading day into 78 successive 5-minute intervals, from the market’s opening at 9:30 a.m. to its close at 4:00 p.m. Since stock returns are computed using intraday data only, we exclude overnight returns from the series. Following Stoll & Whaley (1990) and Darrat & Rahman & Zhong (2003), we exclude the first two 5-minute returns, as they conclude that the average time to open for S&P 500 Index is around 5–7 minutes. Return prices during this period reflect the stale closing price of the previous day. Therefore, excluding the first two 5-minute return observations as well as those after 4:00 p.m. helps mitigate the effects of stale price information. We also exclude dates with late openings or early closings on holidays. In line with the methodology of Dobrev & Szerszen (2010) and Visser (2011), we exclude dates with early closing times (e.g., 12:15 p.m. on days before holidays) or late openings, resulting in the exclusion of 45 dates.

The GARCH and EGARCH residuals and conditional variances are computed using close-to-close values from the high-frequency data, transformed into daily returns. Specifically, we use today's last intraday close price divided by yesterday's last intraday close price. We focus on the period from January 1st, 2008 to January 1st, 2023. The rolling GARCH( $p, q$ )-type models are computed with a rolling window of 22 days to capture information over the course of a month, testing a range of values for  $p$  and  $q$ , where  $p$  is the lag order of asymmetric shocks (innovations) and  $q$  is the lag order of lagged volatility. The residuals and conditional variances extracted from this analysis are key inputs for the LSTM model as shown in Table 4.3.

Regarding macroeconomic data, we include the interest rate, specifically the 3-Month Treasury Constant Maturity Rate (DGS3MO), sourced from the Federal Reserve Economic Data (FRED).<sup>1</sup> The financial and political events used for analysis are a subset of the database of events identified by Bloom (2009), as well as those identified by Piffer & Podstawski (2018). The dataset includes observations from February 15th, 2008 to July 8th, 2015.

---

<sup>1</sup><https://fred.stlouisfed.org>

## 4.4 Methodology

The methodology describes the methods and integration of the Hurst exponent, GARCH-type models, and Long Short-Term Memory (LSTM) networks.

### 4.4.1 Fractional Brownian and Hurst Exponent

The most common methodology applied towards establishing randomness in stock and option pricing deals with the incorporation of Brownian motion as a source of randomness. The seminal article by Brown (1828) discovered this phenomenon in 1827 while looking through a microscope at pollen, and Black & Scholes (1973) popularised this method within finance. However, there was early criticism regarding the inadequacy of Brownian motion with regard to real market data; B. Mandelbrot (1963) challenges the assumption that price changes in financial markets follow a normal distribution which is a key characteristic of Brownian motion and observes that price changes exhibit fat tails and long memory, Fama (1965) argues the efficacy of the stochastic process of Brownian motion discusses the nature of real market distributions and how the distributions were shown to be non-Gaussian and possess leptokurtic properties, and Lo & MacKinlay (1988) observes that the processes of real market data exhibits serial correlation. In the Brownian motion framework serial correlation does not apply; however, the fractional Brownian motion (fBM) stochastic process can be used to analyse this market data behaviour.

The ways in which fBM has influenced stochastic models is in how its able to model long-range

dependence which is measured by the slow decay of auto-correlation functions of increments. Initially introduced by B. B. Mandelbrot & Van Ness (1968), the fBM is a gaussian stochastic process that is a generalisation of the Brownian motion with one additional parameter, the Hurst parameter which has a range between zero and one. Over the range of parameter values of the Hurst component, the process shows different shapes of inter-temporal correlation.

The fBM shows three different states:

1. if  $H = 1/2$  then the process presents a Brownian motion.
2. if  $H > 1/2$  then the process presents positive autocorrelation.
3. if  $H < 1/2$  then the process presents negative autocorrelation.

Rough volatility models employ a fractional Brownian motion (fBm). The uniqueness of the fBm is that the increments are not independent. As previously mentioned,  $H$  is a continuous Gaussian process  $\{B_t^H, t \in \mathbf{R}\}$ ,  $B_t^H = 0$  with a mean of  $\mathbb{E}[B_t^H] = 0$ , therefore, the covariance is:

$$cov(B_t^H, B_s^H) = \mathbb{E}[B_t^H B_s^H] = \frac{1}{2} (|t|^{2H} + |s|^{2H} - |t - s|^{2H}). \quad (4.2)$$

Here the parameter  $H$  is called Hurst parameter, Hurst exponent or Hurst index and this governs the roughness of the process. The relationship between  $H$  and the fBm process is the higher the value of  $H$ , the smoother the process. Theorised by Hurst (1951), the Hurst exponent serves as a tool for assessing the long-term memory in time-series. It is connected

to the autocorrelations within the series, specifically focusing on how these autocorrelations diminish as the lag, or the time interval, between pairs of data points increases. The increment process,  $X_t = B_{t+1}^H - B_t^H$ , is known as fractional Gaussian noise.

According to Hurst (1951), and as shown in Weron (2002) and Caporale & Plastun (2024), the Hurst exponent is determined using rescaled range analysis (R/S analysis). This process starts by segmenting a time-series of returns with total length  $L$  into smaller sub-series, each of length  $n$ , resulting in  $d$  such sub-series. For each sub-series, labeled  $m = 1$  to  $d$ , we first calculate its mean ( $E_m$ ) and standard deviation ( $S_m$ ). Next, we normalise the data in each sub-series ( $Z_{i,m}$ ) by deducting its sample mean to obtain  $K_{i,m} = Z_{i,m} - E_m$ , where  $i$  ranges from 1 to  $n$ . Subsequently, a cumulative time-series  $Y_{i,m}$  is created for each  $i$  in the sub-series, defined as  $Y_{i,m} = \sum_{j=1}^i K_{j,m}$ . We determine the range  $R_m$ , which is the difference between the maximum and minimum of the cumulative time-series  $Y_{i,m}$ . After finding  $R_m$ , we rescale it by dividing by the standard deviation  $S_m$  to get  $R_m/S_m$ . The final step involves computing the average of these rescaled ranges across all sub-series of length  $n$ . The  $R/S$  is defined as,

$$(R/S)_n = \frac{1}{d} \sum_{m=1}^d R_m/S_m. \quad (4.3)$$

From Weron (2002) the  $R/S$  asymptotically follows the relation:  $(R/S)_n \sim cn^H$ . Therefore, the value of  $H$  can be obtained by running a simple linear regression over a sample of increasing time horizons  $\log(R/S)_n = \log c + H \log n$ . Here,  $c$  is a constant and the slope of the regression ( $H$ ) is an estimate of the Hurst exponent.

We follow Caporale & Plastun (2024) R/S methodology to compute the Hurst exponent, and

we use Equation 4.1 detailed in the methodology section to calculate the returns.

#### 4.4.2 GARCH-Type Models

Engle (1982) contributed a novel framework for forecasting the volatility in financial markets, producing an approach to financial time-series analysis and leading to several developments in the field. The GARCH model was first introduced in Bollerslev (1987) seminal paper, he generalised Engle (1982) ARCH model, introducing a processes to capture the impact of past shocks (squared residuals) on current volatility whilst capturing influence of past volatility. Regarding the limitations of the GARCH models, these models may not always capture extreme events effectively. This is due to the assumption of a normal distribution for residuals detailed in Bollerslev (1987). This assumption is a limitation because the normal distribution has thin tails and a kurtosis of 3, while financial data often exhibits excess kurtosis with higher peaks and is often skewed, meaning financial returns are not perfectly symmetrical around the mean. The development of variations like EGARCH allow for skewness and leptokurtosis in the distribution of residuals, which have more capability in capturing extreme events. The EGARCH model, primarily introduced by Nelson (1992) captures an asymmetric phenomenon known as the leverage effect puzzle, thus, the models captures both negative and positive shocks (good and bad news) relative to a positive or negative relationship to returns. Given EGARCH is a more flexible framework for modelling financial time-series volatility, we include this model in our analysis. The GARCH Model specification is expressed as:

$$y_t = \mu_t + \sigma_t \epsilon_t, \quad \epsilon_t \sim i.i.d(0, 1), \quad (4.4)$$

$$r_t = \sigma_t \epsilon_t, \quad r_t | \chi_{t-1} \sim N(0, \sigma_t^2), \quad (4.5)$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 r_{t-1}^2 + \dots + \alpha_q r_{t-q}^2 + \beta_1 \sigma_{t-1}^2 + \dots + \beta_p \sigma_{t-p}^2. \quad (4.6)$$

The GARCH( $p, q$ ) Model is expressed as:

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i r_{t-i}^2 + \sum_{i=1}^p \beta_i \sigma_{t-i}^2. \quad (4.7)$$

From this equation, the GARCH(1, 1) is expressed as:

$$\sigma_t^2 = \alpha_0 + \alpha_1 r_{t-1}^2 + \beta_1 \sigma_{t-1}^2, \quad (4.8)$$

where  $r_t$  is the return time-series value at time,  $\mu$  is the mean of the GARCH model,  $\epsilon_t$  is the model's residual at time  $t$ ,  $\sigma_t$  is the conditional volatility at time  $t$ ,  $p$  is the order of the ARCH component model,  $\alpha_i$  are the parameters of the ARCH component model,  $q$  is the order of the GARCH component model,  $\beta_i$  are the parameters of the GARCH component model.

The EGARCH( $p, q$ ) model is defined as follows:

$$\ln \sigma_t^2 = w + \sum_{i=1}^p \alpha_i \ln \sigma_{t-i}^2 + \sum_{i=1}^q \beta_i g(Z_{t-i}), \quad (4.9)$$

where  $g(Z_t) = \theta Z_t + \lambda(|Z_t| - E(|Z_t|))$ ,  $\sigma_t^2$  is the volatility,  $\omega$ ,  $\beta$ ,  $\alpha$ ,  $\theta$  and  $\lambda$  are coefficients.

$Z_t$  represents the density of the generalised error distribution.

#### 4.4.2.1 Estimating GARCH-Type Model Parameters

We consider the entire time-series of daily returns using close-to-close values from the high-frequency data. To compute daily returns, we compute the logarithm of today's last intraday close price divided by yesterday's last intraday close price (Visser (2011)). We chose to consider the dates from 1<sup>st</sup>, January, 2008 to 1<sup>st</sup>, January, 2023. We compute rolling GARCH( $p, q$ )-type models with a rolling window of 22 to constitute for information over the course of a month. We test a range of values for  $p$  and  $q$ , where  $p$  is the lag order of asymmetric shocks (innovations) and  $q$  is the order of lagged volatility. We fit our model to the data using the maximum likelihood method and set  $p$  and  $q$  to various integers and chose our model depending on the lowest AIC and BIC. We set the distribution for the GARCH-type models, where the Normal distribution is selected for the GARCH model and Generalised Error Distribution (GED) for EGARCH models. To guarantee successful optimisation convergence, if needed we scale the  $R_t$ . We extract the residuals and conditional volatility values from our GARCH-type computations and rescale these values if they were needed to be scaled for optimisation convergence. Therefore,  $R_t$  is rescaled if the scale of  $R_t$  is likely to produce convergence issues when estimating model parameters. The volatility conditional variance and residuals at time  $t$  are extracted, and computed using time  $(t - 1)$  information. Our methodology differs from Kim & Won (2018), as we fit our models using daily returns based on the lowest AIC and BIC in a rolling window.

### 4.4.3 Long Short-Term Memory

#### 4.4.3.1 Neural Network Models

Long Short-Term Memory (LSTM) networks, a type of recurrent neural network, were designed to recognise and learn sequential patterns and is capable of learning long-term dependencies. LSTMs process data sequences through internal loops which undergo training using back-propagation, where weight adjustments are made based on the gradient computed through the chain rule. In standard recurrent networks, using activation functions like sigmoid and tanh can lead to extremely small or large gradients, resulting in the vanishing or exploding gradient problem. This issue makes it challenging for the network to learn from long-range data dependencies. To overcome these limitations, LSTMs, introduced by Hochreiter & Schmidhuber (1997) employ memory cells and gates. Memory cells and gates enable the network to retain information over extended periods or discard irrelevant data. In the context of this study, we compute hybrid models similarly to Kim & Won (2018) that integrates parameters from multiple GARCH-type models with interest rates into a neural network framework. In their case they apply the 3-year Korea Treasury Bond (KTB) interest rate and 3-year AA-grade corporate bond (CB) interest rates.

This approach contrasts with conventional econometric methods, which typically involve more assumptions and modelling constraints. Neural network models, including LSTMs, are advantageous as they autonomously learn high-level features with less reliance on predetermined assumptions.

The architecture of the LSTM cell relies on a mechanism which holds the past inputs with the network or decides to reset the inputs to dismiss past information held in the network. Regarding the LSTM model, the cell state which is updated at each time  $t$  is represented as  $c_t$  by Equation 4.10:

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t, \quad (4.10)$$

where  $c_t$  is the cell state,  $f_t$  is the forget gate, computed from  $h_{t-1}$  which is the last hidden state, and  $i_t$  is the input gate which is computed in a similar method.  $\tilde{c}_t$  represents candidate for cell state at time  $t$  and  $\odot$  denotes the Hadamard product (element-wise product), and  $x_t$  is the current input. The mathematical equations for the aforementioned variables are defined below:

$$f_t = s(W_f x_t + U_f h_{t-1} + b_f), \quad (4.11)$$

$$i_t = s(W_i x_t + U_i h_{t-1} + b_i), \quad (4.12)$$

$$o_t = s(W_o x_t + U_o h_{t-1} + b_o), \quad (4.13)$$

$$h_t = o_t \odot \tanh(c_t), \quad (4.14)$$

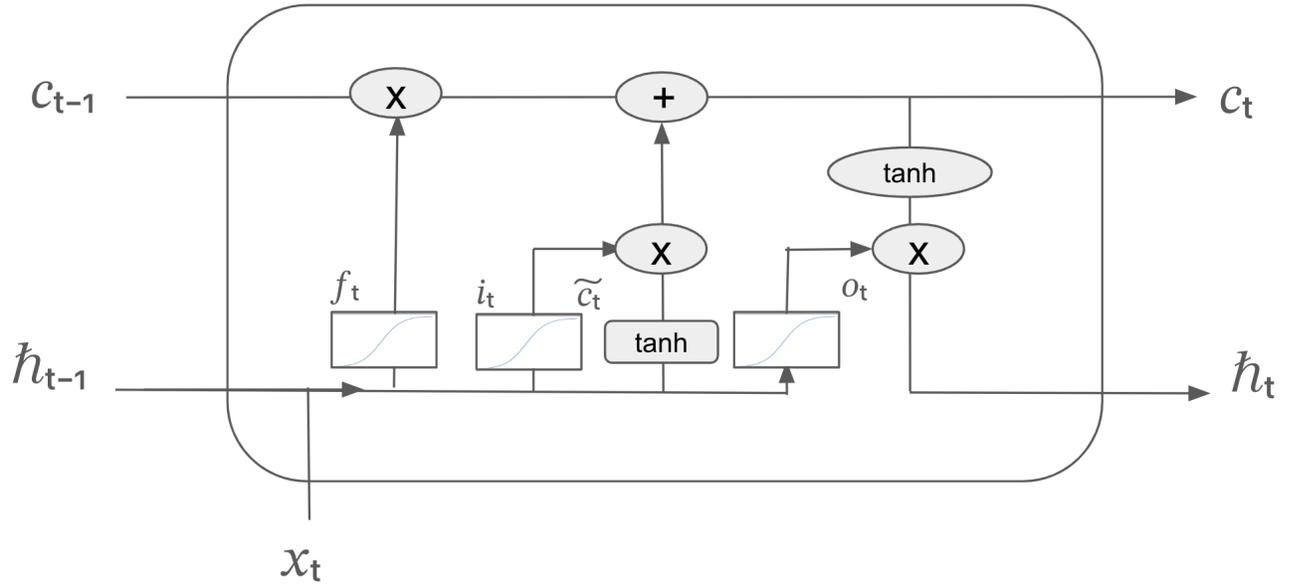
$$\tilde{c}_t = \tanh(W_c x_t + U_c h_{t-1} + b_c), \quad (4.15)$$

where the initial values are  $c_0 = 0$  and  $h_0 = 0$ . The subscript  $t$  indexes the time step, and the superscripts  $d$  and  $h$  refer to the number of input features and number of hidden units, respectively:

- $x_t \in \mathbb{R}^d$  : denotes the input vector to the LSTM unit.
- $f_t \in (0, 1)^h$ : is the forget gate's activation vector.
- $i_t \in (0, 1)^h$ : is the input or update gate's activation vector.
- $o_t \in (0, 1)^h$ : describes the output gate's activation vector.
- $h_t \in (-1, 1)^h$ : is the hidden state vector also known as output vector of the LSTM unit.
- $\tilde{c}_t \in (-1, 1)^h$ : is the cell input activation vector.
- $c_t \in \mathbb{R}^h$ : is the cell state vector.
- $W \in \mathbb{R}^{h \times d}$ ,  $U \in \mathbb{R}^{h \times h}$  and  $b \in \mathbb{R}^h$ : are the weight matrices and bias vector parameters which are learned during training.
- $s$ : signifies the sigmoid function.

The diagram of an LSTM cell is shown in Figure 4.4. LSTM layers have three gates that regulate the flow of information which are, the forget gate, input gate, and the output gate. The forget gate  $f_t$  in Equation 4.11 computes a weighted sum of  $x_t$  and  $h_{t-1}$ , and incorporates a bias. It then transforms this sum into a value ranging from 0 to 1 via the sigmoid function ( $s$ ). A value of 0 for  $f_t$  implies that no input information is being carried forward, while a value of 1 signifies that all input information is retained. Thus, the forget gate effectively modulates the extent of past cell state information ( $c_{t-1}$ ) to be included in the current cell state update at time  $t$ . The input gate, denoted as  $i_t$ , governs the assimilation

Figure 4.4: LSTM Cell Illustration



This illustration depicts a Long Short-Term Memory (LSTM) cell process through time.

of new information into the cell state ( $c_t$ ). Meanwhile,  $o_t$  determines the new information to be considered at time  $t$ , with its output after passing through the  $\tanh$  function ranging between -1 and 1. The cell state at time  $t$ , denoted as  $c_t$ , is derived from the integration of past cell state and new information as regulated by the forget and input gates. The output value  $h_t$  is then produced by filtering this updated cell state  $c_t$  through the output gate  $o_t$ , where  $c_t$  is also passed through the  $\tanh$  function to ensure its values lie between -1 and 1. The selected values are then multiplied by  $o_t$  to form the final output. This entire process updates the cell state from  $c_{t-1}$  to  $c_t$ , differentiating between relevant and irrelevant information, and resulting in the output  $h_t$ . The LSTM model, composed of these memory blocks, learns through back-propagation.

Empirically, we propose the data ( $R_t$  the daily log returns composed of close-to-close high frequency values), and specific statistical methods and econometric factors (GARCH con-

ditional volatility, EGARCH conditional volatility, GARCH residuals, EGARCH residuals, and Hurst exponent) of the stock index with macroeconomic data (interest rate) at each daily time  $t$  and feed to the LSTM model. Namely,  $x_t$  now consists of the specific factors of the stock index. Ultimately, the implicit connections among the stock index factors will be modelled to provide beneficial information in forecasting volatility.

#### 4.4.4 Realised Volatility

From Barndorff-Nielsen & Shephard (2002), the Realised Volatility (RV) theoretical properties are constructed from high-frequency data. To calculate the RV of day  $t$ , we use the following Equation 4.16:

$$RV_t = \sum_{\tau=1}^n R_{\tau}^2, \quad (4.16)$$

where  $R_{\tau}$  is the 5-minute log-return at time  $\tau = 1$  and  $n = 78$ , this is to represent 1 day given 78 five-minute time periods is equivalent to 1 day.  $R_{\tau} = \ln \left( \frac{P_{\tau}}{P_{\tau-1}} \right)$ , where  $P_{\tau}$  is the price of the index at the 5-minute time  $\tau$ .

#### 4.4.5 Development of the Models

In the development of our hybrid model, we delineate several pivotal variables. The model's input gate incorporates explanatory variables denoted by  $x_t$ , encompassing 'log returns', 'GARCH conditional volatility', 'GARCH residuals', 'EGARCH conditional volatility', 'EGARCH residuals', 'rolling Hurst' exponent, and the 'interest rate'. The dimensionality of the LSTM input is configured to reflect the count of these explanatory variables. Additionally,

we establish the training parameters by setting the number of epochs and learning rate at 100 and 0.001 respectively, and specifying the architecture to contain various LSTM layers. The combination of 100 epochs and a learning rate of 0.001 provides a balance in terms of learning efficiency and model performance. Each LSTM layer contains 50 neurones, where each neurone in the LSTM layer has its own memory cell to maintain information, and the dense (fully connected) layer contains 1 neurone for making a single output prediction. The batch size is set to 32 to reflect the number of training samples processed before the model's internal parameters (weights) are updated. The criterion for the LSTM's loss function is designated as the mean-squared error, aligning with a regression framework. We deploy the Adam optimiser from Kingma & Ba (2014) to train the LSTM.

Fundamentally, the next step-ahead forecasts are influenced by the trained values. The LSTM model uses the learned patterns from the training data to make predictions on the test data. During training, the model learns the underlying temporal patterns and dependencies, which it then uses to predict future values. An illustration of the sliding window design is shown in Figure 4.5. It worth mentioning with our methodology, this approach helps the model concentrate on the latest trends and patterns. Since financial market time-series data are highly dynamic, recent data is generally more relevant for short-term forecasting. Sliding windows enable the model to swiftly adapt to new patterns and changes in the market. Expanding windows include all historical data up to the current point, which can reduce the influence of recent information with older, potentially less relevant data. This can make the model less responsive to recent changes, thereby reducing its predictive accuracy for short-term forecasts. In terms of computational load, sliding windows have a constant

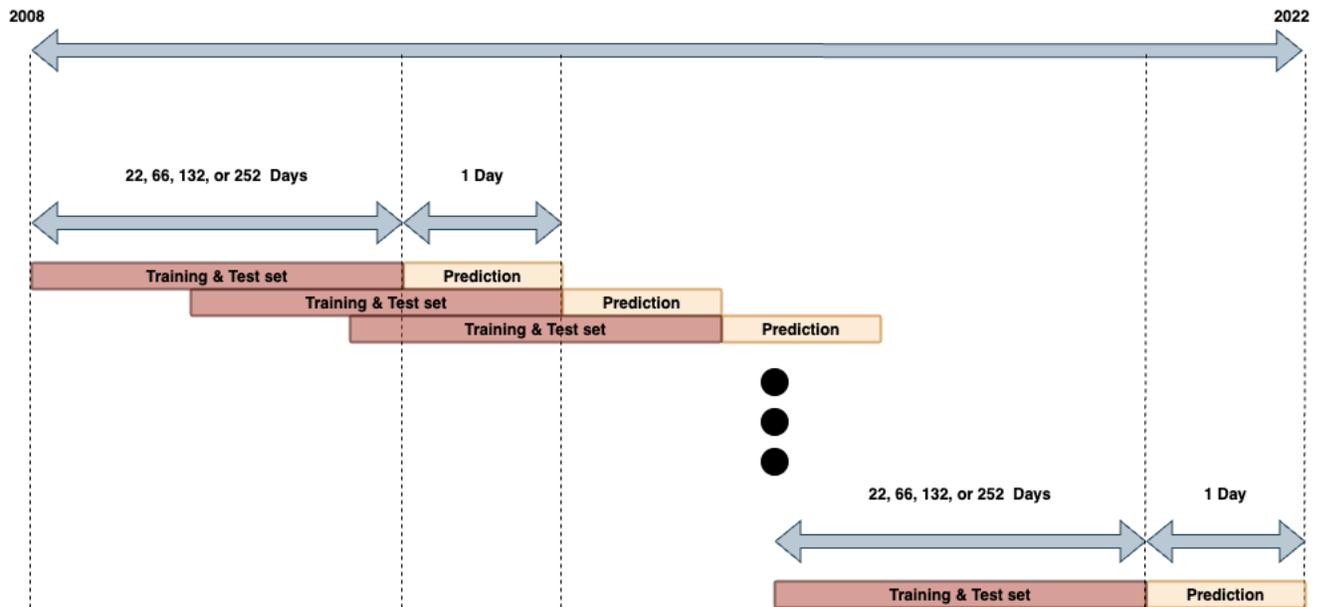
complexity over time due to their fixed size. In contrast, expanding windows increase in size with each step, leading to greater computational demands as more data is incorporated. By restricting the window size, sliding windows also serve as a form of regularisation, helping to prevent the model from overfitting to the entire dataset. Expanding windows, however, are at risk of overfitting, as the model may learn noise and irrelevant patterns from older data points, diminishing its ability to generalise and accurately predict new, unseen data. The time complexity of the sliding window is expressed as,  $O((n - k) \cdot k)$ . The outer loop runs  $O(n - k)$  times, where  $n$  is length of the data,  $k$  is window size, and  $O$  is the big  $O$  notation. For each iteration, the slicing operation takes  $O(k)$  time as it involves copying  $w$  rows, and appending to arrays  $X$  and  $y$  is  $O(1)$ .<sup>2</sup> Therefore, the time Complexity is:  $O((n - k) \cdot k)$ . For the expanding window method, the time complexity is expressed as  $O(n^2)$ . The functions starts with a minimum window size and iterates to the total length of the data. For  $i = m$ , it slices  $m$  rows. For  $i = m + 1$ , it slices  $m + 1$  rows, and so on, up to  $i = n - 1$ . Thus, the total number of operations can be represented as,  $\sum_{i=m}^{n-1} i = (m) + (m + 1) + (m + 2) + \dots + (n - 1)$ . This is an arithmetic series where the sum of an arithmetic series can be calculated as,  $(n - m)/2 \times (m + (n - 1)) = (n - m)(n + m - 1)/2$ . For a large  $n$  and  $m$ , the dominant term will be  $n^2/2$ . Therefore, the time complexity simplifies to  $O(n^2)$ . For computational complexity and predictive accuracy reasons, we focus on implementing the sliding window method.

Furthermore, regarding the LSTM model, values used for prediction at  $t + 1$  are indirectly used to predict values at  $t + 2$ . The model relies on the input sequences to make predictions.

---

<sup>2</sup> $X$  and  $y$  are the input features and  $RV_t$ , respectively.

Figure 4.5: Sliding Window Design for Training and Test Set and Predictions over the Entire Period



This illustration depicts the sliding window design, where in different scenarios varying lengths of days for the training set and test set is used for the next day ahead prediction.

We set our training set time-series to consist of various percentages (67%, 75%, and 80%) of time-series of the explanatory variables, and the test set time-series to be the remaining time-series. The training set is split into training and testing sets for the explanatory variables, using a MinMaxScaler function to transform features by scaling each feature to a given range. The training process involves calculating the loss with 100 epochs, and updating the parameters using back-propagation to minimise the loss, with the realised volatility is set as the target value. Utilising the trained network alongside subsequently acquired data, leveraging the LSTM model’s ability to recall and utilise long-term dependencies in the data, we facilitate a one-step-ahead forecast.

Regarding the GARCH-LSTM model, the inputs are defined to include 'log returns', 'GARCH conditional volatility', 'GARCH residuals', and 'interest rate'. Similarly, the EGARCH-LSTM

Table 4.3: Input Variables for Models

Models	Input Variables( $x_t$ )							
	$RV_t$	Log Returns	Interest Rate	GARCH Conditional Volatility	GARCH Residuals	EGARCH Conditional Volatility	EGARCH Residuals	Rolling Hurst Exponent
GARCH-LSTM	x	x	x	x	x			
EGARCH-LSTM	x	x	x			x	x	
rGARCH-LSTM	x	x	x	x	x			x
rEGARCH-LSTM	x	x	x			x	x	x
rGE-LSTM	x	x	x	x	x	x	x	x

This table describes the inputs for the LSTM models.

model incorporates 'log returns', 'EGARCH conditional volatility', 'EGARCH residuals', and 'interest rate' as its inputs. The rGARCH-LSTM, rEGARCH-LSTM and rGE-LSTM (which encompasses both EGARCH and GARCH features) models further extends this input set by integrating the 'rolling Hurst exponent'. We use these input variables to compute the target variable the forecasted realised volatility with the emphasis to minimise loss functions MAE and RSME to validate the forecasting power of the different realised volatility models of the SPX log-returns. These methodological choices to model volatility leverage econometric, statistical, and machine learning techniques to capture market volatility. Table 4.3 lists all input variables for the models.

#### 4.4.6 Loss Functions and Tests

In this sub-section, we employ a comprehensive array of loss functions to evaluate model performance. Specifically, we utilise the Mean Absolute Error (MAE), the Mean Squared Error (MSE), and the Root Mean Squared Error (RMSE). Collectively, these three loss functions constitute the evaluative framework for our analysis. Additionally, we compute the

Diebold-Mariano (DM) Test from Diebold & Mariano (1995) to compare forecasting accuracy, and the Wilcoxon signed-rank test from Wilcoxon (1947).

1. Mean Absolute Error (MAE)

$$MAE = \frac{1}{n} \sum_{t=1}^n |RV_t - \widehat{RV}_t|, \quad (4.17)$$

2. Mean Squared Error (MSE)

$$MSE = \frac{1}{n} \sum_{t=1}^n \left( RV_t - \widehat{RV}_t \right)^2, \quad (4.18)$$

3. Root Mean Squared Error (RMSE)

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n \left( RV_t - \widehat{RV}_t \right)^2}, \quad (4.19)$$

where  $RV_t$  denotes the realised volatility at time  $t$ , and  $\widehat{RV}_t$  denotes the predicted realised volatility at time  $t$ .

To rigorously compare the forecasting accuracy of the LSTM models, we employ the Diebold-Mariano (DM) Test, as proposed by Diebold & Mariano (1995). This statistical test evaluates the null hypothesis that two competing forecasting models exhibit equivalent predictive accuracy. A positive DM statistic suggests that the primary model yields smaller forecast

errors than the comparative model, while a negative DM statistic indicates the opposite. A DM statistic close to zero denotes minimal difference in predictive accuracy between the two models. Several journal papers focused on LSTM predictive accuracy, including Kim & Won (2018), Petrozziello et al. (2022), and Garcia-Medina & Aguayo-Moreno (2024), employ the DM test. The Wilcoxon Signed-Rank test, introduced by Wilcoxon (1947), is a non-parametric statistical test used to compare two related samples or matched pairs to determine whether their population mean ranks differ. The null hypothesis for both the DM (Diebold-Mariano) and WS (Wilcoxon Signed-Rank) tests is that the two predictive models have the same level of accuracy. If the p-value is less than 0.05, we reject the null hypothesis, indicating that the predictive accuracy of the two competing models is significantly different. Conversely, if the p-value is greater than 0.05, we fail to reject the null hypothesis at a significance level of 5% or lower, suggesting no significant difference in predictive accuracy between the models. For macroeconomic events, we compare forecasting accuracy of the LSTM models by measuring the error metrics.

#### **4.4.7 Three-State First-Order Markov Switching Regression Model**

Lastly, we evaluate the performance of the LSTM models across high, medium, and low volatility states to assess the efficacy of our approach. Initial tests are conducted to calculate the Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) associated with political and financial events. The analysis employs a configuration for one-day step forecasts and volatility decomposition is achieved through a three-state first-order Markov switching

regression model. The specification of the Markov switching model is defined in Subsection 4.4.8. These volatility states are tested to determine how our models behave in the presence of volatility clustering, which is prominent during these periods of elevated market volatility triggered by significant events, thereby facilitating the emergence of volatility clustering phenomena. We then map the identified volatility states to the event days and proceed with forecasting using the LSTM models. Specifically, we generate one-day-ahead forecasts corresponding to or immediately following the event dates. We evaluate the accuracy of these forecasts by calculating the Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) using the last five days of data, including the event day.

#### 4.4.8 Markov Switching Regression Model Specification

Consider the evolution of a series  $y_t$ , where  $t = 1, 2, \dots, T$ , is characterised by three states, as shown in the models below:

- State 1:

$$y_t = \mu_1 + \varphi_1 y_{t-1} + \varepsilon_t, \quad (4.20)$$

- State 2:

$$y_t = \mu_2 + \varphi_2 y_{t-1} + \varepsilon_t, \quad (4.21)$$

- State 3:

$$y_t = \mu_3 + \varphi_3 y_{t-1} + \varepsilon_t, \quad (4.22)$$

where  $\mu_1$ ,  $\mu_2$  and  $\mu_3$  are the intercept terms in state 1, state 2, and state 3, respectively;  $\varphi$  is the AR parameter; and  $\varepsilon_t \sim N(0, \sigma_t^2 s_t)$ . In the simplest case, we can express this model as a state-dependent intercept term for  $k$  states:

$$y_t = \mu_{s_t} + \varphi_{s_t} y_{t-1} + \varepsilon_t, \quad (4.23)$$

where  $\mu_{s_t} = \mu_1$  when  $s_t = 1$ ,  $\mu_{s_t} = \mu_2$  when  $s_t = 2$ , ..., and  $\mu_{s_t} = \mu_k$  when  $s_t = k$ . The probability that  $s_t$  is equal to  $j \in (1, \dots, k)$  depends only on the most recent realisation,  $s_{t-1}$ , and is given by

$$\mathbb{P}(s_t = j | s_{t-1} = i) = p_{ij}. \quad (4.24)$$

All transitions from one state to the other can be sourced in a  $k \times k$  transition matrix which governs the evolution of the Markov chain. All elements of  $\mathbf{P}$  are nonnegative and each column sums to 1:

$$\mathbf{P} = \begin{bmatrix} p_{11} & \cdots & p_{k1} \\ p_{12} & \cdots & p_{k2} \\ \vdots & \ddots & \vdots \\ p_{1k} & \cdots & p_{kk} \end{bmatrix} \quad (4.25)$$

## 4.5 Empirical Analysis

### 4.5.1 Model Configuration

This section focuses on the empirical analysis. We tune our model by comparing performance across various configurations. Thereafter, we compare performance across the training set percentage, the number of LSTM layers to process our forecasts, various  $n$ -step ahead forecasts, the entire dataset, and finally a comparison against political and financial events. For robustness, we employ various loss functions and statistical tests.

#### 4.5.1.1 Training Set Percentages

The results presented in Table 4.4 indicate that optimal performance is achieved when 75% of the data is used for training to forecast the one-step ahead forecast of  $RV_t$ . Specifically, the 75% training set proportion yields the lowest Mean Absolute Error (MAE) in 12 out of 20 cases, also, 75% training set proportion yields the lowest Root Mean Squared Error (RMSE) in 10 out of 20 cases. Thus, in subsequent analysis we choose to utilise 75% of the data for training the LSTM model to ensure optimal forecasting accuracy.

Table 4.4: Comparison of  $RV_t$  Forecasting Errors of Various Models over Different Time Horizons for Various Training Set Percentages

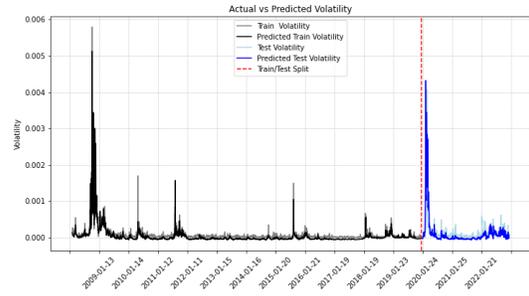
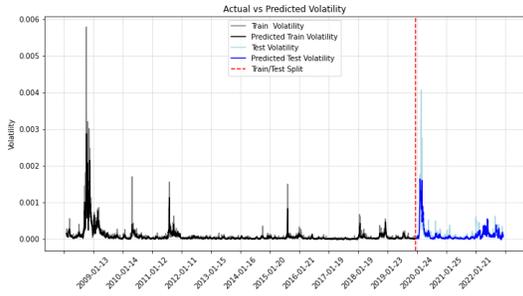
Model	Training Set	MAE				RMSE			
		Window Length				Window Length			
		22	66	132	252	22	66	132	252
GARCH-LSTM Model	67%	0.3519	0.3791	0.5045	0.3080	0.9950	0.9995	1.1490	1.1250
	75%	0.4063	0.4518	<b>0.3519</b>	<b>0.2934</b>	<b>0.9766</b>	0.9576	<b>0.9258</b>	1.0920
	80%	<b>0.3405</b>	<b>0.3458</b>	0.3567	0.2983	1.0420	<b>0.9146</b>	0.9731	<b>0.6660</b>
EGARCH-LSTM Model	67%	0.4490	<b>0.3224</b>	0.3910	<b>0.2951</b>	0.9971	<b>1.0030</b>	1.0480	1.0510
	75%	<b>0.3287</b>	0.3883	<b>0.3167</b>	0.2956	<b>0.8572</b>	1.0470	<b>0.8731</b>	1.0630
	80%	0.3909	0.3500	0.3929	0.3399	1.2190	1.0670	0.9818	<b>0.6686</b>
rGARCH-LSTM Model	67%	<b>0.3535</b>	0.3343	0.6268	0.3008	0.8793	0.8771	1.0710	1.0210
	75%	0.3664	0.3933	0.4043	<b>0.2915</b>	<b>0.7813</b>	0.9979	0.8980	1.0600
	80%	0.3748	<b>0.3169</b>	<b>0.3181</b>	0.2942	0.8686	<b>0.8133</b>	<b>0.8511</b>	<b>0.5931</b>
rEGARCH-LSTM Model	67%	0.4059	0.3425	0.3790	0.2784	0.9747	0.8393	1.0340	0.9808
	75%	0.3485	<b>0.3110</b>	<b>0.2955</b>	<b>0.2773</b>	<b>0.7857</b>	<b>0.8073</b>	<b>0.8371</b>	1.0750
	80%	<b>0.3457</b>	0.3437	0.3251	0.4343	0.8910	0.8457	0.9750	<b>0.7012</b>
rGE-LSTM Model	67%	0.3993	0.4240	0.3442	0.2734	1.0100	0.9788	0.8545	1.0320
	75%	<b>0.3412</b>	<b>0.3224</b>	<b>0.3286</b>	<b>0.2877</b>	<b>0.8352</b>	<b>0.8159</b>	0.9818	1.0420
	80%	0.6858	0.5031	0.4033	0.2805	1.0260	1.1490	<b>0.8370</b>	<b>0.6079</b>

This table shows the MAE and RMSE of the entire dataset of one-day step forecasts using 2/3, 3/4, and 4/5 of the data. The values are computed over four separate sliding windows. Where window lengths are 22, 66, 132, and 252, respectively. We set configurations for the LSTM model which include; 1 layer, 100 epochs, and 0.001 learning rate. Errors are calculated by MAE and RMSE and have been adjusted with a  $1 \times 10^4$  multiple for readability.

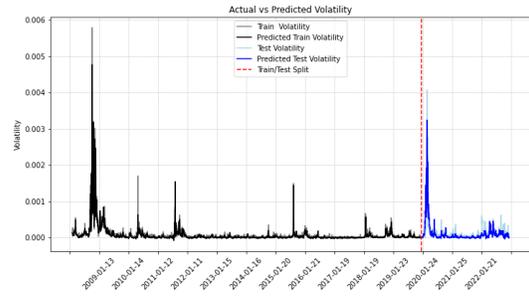
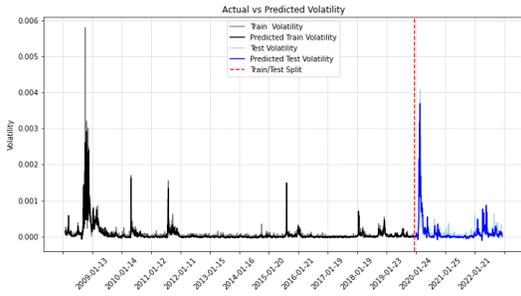
Figure 4.6 shows a comparison of LSTM results using the entire dataset for one-day step ahead forecasts with the sliding window method and 3/4 of the data, where the window length is 22. This figure illustrates how rGE-LSTM forecasts captures  $RV_t$  most accurately. The EGARCH-LSTM model impressively captures large shocks like the rGE-LSTM, however, there are instances where it is unable to forecast smaller shocks in comparison to the rGE-LSTM. From Tables 4.4, 4.5 we can see the optimal configuration for 1-step forecasts are detailed in Table 4.6.

Figure 4.6: RV True Values vs. RV Predicted Values of Forecasting Models

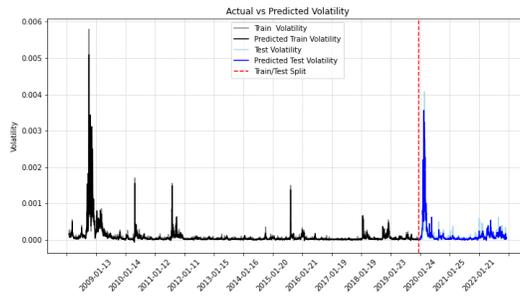
(a) GARCH-LSTM realised volatility vs. trained and predicted realised volatility



(c) rGARCH-LSTM realised volatility vs. trained and predicted realised volatility



(e) rGE-LSTM realised volatility vs. trained and predicted realised volatility



This figure depicts a collection of realised volatility vs. trained and predicted realised volatility for the LSTM models.

#### 4.5.1.2 Number of LSTM Layers

Next, we analyse the impact of the number of LSTM layers in relation to varying window lengths. In Table 4.5, for each window length, we observe that 12 out of 20 instances with the lowest Mean Absolute Error (MAE) occur with a two LSTM layers. Forecasts with one and three LSTM layers both account for 8 out of 20 of the lowest MAEs. However, for each window length, we observe that 12 out of 20 instances with the lowest Root Mean Squared Error (RMSE) occur with a single LSTM layer.

From Tables 4.4 and 4.5 it is evident that, in most cases, larger window lengths correspond to lower MAEs. This finding aligns with the results presented in Table 6 by Kim & Won (2018), implying that larger sliding windows capture more information, thereby reducing MAE errors.

Table 4.5: Comparison of  $RV_t$  Forecasting Errors of Various Models over Different Time Horizons for Various Layers

Model	Number of LSTM Layers	MAE				RMSE			
		Window Length				Window Length			
		22	66	132	252	22	66	132	252
GARCH-LSTM Model	1	0.3777	<b>0.3459</b>	0.3990	<b>0.2854</b>	0.9556	<b>1.0140</b>	<b>0.9399</b>	1.0470
	2	0.3468	0.5474	<b>0.3892</b>	0.2862	1.0170	1.0710	1.0540	<b>1.0150</b>
	3	<b>0.3465</b>	0.4697	0.6739	0.3090	<b>0.9052</b>	1.1190	1.2470	1.0960
EGARCH-LSTM Model	1	<b>0.3156</b>	<b>0.3264</b>	0.4692	<b>0.2713</b>	<b>0.9253</b>	<b>0.9512</b>	<b>0.9504</b>	<b>1.0530</b>
	2	0.7608	0.3813	<b>0.3487</b>	0.3103	1.1850	1.2610	1.0640	1.1170
	3	0.3451	0.3317	0.3703	0.2808	1.0050	1.0070	1.1750	1.0920
rGARCH-LSTM Model	1	0.4628	0.4362	0.4033	0.2942	<b>0.8719</b>	<b>0.8953</b>	0.9345	1.0390
	2	0.3181	<b>0.3832</b>	<b>0.3324</b>	<b>0.2881</b>	0.9475	0.9362	0.9681	<b>1.0310</b>
	3	<b>0.3151</b>	0.6935	0.3519	0.3249	0.9412	1.1890	<b>0.9157</b>	1.0960
rEGARCH-LSTM Model	1	0.4094	0.3424	0.3355	0.2658	0.9317	<b>0.9139</b>	1.0190	<b>1.0100</b>
	2	<b>0.3299</b>	0.4790	<b>0.3144</b>	<b>0.2530</b>	<b>0.8569</b>	1.0570	<b>0.9031</b>	1.0490
	3	0.4530	<b>0.3293</b>	0.4373	0.2836	0.9332	1.0060	1.1270	1.0660
rGE-LSTM Model	1	0.4705	0.3559	0.3243	0.2855	0.9257	0.9272	<b>0.8225</b>	<b>1.0110</b>
	2	<b>0.2834</b>	<b>0.3007</b>	<b>0.2907</b>	<b>0.2777</b>	<b>0.8744</b>	<b>0.8476</b>	0.8376	1.0590
	3	0.3565	0.3080	0.3608	0.3009	0.9941	0.9912	1.0790	1.1060

This table shows the MAE of the entire dataset of one-day step forecasts using 1, 2 and 3 layers in the LSTM model. The values are computed over four separate sliding windows. Where window lengths are 22, 66, 132, and 252, respectively. We train the LSTM models with 75% of the data. We set configurations for the LSTM model which include; 1 layer with 50 neurones and a dense layer with one neurone, 2 layers with 50 neurones each and a dense layer, and 3 layers with 50 neurones each and a dense layer with one neurone. Additionally, we have 100 epochs, and 0.001 learning rate. Errors are calculated by MAE and RMSE have been adjusted with a  $1 \times 10^4$  multiple for readability.

Table 4.6: Optimal configurations for forecasts

Model	GARCH-LSTM		EGARCH-LSTM		rGARCH-LSTM		rEGARCH-LSTM		rGE-LSTM	
	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
Training Set %	75%, 80%	75%, 80%	67%, 75%	75%	80%	80%	75%	75%	75%	75%, 80%
Number of LSTM Layers	1	1	1	1	2	1	2	1 & 2	2	1 & 2

This table shows the best configurations for the LSTM models.

### 4.5.1.3 $n$ Days-Ahead Forecasts

We then proceed to analyse  $n$ -step ahead forecasts, where we have  $n$  representing the following values 5, 10, 132, 252. Tables 4.7 and 4.8 demonstrate that  $n$ -step ahead forecasts produce similar outcomes across different models, with the average MAE errors decreasing as the window length increases. However, we notice higher  $n$  from our  $n$ -step ahead prediction set (5, 10, 132, 252), are associated with higher MAE and RMSE errors, indicating that forecasting accuracy declines as the forecast horizon increases. Notably, models incorporating the roughness parameter, as measured by the Hurst exponent, consistently outperform non-rough models. Specifically, in Table 4.7, rough models achieve the lowest MAE errors in 10 out of 16 cases across all four sliding windows and four forecasting steps. Similarly, in Table 4.8, rough models record the lowest RMSE errors in 13 out of 16 instances under the same conditions. These results underscore the superior predictive performance of models that account for roughness.

Table 4.7: Comparison of  $RV_t$  MAE Forecasting Errors of Various Models over Different Time Horizons for Various  $n$  Days-Ahead

Model	$n$ -Step Ahead Prediction (Days)							
	5-Step				10-Step			
	Window Length (Days)				Window Length (Days)			
	22	66	132	252	22	66	132	252
GARCH(1,1)	$2.55 \times 10^3$	$9.91 \times 10^5$	$5.31 \times 10^2$	$5.91 \times 10^1$	$2.55 \times 10^3$	$9.92 \times 10^5$	$5.32 \times 10^2$	$5.92 \times 10^1$
GARCH-LSTM	0.4714	0.4484	0.4570	0.3646	0.4668	0.4920	0.4796	0.3734
EGARCH-LSTM	0.4591	0.4488	0.4769	0.3454	0.4574	0.4458	0.4943	0.3518
rGARCH-LSTM	0.4693	0.4865	0.4464	0.3429	0.4621	0.4863	0.4615	<b>0.3415</b>
rEGARCH-LSTM	0.4330	0.4329	0.4481	<b>0.3425</b>	0.4640	0.4817	0.4633	0.3489
rGE-LSTM	<b>0.4155</b>	<b>0.4309</b>	<b>0.4400</b>	0.3689	<b>0.4489</b>	<b>0.4457</b>	<b>0.4548</b>	0.3478
Model	132-Step				252-Step			
	Window Length (Days)				Window Length (Days)			
	22	66	132	252	22	66	132	252
	GARCH(1,1)	$2.64 \times 10^3$	$1.03 \times 10^6$	$5.50 \times 10^2$	$6.13 \times 10^1$	$2.73 \times 10^3$	$1.06 \times 10^5$	$5.70 \times 10^2$
GARCH-LSTM	<b>0.6490</b>	0.8126	0.5804	0.4780	<b>0.6851</b>	<b>0.6331</b>	0.5998	0.5072
EGARCH-LSTM	0.6515	<b>0.6528</b>	<b>0.5711</b>	0.4976	0.6962	0.6557	0.6373	<b>0.4974</b>
rGARCH-LSTM	0.6547	0.6937	0.6358	0.4681	0.7238	0.6817	<b>0.5789</b>	0.5070
rEGARCH-LSTM	0.6653	0.6863	0.6168	0.4823	0.7039	0.7060	0.6767	0.5226
rGE-LSTM	0.6802	0.6900	0.5860	<b>0.4685</b>	0.7069	0.7360	0.6715	0.5531

This table shows the MAE of the entire dataset of five-days, ten-days, one-hundred and thirty-two, and two-hundred and fifty-two step forecasts. The values are computed over four separate sliding windows. Where window lengths are 22, 66, 132, and 252, respectively. We train the LSTM models with 75% of the data. We set configurations for the LSTM model which include; 1 layer, 100 epochs, and 0.001 learning rate. Errors are calculated by MAE have been adjusted with a  $1 \times 10^4$  multiple for readability.

Table 4.8: Comparison of  $RV_t$  RMSE Forecasting Errors of Various Models over Different Time Horizons for Various  $n$  Days-ahead

Model	$n$ -Step Ahead Prediction (Days)							
	5-Step				10-Step			
	Window Length (Days)				Window Length (Days)			
	22	66	132	252	22	66	132	252
GARCH(1,1)	$1.43 \times 10^5$	$5.94 \times 10^7$	$2.20 \times 10^4$	$2.71 \times 10^3$	$1.43 \times 10^5$	$5.94 \times 10^7$	$2.20 \times 10^4$	$2.71 \times 10^3$
GARCH-LSTM	1.2940	1.2950	1.2620	1.2420	1.3610	1.5350	1.4580	1.2090
EGARCH-LSTM	1.3180	1.2870	1.2400	1.2490	1.3860	1.3560	1.3840	1.2070
rGARCH-LSTM	1.3740	1.5250	1.3140	1.2180	1.4020	1.5270	1.4360	1.1890
rEGARCH-LSTM	<b>1.2030</b>	1.2670	1.3240	1.2360	1.3940	1.3810	1.4630	1.2130
rGE-LSTM	1.2260	<b>1.2640</b>	<b>1.2010</b>	<b>1.1860</b>	<b>1.3130</b>	<b>1.3370</b>	<b>1.3340</b>	<b>1.1870</b>
Model	132-Step				252-Step			
	Window Length (Days)				Window Length (Days)			
	22	66	132	252	22	66	132	252
	GARCH(1,1)	$1.46 \times 10^5$	$6.04 \times 10^7$	$2.204 \times 10^4$	$2.76 \times 10^3$	$1.48 \times 10^5$	$6.15 \times 10^7$	$2.28 \times 10^4$
GARCH-LSTM	1.9900	2.2690	1.8230	1.5670	<b>2.0320</b>	<b>1.9430</b>	1.8250	<b>1.6170</b>
EGARCH-LSTM	2.0020	2.0550	1.8330	1.5760	2.0620	2.0400	1.9250	1.6480
rGARCH-LSTM	<b>1.9890</b>	2.0860	1.8790	1.5770	2.0960	2.0080	<b>1.7820</b>	1.6290
rEGARCH-LSTM	2.0030	2.0630	1.8680	1.5740	2.0590	2.0340	2.0070	1.6370
rGE-LSTM	2.0010	<b>2.0530</b>	<b>1.7890</b>	<b>1.5550</b>	2.0380	2.0810	1.9530	1.6610

This table shows the RMSE of the entire dataset of five-days, ten-days, one-hundred and thirty-two, and two-hundred and fifty-two step forecasts. The values are computed over four separate sliding windows. Where window lengths are 22, 66, 132, and 252, respectively. We train the LSTM models with 75% of the data. We set configurations for the LSTM model which include; 1 layer, 100 epochs, and 0.001 learning rate. Errors are calculated by RMSE have been adjusted with a  $1 \times 10^4$  multiple for readability.

#### 4.5.1.4 Model Comparison

In Table 4.9, we present the Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) metrics for a variety of models, including the GARCH(1,1) model. The inclusion of the GARCH(1,1) model allows for a comparative analysis of its forecasting performance across multiple years within the dataset. The evaluation focuses on the year 2008, a period marked by substantial financial and political events, and extends to sub-periods covering the years 2009 and 2010. A sliding window of 22 days is utilised in this analysis, reflecting the data's 252 trading days, and the models are assessed based on one-step-ahead forecasts. This approach is particularly well-suited for predicting the next day's outcomes in highly volatile market conditions.

Table 4.9, reveals that the rGE-LSTM model consistently delivers the lowest Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) values. Specifically, the rGE-LSTM model achieves the lowest MAEs and RMSEs for the years 2008, and the lowest RMSE in 2010. The errors captured by the LSTM-based models during the 2008 sample period are significantly higher—by an order of magnitude—compared to the errors in 2009 and 2010, a disparity attributed to the heightened volatility of events during 2008. The rEGARCH-LSTM model follows closely in performance. Furthermore, all LSTM models demonstrate superior performance relative to the GARCH(1,1) model, except for the year 2009 where GARCH(1,1) presents the lowest MAE and RMSE.

Table 4.9: Single Models vs. Hybrid GARCH-Type LSTM Comparison

Model		Loss Function	2008	2009	2010
Single Model	GARCH(1,1)	MAE	$0.1952 \times 10^4$	<b><math>0.5426 \times 10^{-2}</math></b>	$0.9729 \times 10^2$
		RMSE	$0.2311 \times 10^5$	<b><math>0.8326 \times 10^{-2}</math></b>	$0.1014 \times 10^2$
LSTM-Based Hybrid Model	GARCH-LSTM Model	MAE	2.2430	0.3394	0.3777
		RMSE	5.8910	0.5002	0.9556
	EGARCH-LSTM Model	MAE	2.5700	0.363	0.4662
		RMSE	6.7530	0.4854	1.0340
LSTM-Based Hybrid Model with Hurst	rGARCH-LSTM Model	MAE	2.2320	0.3175	0.3819
		RMSE	6.2490	0.4573	1.0070
	rEGARCH-LSTM Model	MAE	2.3670	0.2915	<b>0.3380</b>
		RMSE	6.2290	0.4174	0.9289
rGE-LSTM Model	MAE	<b>2.0650</b>	0.2589	0.3412	
	RMSE	<b>5.6390</b>	0.3614	<b>0.8352</b>	

This table shows the MAE and RMSE results of a sample data representing the years 2008, 2009 and 2010, with one-day step forecasts. Where window length is 22, and the training set percentage is 75%. We set configurations for the LSTM model which include; 1 layer, 100 epochs, and 0.001 learning rate. The single model (GARCH(1,1) is computed with a window length of 22. Errors are calculated by MAE and RMSE have been adjusted with a  $1 \times 10^4$  multiple for readability.

#### 4.5.1.5 Model Tests

To evaluate the statistical significance of these models, we apply the Diebold-Mariano (DM) test, as detailed in Table 4.10, which encompasses the full sample data from 2008 to 2022. The table reports 26 statistically significant results. If the test result is not significant, it suggests that the null hypothesis cannot be rejected, indicating insufficient evidence to assert that one model outperforms another in terms of forecasting accuracy. It is important to note that the magnitude of the DM statistic does not directly measure the superiority of one model over another; however, the sign of the statistic is indicative. A markedly negative value suggests a substantial deviation from the null hypothesis of equal predictive accuracy. The rGE-LSTM model exhibits the highest negative DM statistics among all models. These results, however, were statistically not significant when the rGE-LSTM and rEGARCH-LSTM

Table 4.10: DM Test Pair-Wise Comparison of GARCH-Type LSTM Models

Loss Function	Primary Model	Secondary Model				
		GARCH-LSTM	EGARCH-LSTM	rGARCH-LSTM	rEGARCH-LSTM	rGE-LSTM
MAE	GARCH-LSTM	-	-10.7749***	-0.4810	5.9964***	4.7972***
MSE		-	-1.8310	-0.7537	0.6525	2.6462*
MAE	EGARCH-LSTM	10.7749***	-	9.1780***	16.0898***	15.1738***
MSE		1.8310	-	0.5115	2.754*	4.3539***
MAE	rGARCH-LSTM	0.4810	-9.1780***	-	5.4353***	4.7417***
MSE		0.7537	-0.5115	-	1.4632	2.7084*
MAE	rEGARCH-LSTM	-5.9964***	-16.0898***	-5.4353***	-	-0.4568
MSE		-0.6525	-2.754*	-1.4632	-	2.588*
MAE	rGE-LSTM	-4.7972***	-15.1738***	-4.7417***	0.4568	-
MSE		-2.6462*	-4.3539***	-2.7084*	-2.588*	-

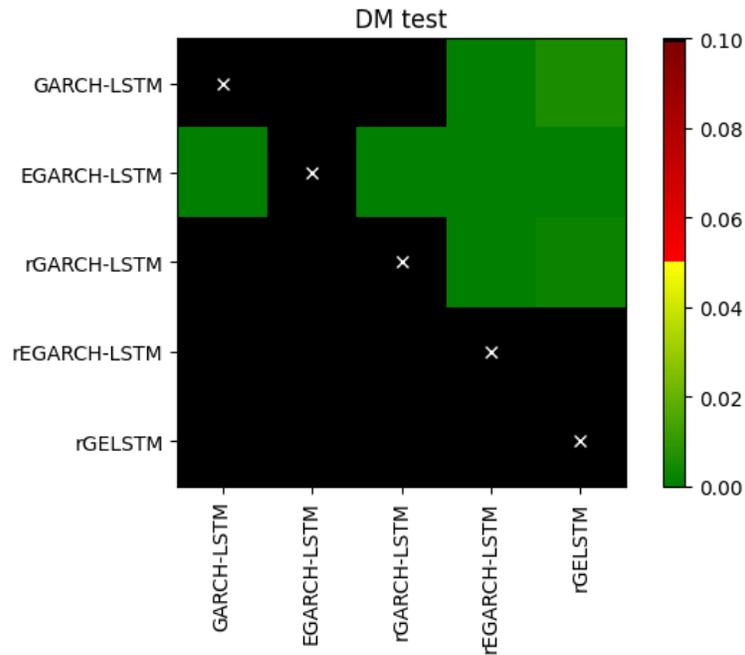
This table shows the DM test of the entire dataset of one-day step forecasts using 1 layer in the LSTM model. Where window length is 22, and train the LSTM models with 75% of the data. Additionally, we have 100 epochs, and 0.001 learning rate. Stars are only intended to flag levels of significance. If a p-value is  $\leq 0.05$  it is flagged with one star (\*), if a p-value is  $\leq 0.01$  it is flagged with two stars (\*\*), subsequently, if a p-value is  $\leq 0.001$  it is flagged with three stars (\*\*\*).

models are tested against each other.

Figure 4.7 presents a heat map depicting the outcomes of applying the Diebold-Mariano (DM) test for comparative forecast evaluation across a range of models detailed in Table 4.10. The multivariate DM test, as utilised by Lago et al. (2021), employs the Mean Absolute Error (MAE) as its loss function to assess the statistical significance of forecast accuracy differences among models. The heat map, presented in a checkerboard format, encodes the p-values derived from testing the null hypothesis that the forecast accuracy of the model on the vertical y-axis is not significantly better than that of the model on the horizontal x-axis.

In this heat map, shades closer to dark green indicate p-values approaching zero, signifying a more pronounced difference in forecast accuracy, with the model on the x-axis outperforming the model on the y-axis. In other words, p-values close to zero represent cases where the

Figure 4.7: Multivariate Diebold-Mariano(DM) Test



This illustration depicts a heat map of the Diebold-Mariano (DM) test of LSTM models using the MAE loss function.

forecast on the x-axis is significantly more accurate than the forecast on the y-axis. Conversely, black squares represent p-values that exceed the defined threshold in the colour scale, typically p-values equal to or higher than 0.10, suggesting no significant difference in forecast accuracy between the models compared.

Figure 4.7 illustrates that the rGE-LSTM and rEGARCH-LSTM models are the most accurate. The rGARCH model is presented as the next best model. However, it is not conclusive whether the rGE-LSTM is better than the rEGARCH-LSTM, given a test statistic of 0.4568 and a not significant p-value. For all rGE-LSTM and rEGARCH-LSTM DM tests, we reject the null hypothesis where p-values are below 0.05. This indicates a significant difference in predictive accuracy of the rGE-LSTM and rEGARCH-LSTM models compared to the other models.

Table 4.11: Wilcoxon Signed-Rank Test(WS) Pair-Wise Comparison of GARCH-type LSTM Models

Primary Model	Secondary Model				
	GARCH-LSTM	EGARCH-LSTM	rGARCH-LSTM	rEGARCH-LSTM	rGE-LSTM
GARCH-LSTM	-				
EGARCH-LSTM	<b>1177741.0***</b>	-			
rGARCH-LSTM	<b>2596888.0 ***</b>	<b>1717108.0***</b>	-		
rEGARCH-LSTM	<b>2929206.5***</b>	<b>1075948.0***</b>	<b>2073970.0***</b>	-	
rGE-LSTM	<b>1055885.0***</b>	<b>743538.0***</b>	<b>734019.0***</b>	<b>1109592.0***</b>	-

The values are p-values for the WS test and boldface values indicate when the p-value is less than 0.05. Here, stars are only intended to flag levels of significance. If a p-value is  $\leq 0.05$  it is flagged with one star (\*), if a p-value is  $\leq 0.01$  it is flagged with two stars (\*\*), subsequently, if a p-value is  $\leq 0.001$  it is flagged with three stars (\*\*\*).

We observe results from the Wilcoxon Signed-Rank Test (WS); it is worth mentioning that a larger test statistic suggests more pronounced observed differences, leading to a greater sum of signed ranks. From Table 4.11, it is evident that all pairs have statistically significant p-values. Hence, we reject the null hypothesis, indicating that one model in each pair has better predictive ability compared to the other model.

#### 4.5.2 Political and Financial Events and Volatility States

The objective of this analysis is to evaluate the performance of our predictive model under conditions of market volatility by comparing the loss metrics MAE and RMSE for the macro-economic events dataset.

We decompose the realised volatility into three states, high, medium, and low volatility states to assess the efficacy of our model. We conduct initial tests to calculate MAE and RMSE of the political and financial events. The analysis uses a configuration for one-day step forecasts using 1 layer in the LSTM model where the window length is 22, and the training

Table 4.12: Markov Switching Model Regression Results

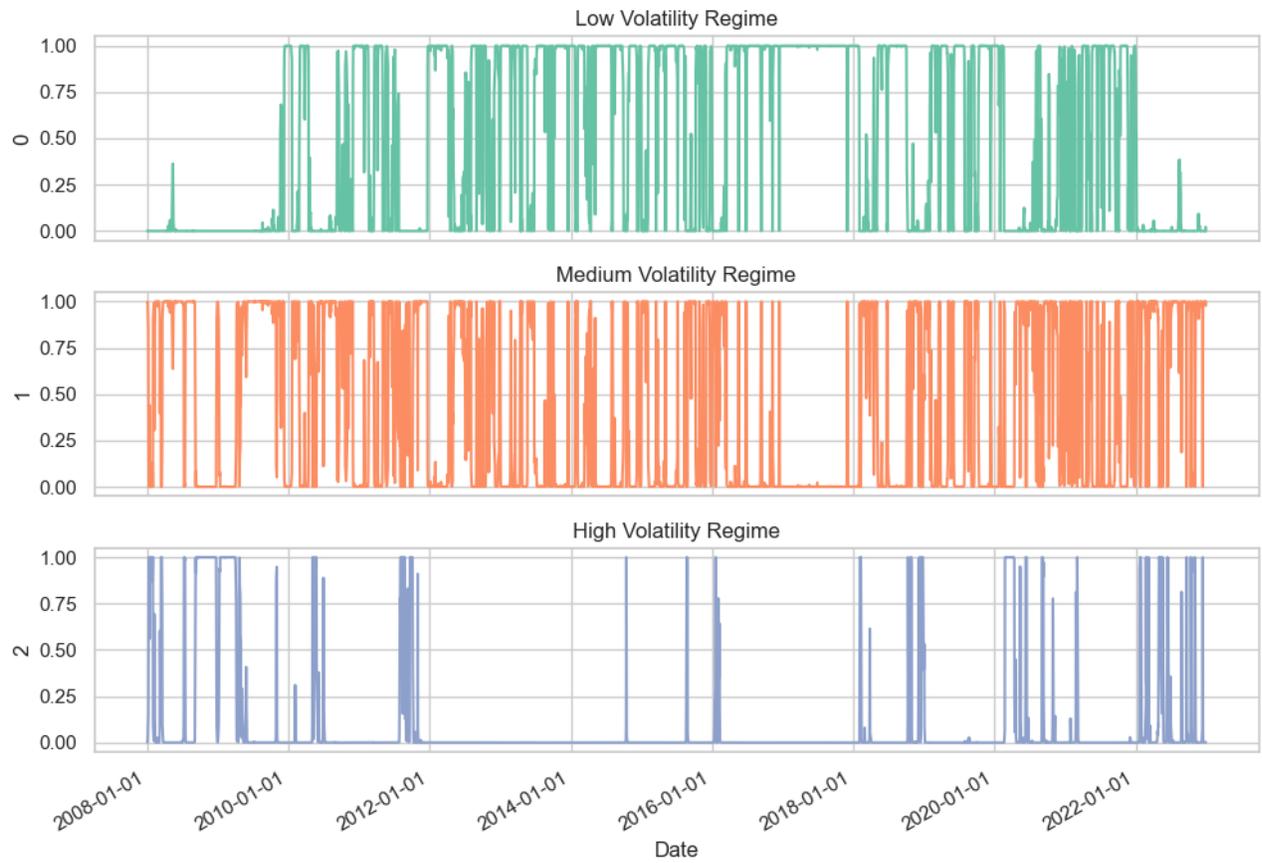
Dependent Variable	$RV_t$
No. Observations	3737
Tail shape 0 parameters	
$\mu$	$1.884 \times 10^{-5***}$
$\sigma^2$	$1.172 \times 10^{-10}$
Tail shape 1 parameters	
$\mu$	$8.23 \times 10^{-5***}$
$\sigma^2$	$2.144 \times 10^{-9}$
Tail shape 2 parameters	
$\mu$	$6.0 \times 10^{-4***}$
$\sigma^2$	$4.143 \times 10^{-7***}$

This table presents results for the 3-state Markov switching model. Stars are only intended to flag levels of significance. If a p-value is  $\leq 0.05$  it is flagged with one star (\*), if a p-value is  $\leq 0.01$  it is flagged with two stars (\*\*), subsequently, if a p-value is  $\leq 0.001$  it is flagged with three stars (\*\*\*).

set percentage is 75%. The method for decomposing volatility is a three-state first-order Markov switching regression model. Figure 4.8 shows how the  $RV_t$  is decomposed into three states, where the x-axis for each chart is the time period from January 1<sup>st</sup>, 2008 to December 31<sup>st</sup>, 2022 and the y-axis are the smoothed probabilities where classifications 0, 1, and 2 represent the low, medium and high volatility states, respectively. We test these states to also ascertain what happens to our models in the presence of volatility clustering given the prevalent volatility clustering during these periods. Such periods are characterised by elevated market volatility, triggered by significant market events, thereby facilitating the emergence of volatility clustering phenomena.

The results presented in Table 4.12 indicate that the average  $RV_t$  in the low volatility state is  $1.884 \times 10^{-5}$ , in the medium volatility state the  $RV_t$  is  $8.23 \times 10^{-5}$ , and in the high volatility state is  $6.0 \times 10^{-4}$ .

Figure 4.8: Markov Switching Model Volatility States



This illustration shows three charts of the volatility decomposition over the period from January 1<sup>st</sup>, 2008 to December 31<sup>st</sup>, 2022.

We map the identified volatility states to the event days and proceed with forecasting using the LSTM models. Specifically, we generate the one-day-ahead forecast corresponding to or immediately following the event date. To evaluate the forecasting accuracy, we calculate the Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) using the last five days of data, including the event day.

We observe some variability in which model yields the lowest MAE and RMSE results across low, medium, and high volatility states. However, it is evident that rough-based LSTM models consistently emerge as the preferred option for next-day forecasts across all volatility states. Tables 4.13, 4.14, and 4.15 demonstrate a decrease in loss function values when the input variables include the Hurst exponent. This suggests that rough LSTM models maintain their predictive efficacy across different volatility states in contrast to non-rough LSTM models. Notably, Table 4.15 reveals that errors in this high volatility state are often ten times greater than those observed in Tables 4.13 and 4.14. In Table 4.13, 13 out of 16, and 12 out of 16 instances of the lowest MAE and RMSE values are achieved by rough-GARCH type LSTM models, with rGELSTM and rEGARCH-LSTM models recording the highest frequency of the lowest MAEs and RMSEs. Similarly, in Table 4.14, 11 out of 13 of the lowest MAEs and 12 out of 13 of the lowest RMSEs are associated with rough-GARCH type LSTM models, with rEGARCH-LSTM models leading in the number of instances with the lowest errors. In the case of Table 4.15 all 6 instances of the lowest MAEs and RMSEs are observed with rough-GARCH type LSTM models, with rEGARCH-LSTM models consistently achieving the lowest error metrics.

Table 4.13: Low Volatility State Events

Event_Date	Event	Event Type	GARCH-LSTM		EGARCH-LSTM		rGARCH-LSTM		rEGARCH-LSTM		rGE-LSTM	
			MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
2011-01-14	Ben-Ali leaves Tunisia	Political	1.21	1.54	3.93	4.06	2.29	2.38	0.86	1.16	<b>0.72</b>	<b>0.90</b>
2011-02-11	Mubarak ousted in Egypt	Political	2.33	2.56	2.72	3.00	2.40	2.52	<b>1.06</b>	<b>1.27</b>	1.66	1.82
2012-02-21	2nd Economic Adjustment Programme for Greece	Financial	1.42	1.49	3.50	3.55	2.27	2.44	1.09	1.48	<b>0.66</b>	<b>0.74</b>
2012-03-08	Debt restructuring in Greece	Financial	1.01	1.22	3.18	3.33	1.13	1.20	<b>0.86</b>	<b>1.007</b>	2.45	2.81
2012-03-30	Tuareg offensive in Mali after coup starts	Political	0.59	0.62	3.15	3.20	1.60	1.72	<b>0.33</b>	<b>0.40</b>	1.29	1.41
2012-09-12	German Court approves ESM	Financial	<b>0.36</b>	<b>0.43</b>	2.38	2.44	0.82	1.06	1.16	1.24	1.60	1.92
2013-10-01	US government shutdown	Financial	<b>0.41</b>	<b>0.48</b>	3.31	3.32	2.10	2.11	0.49	0.53	2.14	2.19
2013-10-09	Yellen nominated chairwoman of the FED	Financial	0.92	1.21	2.27	2.29	<b>0.67</b>	<b>0.80</b>	0.96	1.32	3.48	3.59
2013-11-21	Ukraine rejects EU association agreement	Political	2.42	2.5	3.89	4.02	2.65	2.81	1.88	2.13	<b>1.07</b>	<b>1.20</b>
2014-02-07	German Constitutional Court consults EU Court on OMT	Financial	2.14	2.27	1.99	2.38	<b>1.44</b>	<b>1.51</b>	1.98	2.06	2.96	3.08
2014-02-18	Maidan riots in Ukraine	Political	0.82	0.93	2.71	2.72	1.20	1.26	<b>0.28</b>	<b>0.29</b>	1.87	1.88
2014-02-27	First Russian soldiers in Crimea	Political	2.13	2.11	3.41	3.51	2.04	2.16	2.29	2.43	<b>0.53</b>	<b>0.68</b>
2014-06-10	IS seizes Mosul	Political	3.12	3.14	3.49	3.49	1.89	1.93	1.94	2.04	<b>0.26</b>	<b>0.32</b>
2014-07-08	Israel-Gaza conflict	Political	2.61	2.69	2.88	2.88	1.68	1.71	<b>0.75</b>	<b>0.87</b>	1.52	1.59
2015-06-12	Chinese stock market crash	Financial	<b>0.84</b>	<b>0.93</b>	2.91	2.99	2.01	2.28	1.01	1.17	1.53	2.01
2015-06-30	Greece defaulting on IMF loan	Financial	1.3	<b>1.34</b>	2.95	3.05	<b>1.29</b>	1.49	1.39	1.59	3.28	3.51

This Table shows the MAE and RMSE results of low volatility state one-day step forecasts, where we use the previous 5 days of 1-step ahead forecasts up to the event day to calculate the MAE and RMSE. We use a window length is 22, and the training set percentage is 75%. We set configurations for the LSTM model which include; 1 layer, 100 epochs, and 0.001 learning rate. Errors are calculated by MAE and RMSE have been adjusted with a  $1 \times 10^5$  multiple for readability.

Table 4.14: Medium Volatility State Events

Event_Date	Event	Event Type	GARCH-LSTM		EGARCH-LSTM		rGARCH-LSTM		rGELSTM			
			MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE		
2009-06-15	Iran Green Revolution	Political	8.61	8.72	8.61	8.73	8.54	8.65	5.32	5.51	<b>3.81</b>	<b>4.10</b>
2010-04-27	Downgrading of Greece + Portugal	Financial	3.42	3.41	10.1	10.2	4.13	4.11	3.71	3.72	<b>2.32</b>	<b>2.51</b>
2010-05-03	1st Economic Adjustment Programme for Greece	Financial	4.62	5.21	4.50	5.42	5.33	5.72	<b>3.40</b>	<b>5.11</b>	7.43	7.51
2010-07-21	Dodd-Frank-Act signed by Obama	Financial	2.32	3.51	<b>2.01</b>	3.11	3.52	3.61	2.43	<b>2.51</b>	2.42	2.81
2011-03-07	UN Security Council establishes no-fly zone in Libya	Political	2.21	2.72	4.03	4.01	2.32	2.71	<b>0.99</b>	1.23	1.12	<b>1.10</b>
2011-03-11	Fukushima evacuation order	Political	2.52	2.71	3.81	3.92	3.71	3.82	<b>0.98</b>	<b>1.41</b>	1.91	2.20
2011-09-06	Swiss frank pegged	Financial	9.91	11.2	9.12	11.1	10.1	11.1	<b>6.81</b>	<b>9.32</b>	9.21	10.1
2011-11-09	Berlusconi resignation announced	Political	2.83	3.11	5.11	7.12	2.42	2.71	<b>2.01</b>	<b>2.12</b>	2.41	2.77
2012-06-29	EU Banking Union	Financial	2.01	2.41	<b>1.50</b>	<b>2.01</b>	1.94	2.52	3.51	3.82	4.07	4.57
2012-11-07	Obama re-elected	Political	1.32	1.53	2.81	2.91	1.92	2.01	<b>0.63</b>	<b>0.77</b>	1.76	2.09
2013-04-15	Boston marathon bombing	Political	2.02	2.12	2.13	2.32	1.94	2.21	<b>0.71</b>	<b>0.79</b>	2.34	2.45
2015-01-07	Charlie Hebdo attack	Political	2.61	2.92	6.43	6.81	4.80	5.42	3.11	3.32	<b>2.51</b>	<b>2.80</b>
2015-01-15	Swiss National Bank abolishes currency peg	Financial	3.94	4.51	5.11	5.92	3.91	4.52	3.01	3.81	<b>2.32</b>	<b>2.71</b>

This Table shows the MAE and RMSE results of medium volatility state one-day step forecasts, where we use the previous 5 days of 1-step ahead forecasts up to the event day to calculate the MAE and RMSE. We use a window length is 22, and the training set percentage is 75%. We set configurations for the LSTM model which include; 1 layer, 100 epochs, and 0.001 learning rate. Errors are calculated by MAE and RMSE have been adjusted with a  $1 \times 10^5$  multiple for readability.

Table 4.15: High Volatility State Events

Date	Event	Event Type	GARCH-LSTM		EGARCH-LSTM		rGARCH-LSTM		rEGARCH-LSTM		rGELSTM	
			MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
2008-03-14	Takeover of Bear Stearns by JP Morgan approved	Financial	6.77	8.29	7.36	7.95	5.84	7.63	6.42	8.25	<b>3.31</b>	<b>3.62</b>
2008-10-01	Emergency Economic Stabilisation Act, US bails out banks	Financial	13.7	15.4	12.1	13.8	13.0	16.2	<b>5.17</b>	<b>5.51</b>	11.2	12.3
2008-10-03	TARP passes at the congress	Financial	18.7	21.9	14.6	16.4	14.7	17.5	<b>7.47</b>	<b>8.25</b>	8.37	9.44
2008-11-04	Obama elected	Political	20.1	21.2	9.54	12.4	15.7	18.4	<b>7.62</b>	<b>8.45</b>	9.54	9.63
2009-02-13	US Recovery and Reinvestment Act passes Congress	Financial	7.14	10.4	6.93	7.84	7.45	8.76	<b>5.47</b>	<b>5.99</b>	11.3	16.1
2010-05-10	EFSSF adopted	Financial	41.6	53.4	37.6	52.4	<b>11.8</b>	<b>14.9</b>	24.5	31.2	19.5	23.4

This Table shows the MAE and RMSE results of high volatility state one-day step forecasts, where we use the previous 5 days of 1-step ahead forecasts up to the event day to calculate the MAE and RMSE. We use a window length is 22, and the training set percentage is 75%. We set configurations for the LSTM model which include; 1 layer, 100 epochs, and 0.001 learning rate. Errors are calculated by MAE and RMSE have been adjusted with a  $1 \times 10^5$  multiple for readability.

## 4.6 Conclusion

This Chapter presents the development and evaluative analysis of rough GARCH-type-LSTM models (rGE, rGARCH, and rEGARCH), substantiating the efficacy of rGE-LSTM and rEGARCH as comprehensive predictive tools for forecasting realised volatility under macroeconomic uncertainty. The investigation encompasses a spectrum of modelling techniques, starting with the quantification of interest rates (DGS3MO) and the roughness in high-frequency financial data via the Hurst exponent, a metric indicative of long-term memory attributes. Subsequently, an exploration into GARCH-type models is conducted, where optimal parameters are deduced using the Maximum Likelihood Estimation method. Comparative model assessment is facilitated through the application of the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC), ensuring the selection of models with optimal fit.

The analysis progresses with the implementation of the Long Short-Term Memory (LSTM) model, a variant of recurrent neural networks designed to adeptly capture and learn from the long-term dependencies inherent in the volatility patterns delineated by GARCH-type models and the Hurst exponent features. Our comparative analysis includes GARCH-type LSTM models. The results indicate that in most cases the rGE-LSTM model predicts more accurately and exhibits the lowest forecasting errors in relation to loss functions.

Furthermore, through the deployment of the Wilcoxon Signed-Rank Test (WS) and the Diebold-Mariano (DM) Test, a comparative analysis between all forecasted outcomes of the

GARCH-type LSTM models is undertaken. which indicate high statistical significance with rough-based GARCH-type LSTM models. Our empirical evidence suggests that incorporating a rolling Hurst exponent built on 5-minute data, and combining GARCH-type models that capture volatility clustering, asymmetry, and the leverage effect optimally harnesses the predictive capabilities of forecasting  $RV_t$  using the rGE-LSTM model.

In the context of the macroeconomic events dataset, an evaluation of various loss functions are conducted on three volatility states generated from Markov switching Models. We evaluate low, medium, and high volatility states. The analysis confirms rough GARCH-type LSTM models prove to outperform non-rough LSTM types. This comparative analysis reveals an increase in forecasting errors associated with macroeconomic events, thereby lending empirical support to the phenomenon of volatility clustering. This increase implies that high volatility affects and increases forecasting errors.

The implications of these findings are manifold, particularly accentuating the model's utility for risk managers, hedgers, and monetary policymakers. The incorporation of the Hurst exponent in rough GARCH-type-LSTM models not only enhances the predictive accuracy of volatility but also provides deeper insights into the long-term memory and persistence of financial time series. This inclusion is especially valuable for traders and portfolio managers, as it helps in better understanding the market's underlying dynamics, enabling more informed decision-making in terms of asset allocation, risk assessment, and hedging strategies. Hedgers can benefit from more precise estimates of future volatility, which are crucial for pricing derivatives and other hedging instruments. Monetary policymakers can leverage these models to anticipate and mitigate systemic risks, as the enhanced understanding of volatility dynamics

could inform interest rate decisions and macro-prudential policies. In terms of computational complexity, the inclusion of the Hurst exponent does introduce additional computational load and complexity. Calculating the Hurst exponent involves more sophisticated mathematical procedures, which can increase the overall computational time and resource requirements. However, the benefits often outweigh these drawbacks. The use of the Hurst exponent allows for capturing more intricate features of the data, such as fractional Brownian motion, which traditional models may overlook. This leads to more nuanced and potentially more accurate forecasts, which can significantly enhance decision-making processes. The most computationally intensive part of the function is the calculation of R/S values over all sub-periods, which has a complexity of  $O(M^2)$ , where  $M$  is the length of the input time-series.<sup>3</sup> Despite these advantages, there are also some limitations. While the Hurst exponent provides valuable insights into the persistence of volatility, it may not fully account for all market anomalies or structural breaks. Overall, the inclusion of the Hurst exponent in rough GARCH-type-LSTM models offers a sophisticated approach to volatility forecasting that can significantly benefit traders, portfolio managers, risk managers, and policymakers.

---

<sup>3</sup>To calculate the length of the original series, we retrieve the length which is  $O(1)$ . Next, we transform the series into log returns. Calculating log returns involves subtraction, this occurs over the entire series of length  $M$ . Therefore, the complexity is  $O(M)$ . Thereafter, we initialise of lists for storing R/S values and sub-period lengths. These are simple array initialisations with  $O(1)$  complexity. Then we iterate over sub-period lengths and calculate R/S values: The outer while loop runs while  $n \leq (M - 1)/2$ . The number of iterations is approximately  $M/2$  so the outer loop runs  $O(M)$  times. Within the loop, the data is divided into sub-periods. For each  $n$ , there are approximately  $(M - 1)/n$  sub-periods which calculate deviations, range, and standard deviation which involves linear operations over each sub-period of length  $n$ . Therefore, the complexity of this step can be described as:  $O(\sum_{n=1}^{M/2} (M - 1)) = O(M \cdot M/2) = O(M^2)$

## Chapter 5

### CONCLUSION

This research focuses on the analysis of tail risk and volatility by examining how financial markets respond to various economic and geopolitical events. We evaluate option-implied Probability Density Functions (PDFs), a volatility-managed portfolio approach with a particular emphasis on Extreme Value Theory (EVT), and forecasting realised volatility.

Chapter 2 investigates the evolution of option-implied PDFs during periods of macroeconomic uncertainty, offering insights into market expectations and risk assessments. Option-implied PDFs, which reflect risk-neutral probabilities of future asset prices, are derived from market-traded options and provide an essential lens through which investor perceptions of risk can be analysed. These PDFs, forward-looking by nature, offer a crucial perspective on how market sentiment changes in anticipation of, and in response to, economic events.

Parametric and non-parametric methods are used to estimate these PDFs, with the mixture of log-normals model emerging as the most accurate approach for forecasting European call option prices, as confirmed by the Diebold-Mariano test Diebold & Mariano (1995). The

relationship between ex-ante uncertainty and the resolution of uncertainty post-event is explored using a Structural Vector Autoregressive (SVAR) model on macroeconomic news events. Additionally, the Extreme Value Theory (EVT) Tail Loss Measure (TLM) and Probit regression models are employed to examine the predictive power of low volatility in forecasting financial crises. We find that a low volatility period of four and five years translates to a 19.455%, and 23.052% increase in the probability of market uncertainty. The analysis concludes that periods of low volatility are significant precursors to financial crises, aligning with theories posited by Keynes (1937) and Minsky (1977), which suggest that reduced perception of risk or volatility encourages excessive risk-taking which leads to financial crisis.

Chapter 3 shifts focus to EVT in volatility-managed portfolio optimisation, with a specific emphasis on managing tail risk, downside risk, and enhancing the Sharpe ratio. This Chapter constructs a portfolio based on the Fama-French Five-Factor Model, extending the traditional three-factor framework by incorporating profitability and investment factors. Extreme value theory (EVT) plays a central role in modelling the volatility-managed portfolio and managing tail risk, as traditional Value-at-Risk (VaR) measures often underestimate extreme losses.

Our research demonstrates that EVT-managed (EVTM) portfolios deliver superior performance across a range of volatility-managed strategies, including the Sharpe and Sortino ratios. Sortino-optimised EVTm portfolios perform particularly well, achieving a Sortino-optimised Sharpe ratio 3.460 times higher than that of volatility-managed (VM) portfolios from Moreira & Muir (2017), underscoring the efficacy of incorporating downside risk into portfolio management. The analysis of Conditional Value-at-Risk (CVaR) metrics confirms that EVTm portfolios manage extreme downside risks more effectively than VM portfolios, especially

at the 99% and 95% confidence levels. These results are particularly pertinent to portfolio managers and risk professionals as they provide enhanced methods for handling the risks associated with extreme market movements. Positive skewness and lower kurtosis in EVTm portfolios indicate a more favourable risk-return profile compared to VM portfolios.

Break-even transaction cost analysis reveals that while EVTm portfolios incur higher transaction costs than their VM counterparts, they consistently exhibit superior risk-adjusted returns. The findings align with the principles of Prospect Theory (Kai-Ineman & Tversky (1979)), which highlights investors' tendency to become more risk-averse during periods of loss and more risk-seeking during periods of gain. By dynamically adjusting allocations in response to market conditions, EVTm portfolios offer a robust approach to balancing risk and reward, enhancing portfolio resilience across varying market environments.

Chapter 4 focuses on volatility forecasting with roughness, which is crucial for anticipating financial market stability and investor confidence. The study emphasises the Hurst exponent, GARCH-type models, and LSTM networks for their respective abilities to capture long-term dependencies, volatility clustering, and complex non-linear relationships in time-series data. The rough GARCH-type LSTM model, which integrates the roughness of financial data with GARCH and EGARCH models, is shown to be particularly effective in forecasting realised volatility. Diebold-Mariano (DM) and Wilcoxon Signed-Rank (WS) tests confirm that rough GARCH-type LSTM models outperform traditional models in terms of predictive accuracy.

Volatility states generated by a three-state Markov switching regression model are mapped to financial and political events, allowing for the computation of one day-ahead and  $n$  day-ahead

forecasts. These forecasts are evaluated using Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE), with the rough GARCH-type LSTM model demonstrating superior performance in predicting volatility around major events.

Accurate volatility forecasts enable effective risk management, portfolio optimisation, and policy formulation. This research evaluates various models, including GARCH-type models and machine learning techniques such as Long Short-Term Memory (LSTM) networks, for forecasting realised volatility. While the inclusion of the Hurst exponent increases computational complexity, the benefits in terms of improved forecasting accuracy and deeper insights into volatility dynamics outweigh the costs. This modelling approach provides risk managers with more precise estimates of future volatility, allowing for better pricing of derivatives and hedging strategies, and offering policymakers a tool to forecast volatility.

In conclusion, this research contributes to the fields of volatility forecasting, portfolio optimisation, and tail-risk management. By combining option-implied PDFs, EVT-managed portfolios, and rough GARCH-type LSTM models, this thesis offers novel insights into the dynamics of financial markets under conditions of uncertainty. The findings underscore the importance of incorporating advanced econometric and machine learning techniques into financial risk management strategies to enhance predictive accuracy and portfolio resilience in the face of extreme market events. These insights will be particularly valuable to risk managers, portfolio managers, and policymakers seeking to navigate the challenges of modern financial markets.

## Bibliography

- Abdulali, Bashir Ahmed Albashir et al., 2022. Extreme Value Distributions: An Overview of Estimation and Simulation. *Journal of Probability and Statistics*, vol. 2022,
- Allen, David E & Abhay K Singh & Robert J Powell, 2013. EVT and tail-risk modelling: Evidence from market indices and volatility series. *The North American Journal of Economics and Finance*, vol. 26, pages 355–369.
- Andersen, Torben G. & Tim Bollerslev, 1997. Intraday periodicity and volatility persistence in financial markets. *Journal of Empirical Finance*, vol. 4, pages 115–158.
- Angelidis, Timotheos & Nikolaos Tassaromatis, 2023. The disappearing profitability of volatility-managed equity factors. *Journal of Financial Markets*, vol. 65, pages 100857.
- Bahra, Bhupinder, 1997. Implied risk-neutral probability density functions from option prices: theory and application.
- Baker, Scott R & Nicholas Bloom, 2013. Does uncertainty reduce growth. Using disasters as natural experiments.
- Balkema, A. A. & L. de Haan, 1974. Residual Life Time at Great Age. *The Annals of Probability*, vol. 2, pages 792–804.
- Banz, Rolf W. & Merton H. Miller, 1978. Prices for State-Contingent Claims: Some Estimates and Applications. *The Journal of Business*, pages 653–672.
- Baran, Jaroslav & Jiri Witzany, 2011. A comparison of EVT and standard VaR estimations.
- Bariviera, Aurelio F, 2017. The inefficiency of Bitcoin revisited: A dynamic approach. *Economics Letters*, vol. 161, pages 1–4.
- Barndorff-Nielsen, Ole E & Neil Shephard, 2002. Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 64, pages 253–280.
- Barroso, Pedro & Andrew Detzel, 2021. Do limits to arbitrage explain the benefits of volatility-managed portfolios? *Journal of Financial Economics*, vol. 140, pages 744–767.

- Barunik, Jozef & Ladislav Kristoufek, 2010. On Hurst exponent estimation under heavy-tailed distributions. *Physica A: statistical mechanics and its applications*, vol. 389, pages 3844–3855.
- Bengio, Yoshua & Aaron Courville & Pascal Vincent, 2013. Representation learning: A review and new perspectives. *IEEE transactions on pattern analysis and machine intelligence*, vol. 35, pages 1798–1828.
- Bhattacharya, Sudipto et al., 2015. A reconsideration of Minsky’s financial instability hypothesis. *Journal of Money, Credit and Banking*, vol. 47, pages 931–973.
- Black, Fischer & Myron Scholes, 1973. The pricing of options and corporate liabilities. *Journal of political economy*, vol. 81, pages 637–654.
- Bloom, Nicholas, 2009. The impact of uncertainty shocks. *econometrica*, vol. 77, pages 623–685.
- Bollerslev, Tim, 1987. A conditionally heteroskedastic time series model for speculative prices and rates of return. *The review of economics and statistics*, pages 542–547.
- Bollerslev, Tim & Sophia Zhengzi Li & Bingzhi Zhao, 2020. Good Volatility, Bad Volatility, and the Cross Section of Stock Returns. *Journal of Financial and Quantitative Analysis*, vol. 55, pages 751–781.
- Bookstaber, Richard M & James B McDonald, 1987. A general distribution for describing security price returns. *Journal of business*, pages 401–424.
- Božović, Miloš, 2024. VIX-managed portfolios. *International Review of Financial Analysis*, vol. 95, pages 103353.
- Breeden, Douglas T & Robert H Litzenberger, 1978. Prices of state-contingent claims implicit in option prices. *Journal of business*, pages 621–651.
- Brown, Robert, 1828. XXVII. A brief account of microscopical observations made in the months of June, July and August 1827, on the particles contained in the pollen of plants; and on the general existence of active molecules in organic and inorganic bodies. *The philosophical magazine*, vol. 4, pages 161–173.
- Brunnermeier, Markus K. & Yuliy Sannikov, 2014. A Macroeconomic Model with a Financial Sector. *American Economic Review*, vol. 104, pages 379–421.
- Bu, Ruijun & Kaddour Hadri, 2007. Estimating option implied risk-neutral densities using spline and hypergeometric functions. *The Econometrics Journal*, vol. 10, pages 216–244.
- Caporale, Guglielmo Maria & Luis Gil-Alana & Alex Plastun, 2018. Persistence in the cryptocurrency market. *Research in International Business and Finance*, vol. 46, pages 141–148.

- Caporale, Guglielmo Maria & Luis Gil-Alana & Alex Plastun, 2019. Long memory and data frequency in financial markets. *Journal of Statistical Computation and Simulation*, vol. 89, pages 1763–1779.
- Caporale, Guglielmo Maria & Luis A Gil-Alana, 2013. Long memory and fractional integration in high frequency data on the US dollar/British pound spot exchange rate. *International review of financial analysis*, vol. 29, pages 1–9.
- Caporale, Guglielmo Maria & Alex Plastun, 2024. Persistence in high frequency financial data: the case of the EuroStoxx 50 futures prices. *Cogent Economics & Finance*, vol. 12, pages 2302639.
- Cavalcante, Rodolfo C et al., 2016. Computational intelligence and financial markets: A survey and future directions. *Expert Systems with Applications*, vol. 55, pages 194–211.
- Cederburg, Scott et al., 2020. On the performance of volatility-managed portfolios. *Journal of financial Economics*, vol. 138, pages 95–117.
- Chen, Hao & Qiulan Wan & Yurong Wang, 2014. Refined Diebold-Mariano test methods for the evaluation of wind power forecasting models. *Energies*, vol. 7, pages 4185–4198.
- Corsi, Fulvio, 2009. A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics*, vol. 7, pages 174–196.
- Cotter, John, 2005. Uncovering long memory in high frequency UK futures. *The European Journal of Finance*, vol. 11, pages 325–337.
- Cox, John C & Stephen A Ross, 1976. The valuation of options for alternative stochastic processes. *Journal of financial economics*, vol. 3, pages 145–166.
- Da Silva, Sergio et al., 2007. Hurst exponents, power laws, and efficiency in the Brazilian foreign exchange market. *Economics Bulletin*, vol. 7, pages 1–11.
- Danielsson, Jon & Marcela Valenzuela & Ilknur Zer, 2018. Learning from history: Volatility and financial crises. *The Review of Financial Studies*, vol. 31, pages 2774–2805.
- Danielsson, Jon & Casper G de Vries, 1997. *Extreme returns, tail estimation, and value-at-risk*, vol. 11,
- Darrat, Ali F & Shafiqur Rahman & Maosen Zhong, 2003. Intraday trading volume and return volatility of the DJIA stocks: A note. *Journal of Banking & Finance*, vol. 27, pages 2035–2043.
- De Nicolo, Gianni, 2023. Tail Risk-Managed Portfolios.
- DeMiguel, Victor & Alberto Martin-Utrera & Raman Uppal, 2021. A multifactor perspective on volatility-managed portfolios.
- Deng, Chunliang et al., 2020. Garch model test using high-frequency data. *Mathematics*, vol. 8, pages 1922.

- Diebold, Francis & Roberto Mariano, 1995. Comparing Predictive Accuracy. *Journal of Business and Economic Statistics*, vol. 13, pages 253–63.
- Dobrev, Dobrislav & Pawel Szerszen, 2010. The information content of high-frequency data for estimating equity return models and forecasting risk.
- Engle, Robert F, 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica: Journal of the econometric society*, vol. 50, pages 987–1007.
- Engle, Robert F & Andrew J Patton, 2007. What good is a volatility model? *Forecasting volatility in the financial markets*, pages 47–63.
- Fama, Eugene F, 1965. The behavior of stock-market prices. *The journal of Business*, vol. 38, pages 34–105.
- Fama, Eugene F & Kenneth R French, 1992. The cross-section of expected stock returns. *the Journal of Finance*, vol. 47, pages 427–465.
- Fama, Eugene F & Kenneth R French, 1996. Multifactor explanations of asset pricing anomalies. *The journal of finance*, vol. 51, pages 55–84.
- Fama, Eugene F & Kenneth R French, 2015. A five-factor asset pricing model. *Journal of financial economics*, vol. 116, pages 1–22.
- Fernandez, Viviana, 2003. Extreme value theory and value at risk. *Revista de Análisis Económico*, vol. 18,
- Fisher, Ronald Aylmer & Leonard Henry Caleb Tippett, 1928. Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Mathematical proceedings of the Cambridge philosophical society*, vol. 24, pages 180–190.
- Fréchet, Maurice, 1927. Sur la loi de probabilité de l'écart maximum. *Ann. de la Soc. Polonaise de Math.*
- Garcia-Medina, Andres & Ester Aguayo-Moreno, 2024. LSTM-GARCH hybrid model for the prediction of volatility in cryptocurrency portfolios. *Computational Economics*, vol. 63, pages 1511–1542.
- Garman, Mark B & Steven W Kohlhagen, 1983. Foreign currency option values. *Journal of international Money and Finance*, vol. 2, pages 231–237.
- Gatheral, Jim & Thibault Jaisson & Mathieu Rosenbaum, 2018. Volatility is rough. *Quantitative finance*, vol. 18, pages 933–949.
- Gemmell, Gordon & Apostolos Saffekos, 2000. How useful are implied distributions?: Evidence from stock index options. *The Journal of Derivatives*, vol. 7, pages 83–91.
- Gençay, Ramazan & Faruk Selçuk & Abdurrahman Ulugülyağci, 2003. High volatility, thick tails and extreme value theory in value-at-risk estimation. *Insurance: Mathematics and Economics*, vol. 33, pages 337–356.

- Ghosh, Souvik & Sidney Resnick, 2010. A discussion on mean excess plots. *Stochastic Processes and their Applications*, vol. 120, pages 1492–1517.
- Gnedenko, Boris, 1943. Sur la distribution limite du terme maximum d'une serie aleatoire. *Annals of mathematics*, vol. 44, pages 423–453.
- Goetzmann, William et al., 2007. Portfolio performance manipulation and manipulation-proof performance measures. *The Review of Financial Studies*, vol. 20, pages 1503–1546.
- Grech, Dariusz & Zygmunt Mazur, 2005. Statistical properties of old and new techniques in detrended analysis of time series. *Acta Physica Polonica Series B*, vol. 36, pages 2403.
- Hajizadeh, Ehsan et al., 2012. A hybrid modeling approach for forecasting the volatility of S&P 500 index return. *Expert Systems with Applications*, vol. 39, pages 431–436.
- Hamidieh, Kam, 2014. Estimating the tail shape parameter from option prices.
- Hochreiter, Sepp & Jürgen Schmidhuber, 1997. Long short-term memory. *Neural computation*, vol. 9, pages 1735–1780.
- Hodrick, Robert J & Edward C Prescott, 1997. Postwar US business cycles: an empirical investigation. *Journal of Money, credit, and Banking*, pages 1–16.
- Hu, Yan & Jian Ni & Liu Wen, 2020. A hybrid deep learning approach by integrating LSTM-ANN networks with GARCH model for copper price volatility prediction. *Physica A: Statistical Mechanics and its Applications*, vol. 557, pages 124907.
- Hurst, Harold Edwin, 1951. Long-term storage capacity of reservoirs. *Transactions of the American society of civil engineers*, vol. 116, pages 770–799.
- Jackwerth, Jens Carsten, 1999. Option-implied risk-neutral distributions and implied binomial trees: A literature review. *The Journal of Derivatives*, vol. 7, pages 66–82.
- Jackwerth, Jens Carsten, 2000. Recovering risk aversion from option prices and realized returns. *The Review of Financial Studies*, vol. 13, pages 433–451.
- Jarrow, Robert & Andrew Rudd, 1982. Approximate option valuation for arbitrary stochastic processes. *Journal of financial Economics*, vol. 10, pages 347–369.
- Jondeau, Eric & Ser-Huang Poon & Michael Rockinger, 2007. *Financial modeling under non-Gaussian distributions*,
- Kai-Ineman, DANIEL & Amos Tversky, 1979. Prospect theory: An analysis of decision under risk. *Econometrica*, vol. 47, pages 363–391.
- Kambouroudis, Dimos S & David G McMillan & Katerina Tsakou, 2016. Forecasting stock return volatility: A comparison of GARCH, implied volatility, and realized volatility models. *Journal of futures markets*, vol. 36, pages 1127–1163.

- Kantelhardt, Jan W et al., 2002. Multifractal detrended fluctuation analysis of nonstationary time series. *Physica A: Statistical Mechanics and its Applications*, vol. 316, pages 87–114.
- Ke, Huang et al., 2023. Predicting chinese commodity futures price: An eemd-hurst-lstm hybrid approach. *IEEE Access*, vol. 11, pages 14841–14858.
- Keynes, J. M., 1937. The General Theory of Employment. *The Quarterly Journal of Economics*, vol. 51, pages 209–223.
- Kim, Ha Young & Chang Hyun Won, 2018. Forecasting the volatility of stock price index: A hybrid model integrating LSTM with multiple GARCH-type models. *Expert Systems with Applications*, vol. 103, pages 25–37.
- Kingma, Diederik P & Jimmy Ba, 2014. Adam: A method for stochastic optimization.
- Kışınbay, Turgut, 2010. Predictive ability of asymmetric volatility models at medium-term horizons. *Applied Economics*, vol. 42, pages 3813–3829.
- Koopman, Siem Jan & Borus Jungbacker & Eugenie Hol, 2005. Forecasting daily variability of the S&P 100 stock index using historical, realised and implied volatility measurements. *Journal of Empirical Finance*, vol. 12, pages 445–475.
- Kristjanpoller, Werner & Anton Fadic & Marcel C Minutolo, 2014. Volatility forecast using hybrid neural network models. *Expert Systems with Applications*, vol. 41, pages 2437–2442.
- Kristjanpoller, Werner & Esteban Hernández, 2017. Volatility of main metals forecasted by a hybrid ANN-GARCH model with regressors. *Expert Systems with Applications*, vol. 84, pages 290–300.
- Kristjanpoller, Werner & Marcel C Minutolo, 2016. Forecasting volatility of oil price using an artificial neural network-GARCH model. *Expert Systems with Applications*, vol. 65, pages 233–241.
- Lago, Jesus et al., 2021. Forecasting day-ahead electricity prices: A review of state-of-the-art algorithms, best practices and an open-access benchmark. *Applied Energy*, vol. 293, pages 116983.
- Liu, Xiaoquan et al., 2007. Closed-form transformations from risk-neutral to real-world distributions. *Journal of Banking & Finance*, vol. 31, pages 1501–1520.
- Lo, Andrew W, 1991. Long-term memory in stock market prices. *Econometrica: Journal of the Econometric Society*, pages 1279–1313.
- Lo, Andrew W & A Craig MacKinlay, 1988. Stock market prices do not follow random walks: Evidence from a simple specification test. *The review of financial studies*, vol. 1, pages 41–66.

- Longin, Francois M, 2000. From value at risk to stress testing: The extreme value approach. *Journal of Banking & Finance*, vol. 24, pages 1097–1130.
- Los, Cornelis A & Rossitsa M Yalamova, 2004. Multi-fractal spectral analysis of the 1987 stock market crash.
- Lynch, Damien PG & Nikolaos Panigirtzoglou, 2008. Summary statistics of option-implied probability density functions and their properties.
- MacDonald, R & MP Taylor, 1992. Persistence in stock market returns: Estimates using a weekly data base. *Journal of Business Finance and Accounting*, vol. 19, pages 23–31.
- Mandelbrot, Benoit, 1963. New methods in statistical economics. *Journal of political economy*, vol. 71, pages 421–440.
- Mandelbrot, Benoit B & John W Van Ness, 1968. Fractional Brownian motions, fractional noises and applications. *SIAM review*, vol. 10, pages 422–437.
- Markose, Sheri & Amadeo Alentorn, 2011. The generalized extreme value distribution, implied tail index, and option pricing. *The Journal of Derivatives*, vol. 18, pages 35–60.
- Markowitz, Harry, 1952. Portfolio Selection. *The Journal of Finance*, vol. 7, pages 77–91.
- McDonald, James B & Richard M Bookstaber, 1991. Option pricing for generalized distributions. *Communications in Statistics-Theory and Methods*, vol. 20, pages 4053–4068.
- McNeil, Alexander J & Rüdiger Frey, 2000. Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach. *Journal of empirical finance*, vol. 7, pages 271–300.
- Melick, William R & Charles P Thomas, 1997. Recovering an asset's implied PDF from option prices: an application to crude oil during the Gulf crisis. *Journal of Financial and Quantitative Analysis*, vol. 32, pages 91–115.
- Mertens, Karel & Morten O. Ravn, 2013. The Dynamic Effects of Personal and Corporate Income Tax Changes in the United States. *The American Economic Review*, vol. 103, pages 1212–1247.
- Minsky, Hyman P., 1977. The Financial Instability Hypothesis: An Interpretation of Keynes and an Alternative to "Standard" Theory. *Nebraska Journal of Economics and Business*, vol. 16, pages 5–16.
- Moreira, Alan & Tyler Muir, 2017. Volatility-managed portfolios. *The Journal of Finance*, vol. 72, pages 1611–1644.
- Nelson, Daniel B, 1992. Filtering and forecasting with misspecified ARCH models I: Getting the right variance with the wrong model. *Journal of econometrics*, vol. 52, pages 61–90.

- Odening, Martin & Jan Hinrichs, 2003. Using extreme value theory to estimate value-at-risk. *Agricultural finance review*, vol. 63, pages 55–73.
- Petrozziello, Alessio et al., 2022. Deep learning for volatility forecasting in asset management. *Soft Computing*, vol. 26, pages 8553–8574.
- Pickands, James, 1975. Statistical Inference Using Extreme Order Statistics. *The Annals of Statistics*, vol. 3, pages 119–131.
- Piffer, Michele & Maximilian Podstawski, 2018. Identifying Uncertainty Shocks Using the Price of Gold. *The Economic Journal*, vol. 128, pages 3266–3284.
- Poon, Ser-Huang & Clive W J Granger, 2003. Forecasting volatility in financial markets: A review. *Journal of economic literature*, vol. 41, pages 478–539.
- Qian, Bo & Khaled Rasheed, 2004. Hurst exponent and financial market predictability. *IASTED conference on Financial Engineering and Applications*, pages 203–209.
- Qian, Edward, 2011. Risk parity and diversification. *Journal of Investing*, vol. 20, pages 119.
- Qiao, Xiao & Siboyan & Binbin Deng, 2020. Downside volatility-managed portfolios. *The Journal of Portfolio Management*, vol. 46, pages 13–29.
- Reinhart, Carmen M & Kenneth S Rogoff, 2011. From financial crash to debt crisis. *American Economic Review*, vol. 101, pages 1676–1706.
- Ritchey, Robert J, 1990. Call option valuation for discrete normal mixtures. *Journal of Financial Research*, vol. 13, pages 285–296.
- Roh, Tae Hyup, 2007. Forecasting the volatility of stock price index. *Expert Systems with Applications*, vol. 33, pages 916–922.
- Romer, Christina D & David H Romer, 2004. A new measure of monetary shocks: Derivation and implications. *American economic review*, vol. 94, pages 1055–1084.
- Ross, Stephen A, 1976. Options and efficiency. *The Quarterly Journal of Economics*, vol. 90, pages 75–89.
- Salakhutdinov, Ruslan & Geoffrey Hinton, 2009. Semantic hashing. *International Journal of Approximate Reasoning*, vol. 50, pages 969–978.
- Serletis, Apostolos & Aryeh Adam Rosenberg, 2007. The Hurst exponent in energy futures prices. *Physica A: Statistical Mechanics and its Applications*, vol. 380, pages 325–332.
- Sharpe, William F, 1964. Capital asset prices: A theory of market equilibrium under conditions of risk. *The journal of finance*, vol. 19, pages 425–442.
- Shimko, David, 1993. Bounds of probability. *Risk*, vol. 6, pages 33–37.
- Statman, Meir, 1987. How many stocks make a diversified portfolio? *Journal of financial and quantitative analysis*, vol. 22, pages 353–363.

- Stock, James H & Mark W Watson, 2012. Disentangling the Channels of the 2007-2009 Recession. *National Bureau of Economic Research*,
- Stoll, Hans R & Robert E Whaley, 1990. Stock market structure and volatility. *The Review of Financial Studies*, vol. 3, pages 37–71.
- Tseng, Chih-Hsiung et al., 2008. Artificial neural network model of the hybrid EGARCH volatility of the Taiwan stock index option prices. *Physica A: Statistical Mechanics and its Applications*, vol. 387, pages 3192–3200.
- Urom, Christian et al., 2022. Directional predictability and time-frequency spillovers among clean energy sectors and oil price uncertainty. *The Quarterly Review of Economics and Finance*, vol. 85, pages 326–341.
- Vidal, Andrés & Werner Kristjanpoller, 2020. Gold volatility prediction using a CNN-LSTM approach. *Expert Systems with Applications*, vol. 157, pages 113481.
- Vilasuso, Jon, 2002. Forecasting exchange rate volatility. *Economics Letters*, vol. 76, pages 59–64.
- Vilkov, Grigory & Yan Xiao, 2013. Option-implied information and predictability of extreme returns.
- Visser, Marcel P, 2011. GARCH parameter estimation using high-frequency data. *Journal of Financial Econometrics*, vol. 9, pages 162–197.
- Vitiello, Luiz & Ser-Huang Poon, 2014. Non-monotonic pricing kernel and an extended class of mixture of distributions for option pricing. *Review of Derivatives Research*, vol. 17, pages 241–259.
- Wang, Feifei & Xuemin Sterling Yan, 2021. Downside risk and the performance of volatility-managed portfolios. *Journal of banking & finance*, vol. 131, pages 106198.
- Weron, Rafał, 2002. Estimating long-range dependence: finite sample properties and confidence intervals. *Physica A: Statistical Mechanics and its Applications*, vol. 312, pages 285–299.
- Wilcoxon, Frank, 1947. Probability tables for individual comparisons by ranking methods. *Biometrics*, vol. 3, pages 119–122.
- Xiong, James X & Thomas M Idzorek, 2018. A Case for Tail-Risk-Based Sharpe Ratios. *Journal of Portfolio Management*, vol. 44, pages 114.
- Zakoian, Jean-Michel, 1994. Threshold heteroskedastic models. *Journal of Economic Dynamics and control*, vol. 18, pages 931–955.
- Zhou, Jian & Zhixin Kang, 2011. A comparison of alternative forecast models of REIT volatility. *The Journal of Real Estate Finance and Economics*, vol. 42, pages 275–294.

Zunino, Luciano et al., 2009. Multifractal structure in Latin-American market indices. *Chaos, Solitons & Fractals*, vol. 41, pages 2331–2340.

## Appendix A

### Chapter 2 Appendix

#### A.1 Statistical Procedures to Select the Methodology to Extract the Risk-Neutral PDFs

##### A.1.1 DM Test

In Subsections 2.4.3 and 2.4.2, the parametric and non-parametric methods are discussed mathematically. In this Appendix, we compare densities using the Diebold-Mariano (DM) test, and computational time.

We randomly selected 1,000 dates to assess the computation of density forecasts. The selected dates cover a wide range of expiration dates, capturing the variability across different option time horizons. We consider expiration dates of approximately 14, 30, 90, 180, and 365 days, resulting in roughly 5,000 density forecasts. We use these dates to compare density forecasts with actual call option prices from the selected business day, for a one-day-ahead forecasting horizon ( $h = 1$ ).

Table A.1 presents the percentages of instances where forecast errors from  $\hat{y}_{1t}$  are compared with forecast errors from  $\hat{y}_{2t}$  using the DM test for estimating S&P 500 call prices. The results indicate that the Mixture of Two Log-normals method performs best when compared with all other methods for forecasting S&P 500 European option call prices. Specifically,

Table A.1: Diebold-Mariano Test Applied to Shimko’s Spline Method, Edgeworth Density, Black-Scholes-Merton Model, Mixture of Two Log-normals, and GB2 Forecasts of S&P500 European Call Prices.

$\hat{y}_{1t}$	$\hat{y}_{2t}$			
	Shimko’s Spline Method	GB2	Edgeworth Density	Mixture of Two Log-normals
Shimko’s Spline Method		88.54%	44.27%	31.42%
GB2	14.24%		94.79%	44.77%
Edgeworth Density	55.73%	5.201%		06.42%
Mixture of Two Log-normals	68.58%	55.23%	93.58%	

We compute the DM test to see whether  $\hat{y}_{1t}$  is more accurate than  $\hat{y}_{2t}$  in predicting  $y_t$ , where  $y_t$  is the S&P500 European call prices. We have over 100 observations on average and a forecasting horizon of 1.

the Mixture of Two Log-normals method yields a better density forecast in 68.58% of cases when compared with Shimko’s Spline method, 55.23% of cases when compared with the GB2 method, and 93.58% of cases when compared with the Edgeworth method. This implies that when the Mixture of Two Log-normals method is compared to the other methods, the DM test produces a negative test statistic for 68.58%, 55.23%, and 93.58% of the cases, respectively.

### A.1.2 Time Expense

To report the computational time expense of the estimation methods, it is important to provide both a detailed description of the computational setup and quantitative performance metrics. The computational setup for the estimations comprised an Apple M1 ARM-based system-on-a-chip with 8 GB of RAM, running macOS version 15.3.1 (24D70) and RStudio version 2023.12.1+402, along with the Risk Neutral Density (‘RND’) package. The metrics

Table A.2: Time Expense of Density Extraction Methods

Methods	Average Time In Seconds
Shimko's Spline	0.0056
Edgeworth Density	0.0015
Black-Scholes-Merton Model	0.0028
Mixture of Two Log-normals	0.4540
GB2	0.4911

The computational setup comprises of an Apple M1 ARM-based system-on-a-chip, memory 8 GB RAM, macOS Version 15.3.1 (24D70) operating system, Rstudio Version 2023.12.1+402 language, and 'RND' or Risk Neutral Density Package

presented in Table A.2 represent the average CPU time required by each method over approximately 5000 density computations. Notably, the Edgeworth Density method requires the least amount of time to compute a density estimation, whereas the GB2 method is the most time-consuming.

## A.2 Structural VAR Methodology

According to Piffer & Podstawski (2018), the uncertainty shock proxy is built in two steps. First, a list of events that likely affected economic uncertainty independently of other macroeconomic shocks is compiled. Next, fluctuations in safe haven asset prices around these events are analysed to further refine the proxy.

### A.2.1 Events

To isolate periods when uncertainty likely shifted exogenously relative to the broader economy,

we compile a list of events that were both unexpected and independent of other macroeconomic shocks, and that either increased or decreased uncertainty. We begin with the events identified by Bloom (2009) via peaks in the VXO, and then expand our list using data from natural disaster records, armed conflicts, terrorist attacks, political elections and judicial decisions as reflected in Piffer & Podstawski (2018).

Events that could have been anticipated by economic agents or that may be linked to other macroeconomic shocks are excluded. Bloom (2009) comprises 38 events used to identify the VAR model, which is estimated on approximately 400 monthly observations. As detailed in Table 2.20, our dataset includes 62 events, and our baseline analysis focuses on 16 baseline events, these events are marked with an asterisk (\*).

To determine when news of these events reached the market, we rely on Bloomberg News releases, as Bloomberg is a primary source of information for market participants due to its global data aggregation.

### **A.2.2 Proxy Computation**

We select gold as the most favourable safe haven asset for constructing the proxy. As specified by Bloom (2009), this is firstly because the proxy based on the price of gold Granger-causes several measures of uncertainty, suggesting a higher informational content in the dynamics of uncertainty. Secondly, the gold-price proxy is more closely correlated with the VXO residuals from the estimated VAR model, which indicates the presence of a stronger relationship with

the drivers of the data analysed in the VAR model, as also shown in Piffer & Podstawski (2018). These conclusions were reached after comparing the gold proxy with those based on other precious metals, on the price of Treasury bills, and on the VXO, as well as with proxies constructed as dummy variables taking the value 1 when events occurred, the value 1 when the VXO peaked, or values of 1 or  $-1$  when an event was judged to imply an increase or a decrease in uncertainty, respectively.

We utilise intra-day data from the London spot market for physical gold, using prices from the two daily auctions organised at 10:30 and 15:00 by the London Bullion Market. We compute the proxy for the uncertainty shock as the percentage variation in the price of gold around the selected events. Specifically, given an event  $E_j$ , with  $j = 1, \dots, N$  (where  $N$  is the total number of events considered), let  $\gamma^j$  represent the time when event  $E_j$  became known to the market. For each event, we calculate  $\Delta p^j$  as the percentage change in the price of gold between the last available auction price before  $\gamma^j$  and the first available auction price after  $\gamma^j$ . Following the approach of C. D. Romer & D. H. Romer (2004), these  $N$  realisations of  $\Delta p^j$  are then aggregated into a monthly time series by summing the daily proxies within each month.

### A.2.3 Exogeneity of the Proxy SVAR

Since structural shocks are not directly observable, it is impossible to test whether the proxy for the uncertainty shock satisfies exogeneity conditions through direct observation as outlined in equations A.1 and A.2.

We state  $m_t$  as the proxy for the uncertainty shock, and define the  $k \times 1$  vector  $\varphi$  as  $\varphi = (\varphi_u, \varphi^*)'$ , where  $\varphi_u = E(\epsilon_t^u m_t)$  and  $\varphi^* = E(\epsilon_t^{\mathbf{u}} m_t)$ .

If,

$$E(\epsilon_t^u m_t) \equiv \varphi_u \neq 0, \tag{A.1}$$

$$E(\epsilon_t^{\mathbf{u}} m_t) \equiv \varphi^* = \mathbf{0}, \tag{A.2}$$

then  $m_t$  can be used as an instrument to identify  $\epsilon_t^u$ , to allow for isolating variations in  $\mathbf{u}_t$  that are driven by  $\mathbf{u}_t$  specified in Equation 2.39 rather than by  $\epsilon_t^*$ . However, if  $m_t$  correlates also with some of the structural shocks in  $\epsilon_t^*$ , then  $\epsilon_t^u$  requires further restrictions that prevent the estimated shock  $\epsilon_t^u$  from being influenced by the other structural shocks that  $m_t$  correlates with.

Therefore, we assess the exogeneity condition by documenting the relationship between our proxy and several measures of structural shocks.

Each shock is used at the original frequency provided in the datasets by the respective authors, and we aggregate our proxy  $m_t^g$  to that frequency when necessary. Reassuringly, we find a significant correlation with one of the two uncertainty shock instruments, namely the residual from an AR(2) regression on the VIX.

Uncertainty and news shocks may be interrelated, as certain variations in uncertainty could be linked to first-moment shocks akin to news shocks (Baker & Bloom (2013)). Bloom (2009) applied the tests from the proxy  $m_t^g$  model to news shocks on future productivity estimated. Bloom (2009) observes a statistically significant relationship with the news shocks estimations, indicating some correlation between the proxy for the uncertainty shock and the news shocks (which potentially affects risk appetite of economic agents) reported in the literature. This correlation can be interpreted in several ways:

- One possibility is that our proxy for the uncertainty shock also captures news shocks.
- Alternatively, the identification strategy employed may not fully disentangle news shocks from uncertainty shocks, thereby reflecting uncertainty shocks.

To minimise the risk of conflating uncertainty shocks with news shocks, we identify both types of shocks within a unified framework.

---