

Electronic Thesis and Dissertation Repository

December 2011

A Collection of Portfolio Management Issues

Mike McCausland

The University of Western Ontario

Supervisor

John Knight

The University of Western Ontario

Graduate Program in Economics

A thesis submitted in partial fulfillment of the requirements for the degree in Doctor of Philosophy

© Mike McCausland 2011

Follow this and additional works at: <https://ir.lib.uwo.ca/etd>

 Part of the [Econometrics Commons](#)

Recommended Citation

McCausland, Mike, "A Collection of Portfolio Management Issues" (2011). *Electronic Thesis and Dissertation Repository*. 346.
<https://ir.lib.uwo.ca/etd/346>

This Dissertation/Thesis is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Electronic Thesis and Dissertation Repository by an authorized administrator of Scholarship@Western. For more information, please contact tadam@uwo.ca.

A COLLECTION OF PORTFOLIO MANAGEMENT
ISSUES

(Spine title: A Collection of Portfolio Management Issues)
(Thesis format: Integrated Article)

by

Mike McCausland

Graduate Program in Economics

SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

SCHOOL OF GRADUATE STUDIES
THE UNIVERSITY OF WESTERN ONTARIO
LONDON, ONTARIO, CANADA
DECEMBER 2011

© Mike McCausland, 2011

THE UNIVERSITY OF WESTERN ONTARIO
SCHOOL OF GRADUATE AND POSTDOCTORAL STUDIES

CERTIFICATE OF EXAMINATION

Supervisor

Dr. John Knight

Supervisory Committee

Dr. Chris Robinson

Dr. Youngki Shin

Examiners

Dr. Fuchun Li

Dr. Stephen Sapp

Dr. Youngki Shin

Dr. Maria Ponomareva

The thesis by

Mike McCausland

entitled

A Collection of Portfolio Management Issues

is accepted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Date _____

Chair of the Thesis Examining Board

Abstract

This thesis consists of three chapters of interest to a portfolio manager. The first paper examines how the profitability of trading rules depends on volatility. In particular, a question of interest is whether one rule dominates all others regardless of the level of volatility, or whether it is more profitable to vary the choice of trading rule corresponding to volatility. Certain rules, such as the KST indicator using overbought/oversold levels, appear to excel under highly volatile conditions, while exponential moving average rules perform better with low volatility. In the second paper, a Value-at-Risk (VaR) model capable of producing accurate and robust forecasts is presented. In particular, the model presented here provides an extension to the VARLINEX model of Knight, Satchell, and Wang (2003) (hereafter KSW (2003)). The end result is a model capable of accurately forecasting VaR during the recent stock market crash (2008-09), as well as before and after the crash. The new model outperforms a benchmark model that had been successful prior to the crash, as well as the original VARLINEX model (KSW (2003)). The third paper explicitly spells out the link between independence tests and goodness-of-fit tests that are based on copula functions. However, the primary contribution is the development of a new copula-based goodness of fit test, which involves incorporating a weighting function in one of the test statistics proposed in Genest, Rémillard, and Beaudoin (2009). Guidance is given in terms of how to choose an appropriate weighting function, and an application to Value-at-Risk forecasting is included.

Acknowledgements

I wish to express my deepest gratitude to my supervisor John Knight as well as Stephen Sapp, Youngki Shin, Ivan Medovikov, Daniel Berg, Grant Wang, Ben Lester, Lance Lochner, Jim MacGee, Gregory Pavlov, Peter Streufert, Maria Ponomareva, Hiroyoki Kasahara, Martijn van Hasselt, Tim Conley, Chris Robinson, Todd Stinebrickner, Yvonne Adams, Javier Cano Urbina, Andrew Agopsowicz, Dan Montanera, and Chris Mitchell.

London, Canada

Mike McCausland

December 18, 2011

Table of Contents

Certificate of Examination	ii
Abstract	iii
Acknowledgements	iv
Table of Contents	v
List of Tables	viii
List of Figures	xiv
List of Appendices	xvi
1 Introduction	1
1.1 Bibliography	5
2 How Does the Profitability of Trading Rules Depend on Volatility?	6
2.1 Introduction	6
2.2 Literature Review	8
2.2.1 Profitability	8
2.2.2 Data Mining	9
2.2.3 Complex Rules	11
2.2.4 Intradaily Frequency	11
2.3 Data Mining Tests	12
2.4 Portfolio Selection	17
2.5 Data and Trading Rules	20
2.5.1 Data	20
2.5.2 Trading Rules	21
2.6 Results	26

2.6.1	In-Sample Results	27
2.6.2	Out-of-Sample Results	42
2.6.3	Seasonality	58
2.6.4	Final Thoughts on Results	59
2.7	Extensions	60
2.7.1	Should the Choice of Trading Rule Be Dependent on Volatility?	60
2.7.2	What Can Be Learned by Analyzing the Three ‘Sub-Portfolios’ Separately?	61
2.8	Data Mining	64
2.9	Conclusions	66
2.10	Bibliography	67
3	The VARLINEX Model Extended: Accurate Value-at-Risk Forecasts During the 2008-09 Market Crash	71
3.1	Introduction	71
3.2	VARLINEX Model	74
3.2.1	Initial Structure of VaR Problem	74
3.2.2	Value-at-Risk Linear Exponent (VARLINEX) Model	76
3.3	Literature Review	80
3.3.1	Quantile Estimation	81
3.3.2	Volatility Modelling	82
3.4	Data	83
3.5	Extended VARLINEX Model	84
3.5.1	Estimation method 1	86
3.5.2	Estimation method 2	88
3.6	Results	90
3.6.1	Results for Estimation Method 1	95
3.6.2	Results for Estimation Method 2	96
3.6.3	Which Estimation Method Performs Better?	99
3.6.4	Extension - Varying window lengths	101
3.7	Conclusion	103
3.8	Bibliography	104
4	A Copula-based Goodness-of-fit Test Based on Rosenblatt’s Transformation	107
4.1	Introduction	107
4.2	Introduction to Copula	112
4.2.1	Properties	112
4.2.2	Dependence Concepts	113

4.2.3	Parametric Families of Copulas	121
4.2.4	Parameter Estimation	127
4.3	Copula-based Goodness-of-fit Tests	129
4.3.1	Previous Methods	130
4.3.2	New Test	133
4.3.3	Connection to Independence Test Literature	135
4.3.4	Weighting Function	137
4.4	Power Study	142
4.5	General Application	152
4.6	Application to Pair-copula Decomposition	163
4.7	Value-at-Risk Application	173
4.8	Conclusion	176
4.9	Bibliography	178
5	Conclusion	183
5.1	Bibliography	185
A	List of Stocks	186
B	Legend for Tables	188
C	DJI, NASDAQ, & S&P/TSX Results	190
D	Dependence Functions	218
	Curriculum Vitae	222

List of Tables

2.1	Top 10 Rules for Daily Frequency - in-Sample	28
2.2	Top 10 Rules for High Volatility - in-Sample (Daily Frequency)	29
2.3	Top 10 Rules for Medium Volatility - in-Sample (Daily Frequency) . .	30
2.4	Top 10 Rules for Low Volatility - in-Sample (Daily Frequency)	30
2.5	Top 10 Rules for 130 Minute Frequency - in-Sample	31
2.6	Top 10 Rules for High Volatility - in-Sample (130 Min. Frequency) . .	32
2.7	Top 10 Rules for Medium Volatility - in-Sample (130 Min. Frequency)	32
2.8	Top 10 Rules for Low Volatility - in-Sample (130 Min. Frequency) . .	33
2.9	Top 10 Rules for 78 Minute Frequency - in-Sample	33
2.10	Top 10 Rules for High Volatility - in-Sample (78 Min. Frequency) . .	34
2.11	Top 10 Rules for Medium Volatility - in-Sample (78 Min. Frequency)	34
2.12	Top 10 Rules for Low Volatility - in-Sample (78 Min. Frequency) . .	35
2.13	Top 10 Rules for 30 Minute Frequency - in-Sample	35
2.14	Top 10 Rules for High Volatility - in-Sample (30 Min. Frequency) . .	36
2.15	Top 10 Rules for Medium Volatility - in-Sample (30 Min. Frequency)	36
2.16	Top 10 Rules for Low Volatility - in-Sample (30 Min. Frequency) . .	37
2.17	Top 10 Rules for High Volatility - in-Sample (Average Return Over all Frequencies)	37
2.18	Top 10 Rules for Medium Volatility - in-Sample (Average Return Over all Frequencies)	38
2.19	Top 10 Rules for Low Volatility - in-Sample (Average Return Over all Frequencies)	38

2.20	Top 10 Rules for Daily Frequency - out-of-Sample	44
2.21	Top 10 Rules for High Volatility - out-of-Sample (Daily Frequency) . .	45
2.22	Top 10 Rules for Medium Volatility - out-of-Sample (Daily Frequency)	45
2.23	Top 10 Rules for Low Volatility - out-of-Sample (Daily Frequency) . .	46
2.24	Top 10 Rules for 130 Min. Frequency - out-of-Sample	47
2.25	Top 10 Rules for High Volatility - out-of-Sample (130 Min. Frequency)	47
2.26	Top 10 Rules for Medium Volatility - out-of-Sample (130 Min. Frequency)	48
2.27	Top 10 Rules for Low Volatility - out-of-Sample (130 Min. Frequency)	48
2.28	Top 10 Rules for 78 Min. Frequency - out-of-Sample	49
2.29	Top 10 Rules for High Volatility - out-of-Sample (78 Min. Frequency)	49
2.30	Top 10 Rules for Medium Volatility - out-of-Sample (78 Min. Frequency)	50
2.31	Top 10 Rules for Low Volatility - out-of-Sample (78 Min. Frequency)	50
2.32	Top 10 Rules for 30 Min. Frequency - out-of-Sample	51
2.33	Top 10 Rules for High Volatility - out-of-Sample (30 Min. Frequency)	51
2.34	Top 10 Rules for Medium Volatility - out-of-Sample (30 Min. Frequency)	52
2.35	Top 10 Rules for Low Volatility - out-of-Sample (30 Min. Frequency)	52
2.36	Top 10 Rules for High Volatility - out-of-Sample (Average Return Over all Frequencies)	53
2.37	Top 10 Rules for Medium Volatility - out-of-Sample (Average Return Over all Frequencies)	55
2.38	Top 10 Rules for Low Volatility - out-of-Sample (Average Return Over all Frequencies)	56
2.39	Number of Rules That Beat the Benchmark - in-Sample	62
2.40	Number of Rules That Beat the Benchmark - out-of-Sample	63
2.41	Number of Rules with Superior Predictive Ability - in-Sample	64
2.42	Number of Rules with Superior Predictive Ability - out-of-Sample . .	65
3.1	DJI - Summary Statistics	83
3.2	DJI - Full sample - number of VaR violations (N=1466)	97
3.3	DJI - Pre-crisis - number of VaR violations (N=968)	97

3.4	DJI - Crisis - number of VaR violations (N=355)	97
3.5	DJI - Post-crisis - number of VaR violations (N=143)	97
3.6	DJI - Full sample - number of VaR violations (N=1466)	100
3.7	DJI - Pre-crisis - number of VaR violations (N=968)	100
3.8	DJI - Crisis - number of VaR violations (N=355)	100
3.9	DJI - Post-crisis - number of VaR violations (N=143)	100
4.1	Tail dependence parameters for several copulas	120
4.2	Percentage of rejection of H_0 - comparison between my T_{3b} and Berg's T_{3b}	145
4.3	Percentage of rejection of H_0 - weighted vs. unweighted (n=100) . . .	146
4.4	Percentage of rejection of H_0 - weighted vs. unweighted (n=300) . . .	147
4.5	Percentage of rejection of H_0 - weighted vs. unweighted (n=500) . . .	148
4.6	Summary Statistics	152
4.7	Pairings with High Power	162
4.8	Pairings with Low Power	162
4.9	Summary Statistics	166
4.10	Copula Parameter Values	175
4.11	Number of Violations	176
C.1	DJI - Full sample - number of VaR violations (N=1466)	191
C.2	DJI - Pre-crisis - number of VaR violations (N=968)	191
C.3	DJI - Crisis - number of VaR violations (N=355)	191
C.4	DJI - Post-crisis - number of VaR violations (N=143)	192
C.5	DJI - Full sample - number of VaR violations (N=1466)	192
C.6	DJI - Pre-crisis - number of VaR violations (N=968)	192
C.7	DJI - Crisis - number of VaR violations (N=355)	193
C.8	DJI - Post-crisis - number of VaR violations (N=143)	193
C.9	DJI - Full sample - number of VaR violations (N=1466)	194
C.10	DJI - Pre-crisis - number of VaR violations (N=968)	194

C.11 DJI - Crisis - number of VaR violations (N=355)	194
C.12 DJI - Post-crisis - number of VaR violations (N=143)	195
C.13 DJI - Full sample - number of VaR violations (N=1466)	195
C.14 DJI - Pre-crisis - number of VaR violations (N=968)	195
C.15 DJI - Crisis - number of VaR violations (N=355)	196
C.16 DJI - Post-crisis - number of VaR violations (N=143)	196
C.17 DJI - Full sample - number of VaR violations (N=1466)	196
C.18 DJI - Pre-crisis - number of VaR violations (N=968)	197
C.19 DJI - Crisis - number of VaR violations (N=355)	197
C.20 DJI - Post-crisis - number of VaR violations (N=143)	197
C.21 DJI - Full sample - number of VaR violations (N=1466)	198
C.22 DJI - Pre-crisis - number of VaR violations (N=968)	198
C.23 DJI - Crisis - number of VaR violations (N=355)	198
C.24 DJI - Post-crisis - number of VaR violations (N=143)	199
C.25 NASDAQ - Full sample - number of VaR violations (N=1466)	200
C.26 NASDAQ - Pre-crisis - number of VaR violations (N=968)	200
C.27 NASDAQ - Crisis - number of VaR violations (N=355)	200
C.28 NASDAQ - Post-crisis - number of VaR violations (N=143)	200
C.29 NASDAQ - Full sample - number of VaR violations (N=1466)	201
C.30 NASDAQ - Pre-crisis - number of VaR violations (N=968)	201
C.31 NASDAQ - Crisis - number of VaR violations (N=355)	201
C.32 NASDAQ - Post-crisis - number of VaR violations (N=143)	202
C.33 NASDAQ - Full sample - number of VaR violations (N=1466)	202
C.34 NASDAQ - Pre-crisis - number of VaR violations (N=968)	202
C.35 NASDAQ - Crisis - number of VaR violations (N=355)	203
C.36 NASDAQ - Post-crisis - number of VaR violations (N=143)	203
C.37 NASDAQ - Full sample - number of VaR violations (N=1466)	203
C.38 NASDAQ - Pre-crisis - number of VaR violations (N=968)	204
C.39 NASDAQ - Crisis - number of VaR violations (N=355)	204

C.40 NASDAQ - Post-crisis - number of VaR violations (N=143)	204
C.41 NASDAQ - Full sample - number of VaR violations (N=1466)	205
C.42 NASDAQ - Pre-crisis - number of VaR violations (N=968)	205
C.43 NASDAQ - Crisis - number of VaR violations (N=355)	206
C.44 NASDAQ - Post-crisis - number of VaR violations (N=143)	206
C.45 NASDAQ - Full sample - number of VaR violations (N=1466)	206
C.46 NASDAQ - Pre-crisis - number of VaR violations (N=968)	207
C.47 NASDAQ - Crisis - number of VaR violations (N=355)	207
C.48 NASDAQ - Post-crisis - number of VaR violations (N=143)	207
C.49 NASDAQ - Full sample - number of VaR violations (N=1466)	208
C.50 NASDAQ - Pre-crisis - number of VaR violations (N=968)	208
C.51 NASDAQ - Crisis - number of VaR violations (N=355)	208
C.52 NASDAQ - Post-crisis - number of VaR violations (N=143)	209
C.53 NASDAQ - Full sample - number of VaR violations (N=1466)	209
C.54 NASDAQ - Pre-crisis - number of VaR violations (N=968)	209
C.55 NASDAQ - Crisis - number of VaR violations (N=355)	210
C.56 NASDAQ - Post-crisis - number of VaR violations (N=143)	210
C.57 NASDAQ - Summary Statistics	210
C.58 S&P/TSX Composite - Full sample - number of VaR violations (N=1350)	211
C.59 S&P/TSX Composite - Pre-crisis - number of VaR violations (N=859)	211
C.60 S&P/TSX Composite - Crisis - number of VaR violations (N=350)	211
C.61 S&P/TSX Composite - Post-crisis - number of VaR violations (N=141)	211
C.62 S&P/TSX Composite - Full sample - number of VaR violations (N=1350)	212
C.63 S&P/TSX Composite - Pre-crisis - number of VaR violations (N=859)	212
C.64 S&P/TSX Composite - Crisis - number of VaR violations (N=350)	212
C.65 S&P/TSX Composite - Post-crisis - number of VaR violations (N=141)	213
C.66 S&P/TSX Composite - Full sample - number of VaR violations (N=1350)	213
C.67 S&P/TSX Composite - Pre-crisis - number of VaR violations (N=859)	213
C.68 S&P/TSX Composite - Crisis - number of VaR violations (N=350)	214

C.69 S&P/TSX Composite - Post-crisis - number of VaR violations (N=141)	214
C.70 S&P/TSX Composite - Full sample - number of VaR violations (N=1350)	214
C.71 S&P/TSX Composite - Pre-crisis - number of VaR violations (N=859)	215
C.72 S&P/TSX Composite - Crisis - number of VaR violations (N=350)	215
C.73 S&P/TSX Composite - Post-crisis - number of VaR violations (N=141)	215
C.74 S&P/TSX Composite - Full sample - number of VaR violations (N=1350)	216
C.75 S&P/TSX Composite - Pre-crisis - number of VaR violations (N=859)	216
C.76 S&P/TSX Composite - Crisis - number of VaR violations (N=350)	216
C.77 S&P/TSX Composite - Post-crisis - number of VaR violations (N=141)	217
C.78 S&P/TSX - Summary Statistics	217

List of Figures

2.1	Value of Portfolio - out-of-Sample Period	40
2.2	Value of Portfolio - out-of-Sample Period	54
3.1	Correlograms	85
3.2	Dow Jones - 95% confidence interval of 99% VaR forecasts	98
3.3	Method 1 Vs. Method 2 - 99% VaR forecasts (DJI data)	102
4.1	Minimum copula	113
4.2	Maximum copula	114
4.3	Product copula	114
4.4	Clayton scatter plot $n = 1000, \tau = 0.7$	119
4.5	Gumbel scatter plot $n = 1000, \tau = 0.7$	119
4.6	Normal scatter plot $n = 1000, \tau = 0.7$	120
4.7	Normal copula scatter plot with Student-t(3) and Beta(2,2) marginals	124
4.8	Normal copula density ($\rho = 0.8$)	124
4.9	Clayton copula density ($\theta = 2$)	126
4.10	Gumbel copula density ($\theta = 3$)	126
4.11	Clayton vs. Clayton $n = 500, \tau = 0.2$	139
4.12	Clayton vs. Gumbel $n = 500, \tau = 0.2$	140
4.13	Clayton vs. Normal $n = 500, \tau = 0.2$	140
4.14	Clayton vs. Clayton $n = 100, \tau = 0.2$	146
4.15	Clayton vs. Normal $n = 100, \tau = 0.2$	147
4.16	Gumbel vs. Normal $n = 500, \tau = 0.2$	151

4.17	Gumbel vs. Clayton $n = 500, \tau = 0.2$	151
4.18	Correlograms for the raw MSFT data	154
4.19	Lambda functions ($n = 300$)	158
4.20	Lambda functions ($n = 100$)	160
4.21	Lambda functions	161
4.22	Correlograms for the raw Barrick Gold data	167
4.23	Lambda functions	171
4.24	Lambda functions	172
4.25	VaR forecasts along with log-returns from portfolio	177
D.1	Gumbel vs. Gumbel $n = 500, \tau = 0.2$	218
D.2	Gumbel vs. Normal $n = 500, \tau = 0.2$	219
D.3	Gumbel vs. Clayton $n = 500, \tau = 0.2$	219
D.4	Normal vs. Normal $n = 500, \tau = 0.2$	220
D.5	Normal vs. Gumbel $n = 500, \tau = 0.2$	220
D.6	Normal vs. Clayton $n = 500, \tau = 0.2$	221

List of Appendices

A	List of Stocks	187
B	Legend for Tables	188
C	DJI, NASDAQ, & S&P/TSX Results	190
D	Dependence Functions (<i>DEPFUNC</i>)	218

Chapter 1

Introduction

This thesis consists of three papers dealing with portfolio management issues. Effective portfolio management is a complex task requiring the portfolio manager to consider many important issues. For example, when an asset is being considered for inclusion in a portfolio it is first analyzed on a stand-alone basis, then it must be determined whether the asset is a good fit in the portfolio, and then whether the timing is right to execute the trade from a technical analysis viewpoint. The first paper in this thesis is related to technical analysis. If the asset passes these tests, then the portfolio manager may also need to worry about other issues such as how to achieve best execution for the trade. Integral to deciding whether the asset is a good fit for the portfolio is the issue of risk management, which is a complex field of study in its own right. A Value-at-Risk model is a commonly used tool within risk management. In the second paper, a Value-at-Risk forecasting model is developed. Within several of the sub-fields of portfolio management is the issue of data modelling. Copula functions provide a flexible way to model the dependence structure

between multiple data variables. Data modelling with copula functions is the topic of the third paper.

The first paper is related to technical analysis and the profitability of technical trading rules. There exists a vast amount of literature dedicated to the examination of the profitability of technical trading rules. The results have been mixed; many studies claim trading rules are profitable, while several others claim just the opposite. Rather than adding another paper to the pile, this study puts a different spin on the problem by examining how the profitability of trading rules depends on the level of volatility in the asset being traded. Of particular interest is the issue of whether one rule dominates all others regardless of the level of volatility, or whether it is more profitable to vary the choice of trading rule corresponding to volatility. Several classes of trading rules are tested on a portfolio of 20 stocks from the Russell 3000 index. The results of this paper should provide guidance to a practitioner attempting to develop an investment strategy based on technical trading rules, or to a portfolio manager using trading rules to decide the proper time to enter/exit positions in his portfolio. During times of high volatility, there is plenty of money to be made or lost in financial markets, as a result, it is important that investors are using strategies that are appropriate given the current market conditions. It may be the case that investors need to adjust their trading rules in order to maximize profits/minimize losses as volatility fluctuates. This paper makes three contributions to the literature. First, it is determined how the ranking of trading rules (in terms of profitability) depends on volatility. Second, the trading rules are tested on intradaily data, which does not appear to have been done before with equities. Finally, the profitability of trading rules in ‘young’ vs. ‘mature’ markets is examined.

In the second paper, a Value-at-Risk (VaR) model capable of producing accurate and robust forecasts is presented. The recent market crash of 2008-2009 has made the measurement and forecasting of risk a major issue for both regulators and risk management divisions. Under the Basel II and Basel III accords, banks are required to maintain a specific level of capital which is related to the riskiness of their assets. The loss that is only surpassed $(1 - \alpha)\%$ of the time is referred to as the $\alpha\%$ VaR.¹ A VaR model is recognized as an effective method used in the calculation of required risk capital, in particular in the context of market risk. From a portfolio management perspective, an accurate VaR model can help to identify exposures that carry an unacceptable amount of downside risk, possibly allowing for the portfolio manager to make changes to the portfolio before it loses too much value. Most of the existing VaR literature ignores the asymmetric loss of forecasting VaR for different economic agents. The VARLINEX model developed by Knight, Satchell, and Wang (2003) (hereafter KSW (2003)) takes this asymmetry into account by using the Linex loss function, which is an alternative to the traditionally used symmetric loss function. Specifically, bank managers likely feel that the loss of overestimating VaR is greater than the loss of underestimation. This is because, in the event of overestimation, banks will hold more capital than is required by regulation, which will decrease their return on equity. On the other hand, regulators certainly would suffer a greater loss from underestimation relative to overestimation, since underestimation is more likely to result in systemic bank failures. Additionally, a bank manager may wish to err on the side of caution in a time of crisis, in which case his perspective will be similar to that of the regulator. In this paper, an extension is provided to the VARLINEX

¹Typical values for α are 0.95, 0.975, 0.99, and 0.995.

model. Two alternative methods of estimating the asymmetry parameter from the VARLINEX model are introduced. These new estimation methods improve upon the results of the original model (i.e. KSW (2003)), in addition to outperforming the benchmark model (which was based on Extreme Value Theory (EVT)). As a result, this paper provides a versatile VaR model capable of yielding accurate forecasts during a time of crisis (which the benchmark EVT model was unable to do), as well as under calm market conditions.

In the third paper, a new copula-based goodness-of-fit (g-o-f) test is developed. Copulas are used to model the dependence structure between two or more variables. From the standpoint of a portfolio manager these variables could be, for example, returns on any type of financial asset, or risk factors affecting the returns on financial assets. Copula-based g-o-f tests are useful in several financial applications, such as integrated risk management, optimal portfolio construction, and asset pricing. The tests are also useful in the pair-copula decomposition method, which decomposes a high-dimensional copula into a function of several bivariate copulas. In all of these applications, the tests are used to determine which family of copula functions best describes the data. The new test developed in this paper implements a weighting function in one of the test statistics proposed in Genest et al. (2009). The result is a test statistic with greater power than the original unweighted version, which should be of interest to any practitioners who wish to model financial data using copulas. The methodology used to derive an appropriate weighting function is introduced. An application to Value-at-Risk forecasting shows how the test can be used in practice by the portfolio manager. In addition, a link between copula-based independence tests and copula-based g-o-f tests is made explicit.

1.1 Bibliography

- Genest, C., Rémillard, B., Beaudoin, D. (2009), “Goodness-of-Fit Tests for Copulas: A Review and Power Study”, *Insurance: Mathematics and Economics*, 44, 199-213.
- Knight, J., S. Satchell, and G. Wang (2003), “Value at Risk Linear Exponent (VARLINEX) Forecasts”, *Quantitative Finance*, 3, 332-344.
- McNeil, A. J., and R. Frey (2000), “Estimation of Tail-related Risk Measures for Heteroscedastic Financial Time Series: An Extreme Value Approach”, *Journal of Empirical Finance*, 7, 271-300.

Chapter 2

How Does the Profitability of Trading Rules Depend on Volatility?

2.1 Introduction

There exists a vast amount of literature dedicated to the examination of the profitability of technical trading rules. The results have been mixed; many studies claim trading rules are profitable, while several others claim just the opposite. However, there have been some interesting findings that warrant further exploration. For instance, Hsu and Kuan (2005) find evidence that trading rules are able to generate significant profits in ‘young’ markets such as the Russell 2000 and the NASDAQ Composite, while several studies, including Hsu and Kuan (2005), Sullivan, Timmerman, and White (1999), and Aronson (2007), find that trading rules are unable to generate significant profits in more ‘mature’ markets such as the Dow Jones Industrial Average (DJIA) and the S&P 500. In addition, Hsu and Kuan (2005) find evidence that

complex trading rules consistently outperform simple trading rules.¹ Although the aforementioned findings will be tested here as well, the primary focus of this paper is to determine how the profitability of trading rules depends on the volatility in the asset being traded. Each day is defined to be either ‘high’, ‘medium’, or ‘low’ in terms of volatility. The method used to determine which volatility range a day is in is explained in Section 2.4. Several simple and complex trading rules are tested on a portfolio of 20 stocks from the Russell 3000 index. Specifically, the portfolio consists of six stocks from the Russell Top 200, and seven from each of the Russell Midcap and the Russell 2000.² Further details regarding the portfolio are deferred to Section 2.4. The Stepwise Superior Predictive Ability (Step-SPA) test proposed in Hsu, Hsu, and Kuan (2010) is used to correct the data mining bias.

The results of this paper should provide guidance to a practitioner attempting to develop an investment strategy based on technical trading rules. During times of high volatility, there is plenty of money to be made or lost in the financial markets, as a result, it is important that investors are using strategies that are appropriate given the current market conditions. It may be the case that investors need to adjust their trading rules in line with volatility in order to maximize profits/minimize losses.

The contribution of this paper is three-fold. First, it is established that the ranking of trading rules (in terms of profitability) depends on volatility. Second, the trading rules are tested on intradaily data, which does not appear to have been done

¹In general, complex trading rules take as inputs the output of several simple trading rules. Concrete examples of such rules will be provided in Section 2.5

²The Russell Top 200 is made up of the 200 largest public companies in the United States based on market capitalization, the Russell Midcap Index is made up of the 201st to 1000th largest companies in the US; and the Russell 2000 is made up of the 1001st to 3000th largest companies.

before with equities. Finally, the profitability of trading rules in ‘young’ vs. ‘mature’ markets is examined.

The paper is organized as follows. Section 2.2 reviews the existing literature. Section 2.3 describes the current data mining tests. Section 2.4 describes the portfolio selection methodology. Section 2.5 describes the data and the universe of trading rules. Section 2.6 presents and analyzes the results of the tests. Section 2.7 analyzes some interesting extension questions. Section 2.8 discusses the results of the Step-SPA test for data mining. Section 2.9 concludes.

2.2 Literature Review

2.2.1 Profitability

Whether or not technical trading rules have the ability to generate significant profits has been a topic of hot debate in the finance literature over the past few decades. There are two things that make it such a popular topic. The first is its direct relevance to investment practitioners, and the second is that there is not yet a clear winner. Many studies have surfaced that support the profitability of trading rules, including Sweeney (1986, 1988), Brock, Lakonishok, and LeBaron (1992), Neely, Weller, and Dittmar (1997), Rouwenhorst (1998), Chang and Osler (1999), Neely and Weller (1999), Lo, Mamaysky, and Wang (2000), Korajczyk and Sadka (2004), and Hsu and Kuan (2005). Likewise, there have been many studies to conclude that trading rules

cannot generate significant profits, including Knez and Ready (1996), Sullivan, Timmerman, and White (1999), Ready (2002), and Lesmond, Schill, and Zhou (2003). However, all of these articles contain at least one of the following deficiencies: i) the article did not account for data-mining bias, ii) the article only considered simple trading rules (as opposed to complex), iii) the article only tested the strategies on daily or weekly data (as opposed to intradaily data). All three issues are addressed in this paper.

2.2.2 Data Mining

Sullivan, Timmerman, and White (1999) (hereafter STW) implemented White’s Reality Check bootstrap methodology (WRC), the first test that was able to quantify the data mining bias. STW define data mining as a condition that arises “when a given set of data is used more than once for purposes of inference or model selection”.³ They go on to explain that “when such data reuse occurs, there is always the possibility that any satisfactory results obtained may simply be due to chance rather than any merit inherent in the method yielding the results”. STW applies a universe of 7,846 simple trading rules to 100 years of daily data on the Dow Jones Industrial

³STW refers to this as ‘data snooping’ rather than ‘data mining’. This paper follows the convention of Aronson (2007). Aronson defines data snooping as the use of “results of prior rule studies reported by other researchers”. He contends that data snooping is a greater problem than data mining because the researchers “typically do not disclose the amount of data mining that led to the discovery (of the aforementioned results)”.

Average. They find that significant profits are not generated by any of the rules after taking into account the data mining bias.

However, many improvements have been made in the last decade since WRC was introduced. WRC has two principle flaws; first, it determines whether there is a significant model, but it does not identify all significant models, and second, Hansen (2005) notes that WRC is conservative because its null distribution is obtained under the least favourable configuration. The first flaw is corrected by Romano and Wolf (2005) who propose a stepwise test based on WRC which identifies as many significant models as possible. Unfortunately, the stepwise test still possesses the second flaw since its null distribution is obtained under the least favourable configuration. Hansen (2005) introduces the SPA test which does not rely on the least favourable configuration. Hansen and Lunde (2005), and Hsu and Kuan (2005) show that Hansen's SPA test is more powerful than WRC. However, the SPA test suffers from the first flaw of WRC since it is unable to identify all significant models. Hsu, Hsu, and Kuan (2010) develop a stepwise SPA test which addresses both of the major flaws in WRC. Hsu, Hsu, and Kuan (2010) show that their stepwise SPA test is more powerful than the stepwise reality check test from Romano and Wolf (2005). In this paper, the stepwise SPA test introduced in Hsu, Hsu, and Kuan (2010) is used to correct for data mining bias. The details of the test will be presented in Section 2.3.

2.2.3 Complex Rules

In Tuttle (2006), the author claims that money managers misuse technical indicators for three reasons. One of these reasons is that they employ “only one indicator to make an investment decision”. Tuttle advocates using the output from several different indicators to make an investment decision. This sentiment is shared by many other experts on technical analysis, including Martin Pring (Pring (2002)). Following this line of thinking, it makes sense that money managers may wish to consider the use of complex rather than simple trading rules to make decisions. Thus, a study that exclusively examines the profitability of simple trading rules may not be relevant to investment practitioners. Hsu and Kuan (2005) examine the profitability of three classes of complex rules: i) learning, ii) voting, and iii) fractional.⁴ The authors find that complex rules are more profitable than simple rules in general, and that they are capable of generating significant profits in ‘young’ markets, such as the Russell 2000 and the NASDAQ Composite.

2.2.4 Intradaily Frequency

The vast majority of studies involving the use of intradaily data to test the profitability of trading rules focus on the foreign exchange (FX) market. Curcio, Goodhart, Guillaume, and Payne (1997) find that filter rules are unable to generate significant

⁴The voting and fractional rules in this paper are modified versions of those in Hsu and Kuan (2005)

profits on average. Neely, Weller, and Dittmar (1997) find evidence of excess profits when using genetic programming techniques to select technical trading rules. Osler (2000) finds that signals given by support and resistance levels are able to generate excess profits. Neely (2002) uses intradaily data to determine whether central bank intervention contributes to technical trading profits in the foreign exchange market. Neely and Weller (2003) find that excess profits are not generated in the FX market when either a genetic program, or an optimal linear forecasting model are used to select technical trading rules.

2.3 Data Mining Tests

First, some notation needs to be introduced. Let m indicate the total number of trading rules, and let n denote the number of periods in the sample. Denote r_t as the return of the trading rule at time t ,⁵ and r_t^B as the return on the benchmark at time t .⁶ Then, $d_{k,t} = \frac{1}{n} \sum_{t=1}^n \log(1 + r_t - r_t^B - \delta_{k,t} TC)$ is defined as the return of trading rule k at time t relative to the benchmark, where $k = 1, \dots, m$; $\delta_{k,t}$ is 1 if there is a buy or sell transaction at time t , and 0 otherwise, and TC is the transaction cost. Further discussion of the transaction cost is deferred to Section 2.6. A composite hypothesis test may be formed to test whether any of the trading rules generate significant profits. First, for each k , let $E(d_{k,t}) = \mu_k$ for all t . Then, the hypothesis test is

⁵Transaction costs are ignored for now.

⁶In this paper, the benchmark is always the buy-and-hold strategy.

$$H_0 : \max_{k=1,\dots,m} \mu_k \leq 0, \quad k = 1, \dots, m \quad (2.1)$$

The alternative is that the best performing rule generates significant profits relative to the benchmark.

White's Reality Check

The hypothesis test in (2.1) can be used to determine the existence of a profitable trading rule as long as μ has a continuous limiting distribution. Taking after Hansen (2005) and Hsu, Hsu, and Kuan (2010), the following condition is imposed on $d_t = (d_{1,t}, \dots, d_{m,t})'$ which allows d_t to exhibit weak dependence over time.

Assumption 2.1 $\{d_t\}$ is strictly stationary and α -mixing of size $(2+\eta)(s+\eta)/(s-2)$, for some $s > 2$ and $\eta > 0$, where $E[|d_t|^{(s+\eta)}] < \infty$, and $\text{var}(d_{k,t}) > 0$ for all k .⁷

Given Assumption 2.1, the data obey a central limit theorem:

$$\sqrt{n}(\bar{d} - \mu) \rightarrow N(0, \Omega), \quad (2.2)$$

where $\bar{d} = \frac{1}{n} \sum_{t=1}^n d_t$, $\mu = E(d_t)$, and $\Omega = \lim_{n \rightarrow \infty} \text{var}(n^{1/2}(\bar{d} - \mu))$. WRC uses the test statistic $\bar{V} = \max_{k=1,\dots,m} \{\sqrt{n}\bar{d}_k\}$, where \bar{d}_k is the k_{th} element of \bar{d} . The test uses the least favourable configuration (LFC), i.e., $\mu = 0$, in equation (2.1) to obtain the null distribution. Substituting $\mu = 0$ into equation (2.2) yields $\sqrt{n}\bar{d} \rightarrow N(0, \Omega)$, thus the limiting distribution of \bar{V} is $\max_k \{N(0, \Omega)\}$, which can be approximated using a stationary bootstrap procedure. The bootstrap procedure is used to obtain a p -value

⁷ $|\cdot|$ is the Euclidean norm.

for H_0 . The null hypothesis in (2.1) would be rejected at a significance level of α for a bootstrapped p -value less than α , or when the test statistic, \bar{V} , is greater than the bootstrapped critical value. White (2000) calls this procedure the “bootstrap reality check”, because it “provides an objective measure of the extent to which apparently good results accord with the sampling variation.”

The Stepwise Reality Check (Step-RC)

It was mentioned in Section 2.2 that WRC suffers from two primary flaws; the first of which was that it is unable to identify all significant models. Romano and Wolf (2005) solve this problem with a stepwise procedure that identifies as many significant models as possible. The procedure works as follows:

1. Create a vector that lists the \bar{d}_k 's in descending order.
2. For the best trading rule k , if $\sqrt{n}\bar{d}_k$ is greater than the bootstrapped critical value, reject the rule, otherwise end the procedure.
3. Remove the best rule and repeat step 2, bootstrapping the critical value again using the remaining data. The procedure ends when no more rules can be rejected.

However, this test does not address the second flaw of WRC.

The Superior Predictive Ability (SPA) Test

The second flaw mentioned in Section 2.2 was that WRC is too conservative. The reason for this is that it is thought the inclusion of very poor models may artificially inflate the bootstrapped p -value of WRC. As a result, Hansen (2005) introduces the SPA test which uses the test statistic given by

$$T_n^{SPA} = \max \left[\max_{k=1, \dots, m} \left\{ \frac{\sqrt{n} \bar{d}_k}{\hat{\omega}_k}, 0 \right\}, 0 \right],$$

where $\hat{\omega}_k^2$ is some consistent estimator of $\omega_k^2 = \text{var}(\sqrt{n} \bar{d}_k)$.

The SPA test avoids the LFC by re-centering the null distribution. Let $\hat{\Omega}$ be a consistent estimator of Ω with the (i, j) _{th} element $\hat{\omega}_{ij}$. Also, let $\hat{\sigma}_k^2 = \hat{\omega}_k^2/n$ and $A_{n,k} = -\hat{\sigma}_k \sqrt{2 \log \log n}$. $\hat{\mu}$ is defined as the vector with k _{th} element $\hat{\mu}_k = \bar{d}_k \mathbf{1}(\sqrt{n} \bar{d}_k \leq A_{n,k})$, where $\mathbf{1}(\cdot)$ is the indicator function. Finally, noting that $\sqrt{n} \bar{d} = \sqrt{n}(\bar{d} - \mu) + \sqrt{n} \mu$, Hansen (2005) proposes adding $\sqrt{n} \hat{\mu}$ to the bootstrapped distribution of $\sqrt{n}(\bar{d} - \mu)$. This re-centering of the distribution yields a better approximation to the null distribution of T_n^{SPA} .

The Stepwise SPA Test

Hsu, Hsu, and Kuan (2010) introduce the Stepwise SPA test (hereafter referred to as Step-SPA) which addresses both flaws of WRC. The procedure is the same as that of the stepwise test in Romano and Wolf (2005), except that its critical values are obtained from bootstrapping the re-centered distribution from the SPA test. The Step-SPA test uses the stationary bootstrap of Politis and Romano (1994). Before describing the bootstrap method, some notation is needed. Let $d_t^*(b) = d_{n_b, t}^*$, where

$t = 1, \dots, n$ be the b_{th} re-sample of d_t , where the indices $n_{b,1}, \dots, n_{b,n}$ consist of blocks of $\{1, \dots, n\}$ with random lengths determined by the realization of a geometric distribution with the parameter $Q \in [0, 1)$. The method goes as follows:

1. $n_{b,1}$ is chosen uniformly from $\{1, \dots, n\}$.
2. For all $t > 1$, $n_{b,t} = n_{b,t-1} + 1$ with probability Q ; otherwise, $n_{b,t}$ is chosen uniformly from $\{1, \dots, n\}$ as in step 1. A re-sample is done when n observations are drawn. Let $\bar{d}^*(b) = \sum_{t=1}^n d_t^*(b)/n$ denote the sample average of this re-sample.
3. Repeat steps 1 and 2 B times.

This yields an empirical distribution of \bar{d}^* with B realizations. Now, the bootstrapped T_n^{SPA} critical value is:

$$\hat{q}_{\alpha_0}^* = \max(\hat{q}_{\alpha_0}, 0), \quad (2.3)$$

where

$$\hat{q}_{\alpha_0} = \inf\{q | P^*[\sqrt{n} \max_{k=1, \dots, m} (\bar{d}_k^* - \bar{d}_k + \hat{\mu}_k) \leq q] \geq 1 - \alpha_0\}$$

is the $(1 - \alpha_0)_{th}$ quantile of the re-centered empirical distribution, and P^* is the bootstrapped probability measure. The Step-SPA test proceeds as follows:

1. Arrange the vector of \bar{d}_k 's in descending order.
2. Reject the best trading rule k if $\sqrt{n}\bar{d}_k > \hat{q}_{\alpha_0}$. If the best rule is rejected, go to step 3, otherwise end the procedure.

3. Remove the best rule from the data, and repeat step 2.
4. Repeat step 3 until no rules are rejected.

According to Hsu, Hsu, and Kuan (2010), the Step-SPA test is more powerful than the stepwise test in Romano and Wolf (2005), given Assumption 2.1 holds.

2.4 Portfolio Selection

The main issue related to portfolio selection was to ensure the portfolio consisted of a similar number of high, medium, and low volatility stocks.⁸ The reason is that there need to be enough periods of each type of volatility to test the rules on. For example, if only 15% of the periods in the sample were ‘low’ periods, this would not yield very much information pertaining to the ability of the rules to generate profits in a low volatility environment. In order to ensure there were an approximately equal number of periods of each range of volatility, it was necessary to determine what exactly it meant to be ‘high’, ‘medium’, or ‘low’ in terms of volatility. Essentially, ‘cutoff’ values that separated high from medium and medium from low needed to be computed. The following is a description of the method used to calculate the cutoff values.

First, daily return data was collected for a sample of 60 stocks from the Russell 3000 index over the period July 1, 2003 to June 20, 2007 (1000 days of data). Specifically, 20 stocks were randomly selected from each of the Russell Top 200, Russell

⁸To be more accurate, the portfolio needed to consist of a similar number of high, medium, and low volatility periods

Midcap, and Russell 2000 indexes. The reasoning behind using stocks from each of these three indexes was so that there would be a similar number of high, medium, and low volatility stocks. The assumption being that the volatility of a stock is inversely proportional to the size of the corresponding firm. This turns out to be the case in this portfolio, as the Russell 2000 stocks are, on average, the most volatile, followed by the Russell Midcap, and then the Russell Top 200.⁹ Now, the process used to calculate cutoff values is explained.

1. Split up the 1000 day sample into 20 discrete blocks each containing 50 days ($20 \times 50 = 1000$).
2. Calculate the standard deviation of returns over each 50 day block. Note that there will be 20 standard deviation values for the first stock since there were 20 discrete blocks over the sample.
3. Repeat steps 1) and 2) for the remaining 59 stocks.

When step 3 is complete there will be 1200 (60 stocks \times 20 discrete blocks = 1200) standard deviation calculations. Finally, the cutoff value between high and medium volatility will be the 67th percentile (roughly the 400th largest of the 1200 values) and the cutoff value between medium and low volatility will be the 33rd percentile (roughly the 800th largest of the 1200 values). The high-medium cutoff value is 0.0211 and the medium-low cutoff value is 0.0135. At this point it is important to note that

⁹The average standard deviation of returns was calculated for each stock, and then the average for each group was recorded. The average standard deviation was 18.5% for the Russell 2000 stocks, 11.2% for the Russell Midcap, and 8.22% for the Russell Top 200.

these cutoff values should not be used for data from a time period different than the one used here. The period July 1, 2003 to June 20, 2007 used in this paper had very low volatility according to the VIX index. The cutoff values would certainly need to be higher for a period that included the 2008-09 crisis, for instance.

Now, the issue of portfolio selection may be addressed. The portfolio consists of 20 stocks from the Russell 3000 index. Specifically, there are six stocks from the Russell Top 200, seven stocks from the Russell Midcap, and seven stocks from the Russell 2000. These 20 stocks were chosen from the previous sample of 60 such that there were a similar number of high, medium, and low volatility periods for the entire portfolio. To determine whether a day should be classified as ‘high’, ‘medium’, or ‘low’, the standard deviation of returns for that day is calculated, and then the classification is made according to the cutoff values (for example, < 0.0135 indicates ‘low’). This is repeated for every day in the sample. The sample consists of data for 20 stocks from Jan. 3, 2006 to June 28, 2007 (375 days).¹⁰ To recap, each day is classified as ‘high’, ‘medium’, or ‘low’ volatility, and there are 20 stocks, so there are 7500 ($375 \times 20 = 7500$) total classifications. Ideally, roughly one-third of these would be each of high, medium, and low. In the actual sample, 36.03% of the days are classified as ‘high’, 30.94% are ‘medium’, and 33.03% are ‘low’. These values are considered satisfactorily close to one-third for the purposes of this paper. To reiterate, the main issue is that there need to be enough days of each type of volatility to test the rules on. In other words, the primary concern is that the minimum value (in this case 30.94%) may be too low. However, 30.94% seems to be a very reasonable minimum value, considering the highest value this minimum can attain is 33.33%.

¹⁰Within the 375 day sample, there is an in-sample period of 250 days and an out-of-sample period of 125 days.

The list of stocks used in the portfolio can be found in Appendix A. The following section provides a description of the data and the universe of trading rules.

2.5 Data and Trading Rules

2.5.1 Data

The data used in this paper is one-minute-equity-data (OMED) for the 20 stocks in the portfolio from Jan. 3, 2006 to June 28, 2007. The in-sample period runs from Jan. 3, 2006 to Dec. 28, 2006 (250 trading days), while the out-of-sample period runs from Dec. 29, 2006 to June 28, 2007 (125 trading days). The data was purchased from Tick Data (www.tickdata.com). The trading rules are tested on the data at frequencies of 30, 78, 130, and 390 minutes.¹¹ The Dickey Fuller test is applied to the data to test for stationarity. The null hypothesis of nonstationarity is tested against the alternative that the data follows an AR(1) with drift. The null hypothesis of nonstationarity can be rejected for each of the 20 time series.¹²

¹¹Note that these values are chosen such that there will be an integer number of periods per trading day. There are 390 minutes in a trading day, so a frequency of 30 minutes yields 13 periods, a frequency of 78 minutes yields 5 periods, a frequency of 130 minutes yields 3 periods, and a frequency of 390 minutes yields 1 period (daily data).

¹²Results of the test are available upon request.

2.5.2 Trading Rules

Two simple rules and two complex rules are examined. The simple rules are the exponential moving average (EMA) and the KST (know-sure-thing) indicator. Before describing how the EMA rule works, the calculation of an EMA will be explained. The benefit of using an EMA instead of a simple moving average is that the EMA gives more weight to recent observations, whereas the simple MA is just an arithmetic mean of past prices. The EMA is calculated using the following equation:

$$EMA_t = EMA_{t-1} + \alpha \cdot (price_{t-1} - EMA_{t-1})$$

where $0 < \alpha < 1$ is a smoothing parameter.

Typically two EMA's will be calculated, one short and one long. In this paper, the short moving average is either one or two periods. Note that a one-period moving average is just the price of that period. In the case that the short moving average is two periods, a simple moving average is used, that is, the arithmetic mean of the past two periods. The long EMA ranges from 10 to 60 periods (in increments of 10). When the short MA crosses the long EMA from below, a buy signal is generated. Likewise, when the short MA crosses the long EMA from above, a sell signal is generated, and in some cases the rule may include an instruction to sell the stock short in such an instance. Additionally, the rule can be modified by placing trading bands around the long EMA. The trading band is comprised of an upper and lower bound on either side of the EMA. The value of the upper bound of a 1% trading band at time t is obtained using the following equation:

$$U_t = EMA_t + EMA_t \cdot 0.01$$

where U_t is the value of the upper bound at time t , and EMA_t is the value of the EMA at time t . The lower bound is calculated in a similar fashion:

$$L_t = EMA_t - EMA_t \cdot 0.01$$

Now, a buy signal is generated when the short MA crosses the upper bound of the long EMA from below, and a sell signal is generated when the short MA crosses the lower bound of the long EMA from above. The idea behind using the trading band is to minimize the effect of ‘whipsaws’. A whipsaw is defined as a price movement in one direction (either up or down), followed quickly by a movement in the other direction. In the context of the EMA rule, the investor would be worried about a situation where the short MA moved just above the long EMA, triggering a buy signal, only to immediately turn and move quickly downward, triggering a sell signal, and almost certainly a loss in profit due to the transaction costs involved.¹³ However, there is a tradeoff here; the wider the band, the longer the investor has to watch the price increase (decrease) before the asset is actually purchased (sold), so while the effect of whipsaws may be minimized, the profit on each successful trade will be smaller. In this paper, trading bands of 0.2%, 0.4%, and 0.6% are used.

The second simple trading rule is the KST (know-sure-thing) indicator. The KST was developed by Martin Pring, author of *Technical Analysis Explained* (Pring (2002)), one of the most important books written on the subject of technical analysis. The KST is made up of four ROC (rate of change) indicators, so the first step is to

¹³For large institutional investors the transaction costs would primarily be due to the bid-ask spread, while brokerage commissions would be minimal. On the other hand, individual investors may face significant brokerage fees in addition to the bid-ask spread, hence, magnifying the negative effect of a whipsaw.

define the ROC. The value of the 10-day ROC today is given by today's price divided by the price 10 days ago. For example, if today's price is 95, and the price 10 days ago was 100, then today's ROC value is 0.95. Similarly, tomorrow's ROC value will be given by tomorrow's price divided by the price 9 days ago. The formula for the general n -day ROC at time t is given by:

$$ROC_t = Price_t / Price_{t-n}.$$
¹⁴

In the KST the ROC is smoothed by a moving average, so that the value that goes into the calculation of the KST is not the value of the ROC itself, but the value of the moving average of the ROC. It was mentioned earlier that the KST is made up of four ROC's. Specifically, the version of the KST used in this paper (known as the 'daily KST') is comprised of a 10-day ROC smoothed by a 10-day MA, a 15-day ROC smoothed by a 10-day MA, a 20-day ROC smoothed by a 10-day MA, and a 30-day ROC smoothed by a 15-day MA. Each of these ROC's receives a different weight, so that the calculation of the KST is given by

$$\begin{aligned} KST_t = & 1/12 \times (10\text{-day MA of } 10\text{-day ROC})_t \\ & + 2/12 \times (10\text{-day MA of } 15\text{-day ROC})_t \\ & + 3/12 \times (10\text{-day MA of } 20\text{-day ROC})_t \\ & + 6/12 \times (15\text{-day MA of } 30\text{-day ROC})_t \end{aligned}$$

Pring (2002) provides the following rationale for using the KST indicator:

¹⁴At this point it is important to note that in this paper this ROC value is multiplied by 100 and then 100 is subtracted from the resulting product. This has the intuitively appealing effect of centering the KST around 0 instead of 100.

Price is determined by the interaction of a number of different time cycles at any given point in time. A momentum indicator that's constructed from only one time span, such as a 10-day rate of change (ROC) then, will only reflect the cycles that are close to the defined parameter. The KST indicator was intentionally created using several time cycles that build a broader visual picture of the market. By including moving averages in the formula, you get an oscillator that is smooth, does not turn prematurely, and reflects the underlying cyclical waves.

Now that the KST has been derived, the method of generating buy and sell signals can be discussed. There are a few different possibilities here; namely, moving average crossovers, or the use of overbought/oversold lines. The former involves calculating a moving average of the KST. A buy (sell) signal is generated when the KST crosses the MA from below (above). However, this paper uses the latter method exclusively. An overbought (oversold) region is meant to reflect a situation where the market value of an asset is significantly higher (lower) than its intrinsic value. Thus, when the KST crosses the overbought (oversold) line this generates a sell (buy) signal.¹⁵ The primary flaw of this method is that sometimes the KST will vastly overshoot these lines, causing large losses for an investor using this strategy. To clarify the previous statement, consider the following example. Suppose the market learns of some news that will hurt company ABC, as a result, ABC's stock price plunges. After an initial loss of 20%, the KST has crossed below the oversold line, generating a buy signal.

¹⁵These overbought/oversold lines are chosen subjectively, typically based on values that have worked well in the past. The oversold (overbought) line will be placed somewhere below (above) zero.

However, the stock continues to plunge pushing the KST further and further into oversold territory, the stock loses a further 25% and never fully recovers leading to a significant loss for the investor. It should be clear how a similar situation could arise with a stock's price moving very deep into overbought territory, and an investor deciding to short the stock as soon as the overbought line was crossed. This is clearly a significant problem, so the following modification is employed. Instead of buying (selling) when the KST crosses the oversold (overbought) line from above (below), wait for it to turn around and then buy (sell) when the KST crosses the oversold (overbought) line from below (above). This modification should prevent significant losses in such extreme examples as described above, but it is still vulnerable to whipsaws in the KST. All of the parameterizations of the KST rule presented in this paper allow for shorting.

The complex rules used in this paper are called 'complex voting' and 'complex fractional'. The inspiration for both rules comes from Hsu and Kuan (2005), although there are some differences. The voting rule takes the position (that is, long or short) of the EMA rules as inputs. Using this information, the voting rule generates a decision to take either a long or short position in the asset in question. The decision rule is simple; if the number of EMA rules that are long is larger (smaller) than the number that are short, then the investors goes long (short).

The fractional rule is just a modified voting rule. As before, if the number of EMA rules that are long is larger than the number that are short, then the investors goes long. However, the following rule is implemented:

- if the proportion of rules that are long is between 50% and 75%, then buy one

share

- if the proportion of rules that are long is greater than 75%, but less than or equal to 90%, then buy two shares
- if the proportion of rules that are long is greater than 90%, then buy three shares

An analogous rule is in place for shorting. In order to ensure clarity, it should be mentioned that all of the rules are tested on each stock individually. Then, the portfolio return for rule x is just the arithmetic average of the return from rule x applied to all 20 stocks. This concludes the description of the trading rules. In total, there are 106 total rules; the parameterizations for each rule can be found in Appendix B.¹⁶

2.6 Results

The measure used to determine profitability of the trading rules is the simple return:

$$Return_t = \frac{Price_t - Price_{t-1}}{Price_{t-1}}$$

The portfolio is equal-weighted, which means that the same amount of money is invested in each stock in the portfolio. The reason for using an equal-weighted portfolio as opposed to a dollar-weighted portfolio can be illustrated by a simple example.¹⁷

Suppose there is a portfolio of 20 stocks, call them A, B, C, . . . , T. Further, suppose

¹⁶There are ninety-six moving average rules, six KST rules, and four complex rules

¹⁷In a dollar-weighted portfolio, an equal number of shares of each security are purchased.

stock A trades at an average of \$400 over the sample period, while stocks B through T each trade at an average of \$10. It is clear that in a dollar-weighted portfolio, the profitability of the trading rules will largely be determined by how well the rules perform on stock A. Such a situation would be undesirable, thus, an equal-weighted portfolio is much more appealing. An assumption is made that whenever a trading rule generates a neutral signal on a stock, the investor instead invests at the risk free rate.¹⁸

Previous studies have produced many different estimates of transaction costs. For instance, Fama and Blume (1966) estimate the cost of a one-way trade for floor traders at 0.05%, while Chan and Lakonishok (1993) obtain an estimate of 0.13% for large institutional investors. Hsu and Kuan (2005) also use 0.05%, stating that such a cost may be applicable to large institutional investors. In this paper, transaction costs are assumed to be 0.05% for each one-way trade.

The following results are for the in-sample period, consisting of 250 days of data from Jan. 3, 2006 to Dec. 28, 2006.

2.6.1 In-Sample Results

The trading rules were applied at four different frequencies: 30 minutes, 78 minutes, 130 minutes, and 390 minutes (daily). First, the results from the daily frequency are examined. Table 2.1 shows the returns of the 10 best trading rules over the entire

¹⁸The risk free rate is the average of the 4-week T-bill quoted on a discount basis each month from Jan. 2006, to June, 2007 (the length of the sample). The average annualized value is 4.72111% (http://www.federalreserve.gov/releases/h15/data/Monthly/H15_TB_WK4.txt)

Table 2.1: Top 10 Rules for Daily Frequency - in-Sample

Rank	Rule	Return
1	KST-4	0.1754
2	KST-6	0.1605
3	KST-7	0.1385
4	KST-10	0.1025
5	CF-NS	0.0804
6	CV-NS	0.0733
7	MA20b6-NS	0.0604
8	MA50b6-NS	0.0538
9	MA20b2-NS	0.0522
10	MA10b6-NS	0.0515

Note: the values in the ‘Return’ column can be converted to percentages by multiplying them by 100. So, the annual return for the top performing rule is 17.54%. The return from the buy-and-hold strategy is 0.1315.

in-sample period.

The KST rules dominate, as they account for the four most profitable rules. The best rule, KST-4, yields a return of 17.54%, and the top three rules beat the buy-and-hold strategy which yields a return of 13.15%. Moving average rules that allow for shorting perform the worst.

Of course the purpose of the paper is to examine how the rankings change under different regimes of volatility, so let’s look at the results under high, medium, and low volatility. Table 2.2 shows the top 10 performing rules under high volatility.

The five best rules are KST rules. The return for the top rule, KST-4, is 20.79% compared to a buy-and-hold return of 21.63% over the high volatility portion of the sample period. Note that these values have been annualized. The proportion of the in-sample period that the portfolio spent in highly volatile conditions was 0.4022, so in order to annualize the returns, the returns are multiplied by $\frac{1}{0.4022}$. As you

Table 2.2: Top 10 Rules for High Volatility - in-Sample (Daily Frequency)

Rank	Rule	Return
1	KST-4	0.2079
2	KST-6	0.1922
3	KST-7	0.1698
4	KST-10	0.1571
5	KST-2	0.0793
6	CF-NS	0.0632
7	CV-NS	0.0356
8	2MA60b6-NS	0.0245
9	2MA50b6-NS	0.0229
10	MA20b0-NS	0.0179

The return from the buy-and-hold strategy is 0.2163.

can see, the annualization is calculated without compounding.¹⁹ The moving average rules that allowed for shorting performed very poorly. Table 2.3 shows the 10 most profitable rules under medium volatility.

It is a very similar story under medium volatility, with the KST rules dominating, and the moving average rules that allow for shorting performing very poorly. The buy-and-hold return under medium volatility is -9.99% , as a result, 96 of the 106 rules beat the buy-and-hold. The best rule is again KST-4, which yields a return of 21.74% . Finally, Table 2.4 shows the ranking of the 10 best trading rules under low volatility.

The rankings change quite a bit under low volatility, with the moving average and complex rules (both without shorting) leading the way. The KST rules struggle, which makes sense because a sufficient amount of volatility is required to trigger transactions. This is the first piece of evidence to suggest that certain rules may be

¹⁹It should be noted that the listed returns are risk-unadjusted.

Table 2.3: Top 10 Rules for Medium Volatility - in-Sample (Daily Frequency)

Rank	Rule	Return
1	KST-4	0.2174
2	KST-6	0.2020
3	KST-7	0.1794
4	KST-10	0.0961
5	MA50b4-S	0.0881
6	MA60b4-S	0.0852
7	MA60b2-S	0.0671
8	MA50b4-NS	0.0653
9	MA60b4-NS	0.0653
10	CV-S	0.0641

The return from the buy-and-hold strategy is -0.0999.

Table 2.4: Top 10 Rules for Low Volatility - in-Sample (Daily Frequency)

Rank	Rule	Return
1	MA30b6-NS	0.149
2	MA40b6-NS	0.148
3	MA50b6-NS	0.145
4	MA20b6-NS	0.144
5	CF-NS	0.140
6	MA60b6-NS	0.136
7	MA10b6-NS	0.134
8	MA20b2-NS	0.130
9	MA10b2-NS	0.129
10	CV-NS	0.128

The return from the buy-and-hold strategy is 0.2139.

Table 2.5: Top 10 Rules for 130 Minute Frequency - in-Sample

Rank	Rule	Return
1	MA10b0-NS	0.0514
2	MA40b0-NS	0.0513
3	MA30b0-NS	0.0493
4	MA50b0-NS	0.0491
5	MA20b0-NS	0.0488
6	CF-NS	0.0487
7	MA10b4-NS	0.0469
8	MA60b0-NS	0.0457
9	MA20b4-NS	0.0454
10	MA40b4-NS	0.0454

The return from the buy-and-hold strategy is 0.1505.

better suited for specific volatility ranges. The return of the best rule was 14.9% compared to 21.39% for the buy-and-hold.

A similar analysis was performed for the other three frequencies (30, 78, and 130 minutes). The results can be found in Tables 2.5-2.16. The results from all four frequencies were used to obtain a final average ranking for each level of volatility. For example, to obtain the average ranking for high volatility, first calculate the average return for each rule across all four frequencies. Let's start with the rule 'complex fractional - no shorting': $R_{Avg}^k = [0.0254(daily) + 0.0206(130\ min) + 0.0234(78\ min) + 0.0646(30\ min)]/4 = 0.0335$. Now, perform this calculation for each rule $k = 1, \dots, 106$ and arrange the R_{Avg}^k 's in descending order. Tables 2.17 through 2.19 show the average returns of the 10 most profitable rules for each volatility class.

Table 2.6: Top 10 Rules for High Volatility - in-Sample (130 Min. Frequency)

Rank	Rule	Return
1	MA40b0-NS	0.0991
2	MA10b0-NS	0.0956
3	MA30b0-NS	0.0923
4	MA20b0-NS	0.0889
5	MA50b0-NS	0.0871
6	MA60b0-NS	0.0785
7	CF-NS	0.0512
8	MA40b4-NS	-0.0012
9	MA10b4-NS	-0.0013
10	MA40b0-NS	-0.0018

The return from the buy-and-hold strategy is 0.2297.

Table 2.7: Top 10 Rules for Medium Volatility - in-Sample (130 Min. Frequency)

Rank	Rule	Return
1	KST-4	0.2321
2	MA20b4-S	0.1180
3	MA30b4-S	0.1158
4	MA40b4-S	0.1065
5	MA30b6-S	0.1047
6	MA60b6-S	0.1042
7	MA20b6-S	0.1040
8	KST-10	0.1032
9	KST-7	0.1032
10	MA40b6-S	0.1029

The return from the buy-and-hold strategy is -0.0712.

Table 2.8: Top 10 Rules for Low Volatility - in-Sample (130 Min. Frequency)

Rank	Rule	Return
1	SMA40b6-NS	0.1410
2	SMA50b6-NS	0.1380
3	SMA60b6-NS	0.1350
4	SMA20b6-NS	0.1320
5	SMA30b6-NS	0.1300
6	SMA10b6-NS	0.1300
7	SMA20b2-NS	0.1290
8	SMA10b2-NS	0.1280
9	SMA30b2-NS	0.1240
10	SMA40b2-NS	0.1240

The return from the buy-and-hold strategy is 0.2313.

Table 2.9: Top 10 Rules for 78 Minute Frequency - in-Sample

Rank	Rule	Return
1	KST-2	0.1290
2	MA10b6-NS	0.0673
3	MA40b6-NS	0.0630
4	MA20b6-NS	0.0626
5	MA30b6-NS	0.0624
6	MA50b6-NS	0.0597
7	MA60b6-NS	0.0592
8	KST-6	0.0481
9	MA10b4-NS	0.0458
10	2MA20b4-NS	0.0450

The return from the buy-and-hold strategy is 0.1543.

Table 2.10: Top 10 Rules for High Volatility - in-Sample (78 Min. Frequency)

Rank	Rule	Return
1	KST-2	0.1678
2	MA60b0-NS	0.1020
3	MA50b0-NS	0.1020
4	MA20b0-NS	0.0974
5	MA10b0-NS	0.0974
6	MA30b0-NS	0.0974
7	MA40b0-NS	0.0974
8	MA40b6-NS	0.0929
9	MA20b6-NS	0.0926
10	MA30b6-NS	0.0922

The return from the buy-and-hold strategy is 0.2342.

Table 2.11: Top 10 Rules for Medium Volatility - in-Sample (78 Min. Frequency)

Rank	Rule	Return
1	KST-4	0.1876
2	MA10b0-S	0.1190
3	MA30b0-S	0.1179
4	MA40b0-S	0.1179
5	MA20b0-S	0.1177
6	MA50b0-S	0.1168
7	MA60b0-S	0.1158
8	MA10b0-NS	0.1111
9	MA20b0-NS	0.1090
10	KST-6	0.1085

The return from the buy-and-hold strategy is -0.0623.

Table 2.12: Top 10 Rules for Low Volatility - in-Sample (78 Min. Frequency)

Rank	Rule	Return
1	KST-2	0.1528
2	2MA60b6-NS	0.1008
3	2MA40b6-NS	0.0999
4	2MA30b6-NS	0.0978
5	2MA50b6-NS	0.0970
6	2MA10b6-NS	0.0940
7	2MA20b6-NS	0.0928
8	2MA30b4-NS	0.0478
9	KST-10	0.0472
10	KST-7	0.0472

The return from the buy-and-hold strategy is 0.2307.

Table 2.13: Top 10 Rules for 30 Minute Frequency - in-Sample

Rank	Rule	Return
1	KST-4	0.1604
2	KST-7	0.1386
3	KST-6	0.0726
4	MA10b6-NS	0.0706
5	MA40b6-NS	0.0682
6	MA30b6-NS	0.0680
7	MA20b6-NS	0.0679
8	KST-10	0.0675
9	MA50b6-NS	0.0673
10	MA60b6-NS	0.0672

The return from the buy-and-hold strategy is 0.1443.

Table 2.14: Top 10 Rules for High Volatility - in-Sample (30 Min. Frequency)

Rank	Rule	Return
1	KST-4	0.2849
2	KST-7	0.2556
3	CF-NS	0.1606
4	2MA10b2-NS	0.1455
5	MA10b6-NS	0.1435
6	MA40b6-NS	0.1425
7	2MA20b2-NS	0.1423
8	MA50b6-NS	0.1422
9	MA60b6-NS	0.1422
10	MA20b6-NS	0.1397

The return from the buy-and-hold strategy is 0.2198.

Table 2.15: Top 10 Rules for Medium Volatility - in-Sample (30 Min. Frequency)

Rank	Rule	Return
1	KST-6	0.1289
2	2MA20b6-S	0.0857
3	2MA30b6-S	0.0850
4	2MA40b6-S	0.0850
5	2MA50b6-S	0.0850
6	2MA60b6-S	0.0835
7	MA10b4-S	0.0797
8	MA10b4-NS	0.0771
9	KST-7	0.0754
10	KST-10	0.0750

The return from the buy-and-hold strategy is -0.0705.

Table 2.16: Top 10 Rules for Low Volatility - in-Sample (30 Min. Frequency)

Rank	Rule	Return
1	KST-4	0.0848
2	KST-6	0.0674
3	2MA10b6-NS	0.0476
4	2MA20b6-NS	0.0476
5	KST-10	0.0472
6	KST-7	0.0472
7	2MA30b6-NS	0.0457
8	2MA60b6-NS	0.0454
9	2MA40b6-NS	0.0454
10	2MA50b6-NS	0.0454

The return from the buy-and-hold strategy is 0.2251.

Table 2.17: Top 10 Rules for High Volatility - in-Sample (Average Return Over all Frequencies)

Rank	Rule	Average Return
1	CF-NS	0.0833
2	MA20b6-NS	0.0540
3	MA10b6-NS	0.0517
4	MA20b0-NS	0.0473
5	KST-7	0.0471
6	MA50b6-NS	0.0468
7	MA30b6-NS	0.0429
8	MA60b6-NS	0.0409
9	MA40b6-NS	0.0408
10	MA20b4-NS	0.0404

Table 2.18: Top 10 Rules for Medium Volatility - in-Sample (Average Return Over all Frequencies)

Rank	Rule	Average Return
1	KST-4	0.1745
2	KST-6	0.1150
3	KST-7	0.1006
4	KST-10	0.0877
5	MA20b4-S	.0801
6	MA50b4-S	0.0799
7	MA10b4-S	0.0795
8	MA30b4-S	0.0785
9	MA60b4-S	0.0781
10	MA40b4-S	0.0780

Table 2.19: Top 10 Rules for Low Volatility - in-Sample (Average Return Over all Frequencies)

Rank	Rule	Average Ranking
1	2MA40b6-NS	0.0879
2	2MA50b6-NS	0.0857
3	2MA20b6-NS	0.0847
4	2MA30b6-NS	0.0839
5	2MA60b6-NS	0.0816
6	2MA10b6-NS	0.0802
7	KST-4	0.0693
8	KST-6	0.0627
9	MA30b6-NS	0.0617
10	MA40b6-NS	0.0609

Main Features of in-Sample Results

i) Complex

The complex fractional rule without shorting is the most profitable rule under high volatility, but it is only an average rule under ‘medium’ and ‘low’. This result is in stark contrast to the ‘with shorting’ counterpart which is one of the worst rules regardless of volatility. In general, it is expected that rules allowing shorting will perform poorly since the portfolio trends upwards over the majority of the sample period (see Figure 2.1), thus making profiting off of short sales more difficult.

The complex voting rule without shorting is reasonably consistent across volatility ranges (high - 12th, medium - 42nd, low - 14th), although similar to the complex fractional rule its best results are under high volatility. Once again, the version that allows for shorting performs uniformly worse. Thus, the complex rules appear to excel in highly volatile conditions.

ii) KST

The KST rule performs very well overall. The results are best under medium volatility, followed by ‘high’ and ‘low’. This result roughly conforms with intuition since the KST rule is designed to capture over-reactions, which are commonly accompanied by large volume, although it is surprising to see that it performs better under medium rather than high volatility. A more accurate statement may be that the KST rule requires some threshold amount of volatility to generate signals, but once this threshold has been reached, more volatility isn’t necessarily better.

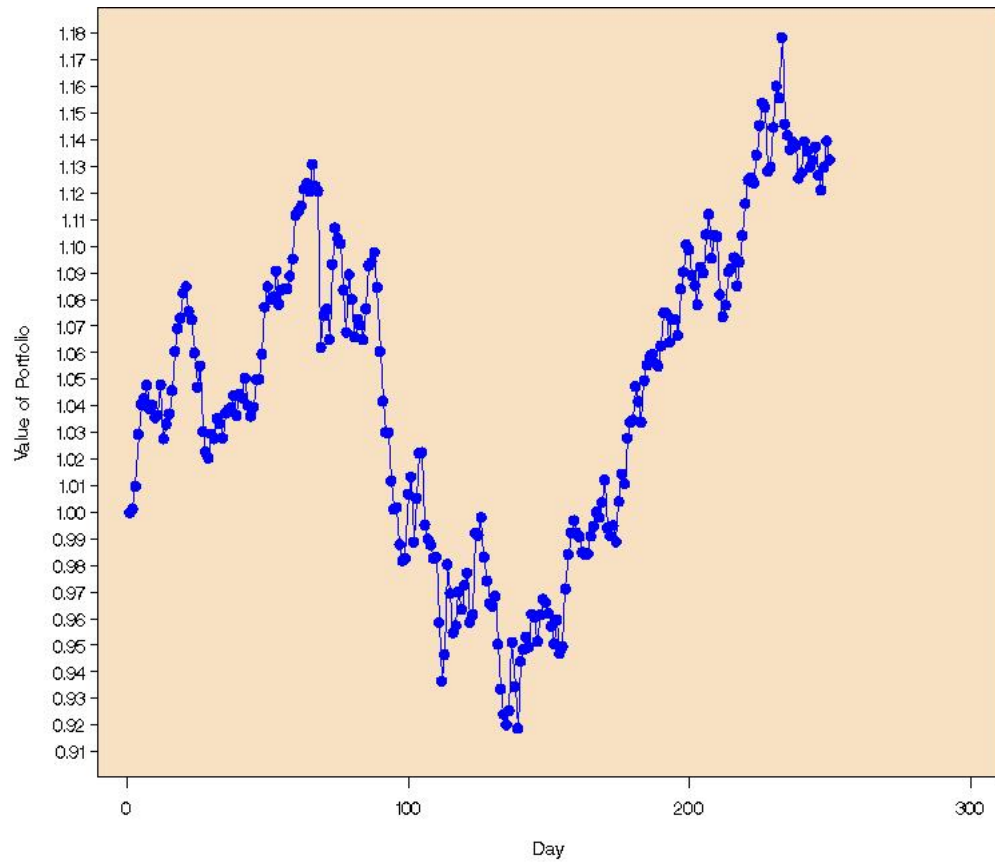


Figure 2.1: Value of Portfolio - in-Sample Period (Assuming 1 dollar is invested in the portfolio on day 1)

An interesting aspect of the results is that the KST rule performs much better on daily data than any other frequency. This should be of no surprise since the rule was parameterized specifically for daily data by the developer Martin Pring. The rule may very well yield consistently significant returns on higher frequency data if re-parameterized in an intelligent way.

Finally, note that the moderate parameterizations (i.e. 4, 6, and 7%) perform much better, in general, than the extreme parameterizations (1, 2, and 10%). The intuition is that the 1 and 2% rules generate far too many signals, and hence, too many ‘whipsaw’ transactions. On the other hand, the 10% rule generates very few transactions. For example, for the daily data, zero buy/sell signals are generated on 17 of the 20 stocks.

iii) Moving Average

- For the moving average rules, the ‘without shorting’ strategies perform much better than those that allow shorting under high and low volatility, while the two versions yield similar results under ‘medium’.
- The rules that have a short MA of 1 perform significantly better than those with a short MA of 2 under high and medium volatility, but the results are similar for ‘low’.
- Finally, with reference to the trading bands, it appears that the smaller bands

(0% and 0.2%) excel under low volatility, while the largest band (0.6%) struggles. There are no clear winners or losers under medium volatility. Under high volatility, the smaller bands (0% and 0.2%) struggle, while the larger bands (0.4% and 0.6%) perform fairly well. These results conform with the intuition behind such trading bands. A small band with high volatility yields too many ‘whipsaw’ transactions, most often resulting in small losses. A larger band gets rid of many of the ‘whipsaws’ experienced under high volatility; however, it sacrifices too much in terms of timing to be profitable in a low volatility environment.

Thus, the best MA rule appears to be one that i) does not allow short selling, ii) has a short MA of one period, and iii) has a trading band such that the width of the band is proportional to the volatility of the asset.

Frequency

Returns for the top rules are higher for the daily and 30 minute frequencies. These are the only frequencies where some rules beat the buy-and-hold. There does not appear to be any pattern in regards to certain trading rules performing better or worse for different frequencies.

2.6.2 Out-of-Sample Results

Before presenting the results for the out-of-sample period, the issue of forecasting the following period's volatility needs to be addressed. The investor needs to know how volatile the following period will be so that he may choose the appropriate trading rule. There is a vast amount of literature dedicated to volatility forecasting and some sophisticated methods have been developed (see, for example Andersen, Bollerslev, Diebold, and Labys (2003)); however, the task at hand is relatively simple as only the *range* in which the next period's volatility will fall needs to be forecasted. A simple method is used whereby the volatility range at time $t + 1$ is assumed to be equal to the known range at time t . The known range at time t is determined by calculating a 50-day moving average of the standard deviation of returns, and then classifying the resulting value as 'high', 'medium', or 'low' based on the cut-off values determined earlier.²⁰ In the out-of-sample data, this forecasting method predicts the correct range 96.72% of the time.²¹ Thus, a more sophisticated method does not appear to be necessary. Now, the out-of-sample results at the daily frequency will be presented.

Overall, the KST rules dominate, accounting for five of the top six performing rules. The best rule, KST-4, yields a return of 15.2%. The moving average, and complex rules that allow for shorting perform poorly. The buy-and-hold return is 34.1%. Again, keep in mind that all of the returns presented in this section have been annualized. The 10 best rules are displayed in Table 2.20.

²⁰The 50-day moving average decreases the number of transitions between volatility regimes (relative to not smoothing with a moving average).

²¹The method is applied to each stock in the portfolio individually. 96.72% is the average success rate over all 20 stocks. The minimum success rate for any stock in the portfolio is 90.4%.

Table 2.20: Top 10 Rules for Daily Frequency - out-of-Sample

Rank	Rule	Return
1	KST-4	0.1510
2	KST-6	0.1168
3	KST-7	0.1044
4	KST-10	0.1034
5	SMA60b6-NS	0.0624
6	KST-2	0.0562
7	MA60b2-NS	0.0523
8	2MA60b0-NS	0.0514
9	2MA50b6-NS	0.0498
10	MA60b6-NS	0.0494

The return from the buy-and-hold strategy is 0.341.

For high volatility, the KST rules are the top performers. There is a huge difference, in terms of returns, between the top four KST rules and the other 102 rules. The best rule, KST-4, yields a return of 26%, compared to the buy-and-hold which yields a massive return of 70.14%. The top 10 rules are displayed in Table 2.21.

There is no clear winner under medium volatility. The top rule, a moving average rule that does not allow shorting, yields a return of 14.77%, compared to the buy-and-hold return of 26.79%. The only clear trend, which is true regardless of volatility, is that rules that allow for shorting are significantly inferior to their ‘no shorting’ counterparts. The top 10 rules are displayed in Table 2.22.

There is a similar story for low volatility, where once again there is no clear winner. The KST and moving average rules perform fairly well. The best rule is KST-4 which yields a return of 16.64%, compared to the buy-and-hold which yields a return of 9.67%. Note that this is the first time in the out-of-sample data that a trading rule beat the benchmark. In fact, six rules beat the benchmark buy-and-hold

Table 2.21: Top 10 Rules for High Volatility - out-of-Sample (Daily Frequency)

Rank	Rule	Return
1	KST-6	0.2600
2	KST-10	0.2060
3	KST-7	0.2000
4	KST-4	0.1400
5	2MA20b6-NS	0.0362
6	2MA10b6-NS	0.0358
7	2MA30b6-NS	0.0327
8	2MA20b0-NS	0.0242
9	2MA20b4-NS	0.0227
10	2MA40b6-NS	0.0158

The return from the buy-and-hold strategy is 0.7014.

Table 2.22: Top 10 Rules for Medium Volatility - out-of-Sample (Daily Frequency)

Rank	Rule	Return
1	MA60b2-NS	0.1477
2	KST-4	0.1466
3	MA60b0-NS	0.1347
4	CF-NS	0.1325
5	MA60b4-NS	0.1291
6	MA50b2-NS	0.1284
7	MA20b0-NS	0.1214
8	MA60b6-NS	0.1210
9	MA40b0-NS	0.1148
10	MA50b4-NS	0.1105

The return from the buy-and-hold strategy is 0.2679.

Table 2.23: Top 10 Rules for Low Volatility - out-of-Sample (Daily Frequency)

Rank	Rule	Return
1	KST-4	0.149
2	2MA20b6-NS	0.1212
3	2MA30b6-NS	0.1212
4	2MA30b6-NS	0.1212
5	2MA60b6-NS	0.1212
6	2MA50b6-NS	0.1142
7	2MA10b6-NS	0.0890
8	KST-2	0.0883
9	2MA20b2-NS	0.0786
10	2MA20b6-S	0.0775

The return from the buy-and-hold strategy is 0.0967.

strategy under low volatility. The 10 best rules are displayed in Table 2.23.

A similar analysis was performed for the other three frequencies (30, 78, and 130 minutes). The results can be found in Tables 2.24-2.35. As in the in-sample results section, the results from all four frequencies were used to obtain a final average ranking for each level of volatility. Tables 2.36 through 2.38 show the average returns of the 10 most profitable rules for each volatility class.

Table 2.24: Top 10 Rules for 130 Min. Frequency - out-of-Sample

Rank	Rule	Return
1	2MA10b4-NS	0.2770
2	2MA60b4-NS	0.2749
3	KST-7	0.2745
4	KST-10	0.2736
5	2MA50b4-NS	0.2734
6	2MA40b4-NS	0.2728
7	2MA20b4-NS	0.2726
8	2MA50b2-NS	0.2722
9	2MA60b2-NS	0.2695
10	2MA30b2-NS	0.2689

The return from the buy-and-hold strategy is 0.3274.

Table 2.25: Top 10 Rules for High Volatility - out-of-Sample (130 Min. Frequency)

Rank	Rule	Return
1	2MA10b4-NS	0.5325
2	2MA10b2-NS	0.5315
3	2MA30b4-NS	0.5278
4	2MA40b2-NS	0.5270
5	2MA50b2-NS	0.5270
6	2MA20b4-NS	0.5261
7	2MA60b4-NS	0.5244
8	2MA40b4-NS	0.5232
9	2MA30b2-NS	0.5226
10	2MA20b2-NS	0.5225

The return from the buy-and-hold strategy is 0.6801.

Table 2.26: Top 10 Rules for Medium Volatility - out-of-Sample (130 Min. Frequency)

Rank	Rule	Return
1	KST-4	0.3215
2	KST-2	0.2902
3	MA60b6-NS	0.2880
4	2MA20b0-NS	0.2866
5	2MA40b2-NS	0.2775
6	2MA50b0-NS	0.2760
7	2MA10b2-NS	0.2743
8	2MA10b0-NS	0.2735
9	MA50b6-NS	0.2715
10	SMA60b2-NS	0.2715

The return from the buy-and-hold strategy is 0.2553.

Table 2.27: Top 10 Rules for Low Volatility - out-of-Sample (130 Min. Frequency)

Rank	Rule	Return
1	KST-4	0.2093
2	KST-2	0.0986
3	2MA30b4-NS	0.0756
4	2MA10b4-NS	0.0748
5	2MA20b4-NS	0.0748
6	2MA50b4-NS	0.0744
7	2MA60b4-NS	0.0744
8	2MA40b4-NS	0.0736
9	2MA50b6-NS	0.0497
10	KST-10	0.0474

The return from the buy-and-hold strategy is 0.0890.

Table 2.28: Top 10 Rules for 78 Min. Frequency - out-of-Sample

Rank	Rule	Return
1	2MA20b4-NS	0.2710
2	2MA10b4-NS	0.2660
3	2MA60b4-NS	0.2660
4	2MA50b4-NS	0.2650
5	2MA40b4-NS	0.2630
6	2MA30b4-NS	0.2620
7	2MA30b6-NS	0.2600
8	2MA20b6-NS	0.2590
9	2MA50b6-NS	0.2560
10	2MA60b6-NS	0.2560

The return from the buy-and-hold strategy is 0.325.

Table 2.29: Top 10 Rules for High Volatility - out-of-Sample (78 Min. Frequency)

Rank	Rule	Return
1	2MA20b4-NS	0.5710
2	2MA30b6-NS	0.5680
3	2MA20b6-NS	0.5670
4	2MA10b4-NS	0.5630
5	2MA30b4-NS	0.5630
6	2MA40b6-NS	0.5600
7	2MA50b4-NS	0.5560
8	2MA60b4-NS	0.5560
9	2MA40b4-NS	0.5520
10	2MA50b6-NS	0.5500

The return from the buy-and-hold strategy is 0.6760.

Table 2.30: Top 10 Rules for Medium Volatility - out-of-Sample (78 Min. Frequency)

Rank	Rule	Return
1	KST-1	0.2563
2	2MA20b4-NS	0.2265
3	2MA50b4-NS	0.2253
4	2MA60b4-NS	0.2250
5	2MA10b4-NS	0.2232
6	2MA40b4-NS	0.2226
7	2MA30b4-NS	0.2124
8	2MA50b6-NS	0.2015
9	2MA60b6-NS	0.2011
10	2MA10b6-NS	0.1997

The return from the buy-and-hold strategy is 0.2533.

Table 2.31: Top 10 Rules for Low Volatility - out-of-Sample (78 Min. Frequency)

Rank	Rule	Return
1	2MA10b6-NS	0.0507
2	2MA20b6-NS	0.0506
3	2MA30b6-NS	0.0506
4	2MA40b6-NS	0.0506
5	2MA50b6-NS	0.0506
6	2MA60b6-NS	0.0506
7	KST-10	0.0474
8	KST-6	0.0474
9	2MA20b4-NS	0.0468
10	2MA60b4-NS	0.0468

The return from the buy-and-hold strategy is 0.0883.

Table 2.32: Top 10 Rules for 30 Min. Frequency - out-of-Sample

Rank	Rule	Return
1	2MA40b6-NS	0.2189
2	2MA50b6-NS	0.2189
3	2MA60b6-NS	0.2180
4	2MA10b6-NS	0.2165
5	2MA20b6-NS	0.2165
6	2MA30b6-NS	0.2165
7	2MA60b4-NS	0.1683
8	2MA20b4-NS	0.1672
9	2MA50b4-NS	0.1662
10	2MA30b4-NS	0.1659

The return from the buy-and-hold strategy is 0.3212.

Table 2.33: Top 10 Rules for High Volatility - out-of-Sample (30 Min. Frequency)

Rank	Rule	Return
1	KST-2	0.4267
2	2MA40b6-NS	0.3936
3	2MA50b6-NS	0.3936
4	2MA60b6-NS	0.3907
5	2MA10b6-NS	0.3880
6	2MA20b6-NS	0.3880
7	2MA30b6-NS	0.3880
8	2MA20b4-NS	0.3871
9	2MA30b4-NS	0.3871
10	2MA40b4-NS	0.3808

The return from the buy-and-hold strategy is 0.6733.

Table 2.34: Top 10 Rules for Medium Volatility - out-of-Sample (30 Min. Frequency)

Rank	Rule	Return
1	2MA40b6-NS	0.2166
2	2MA50b6-NS	0.2166
3	2MA60b6-NS	0.2166
4	2MA10b6-NS	0.2149
5	2MA20b6-NS	0.2149
6	2MA30b6-NS	0.2149
7	2MA60b4-NS	0.1185
8	2MA40b6-S	0.1156
9	2MA50b6-S	0.1155
10	2MA60b6-S	0.1155

The return from the buy-and-hold strategy is 0.2477.

Table 2.35: Top 10 Rules for Low Volatility - out-of-Sample (30 Min. Frequency)

Rank	Rule	Return
1	2MA10b6-NS	0.0584
2	2MA20b6-NS	0.0584
3	2MA30b6-NS	0.0584
4	2MA40b6-NS	0.0584
5	2MA50b6-NS	0.0584
6	2MA60b6-NS	0.0584
7	MA10b6-NS	0.0516
8	MA20b6-NS	0.0516
9	MA30b6-NS	0.0516
10	MA40b6-NS	0.0516

The return from the buy-and-hold strategy is 0.0851.

Table 2.36: Top 10 Rules for High Volatility - out-of-Sample (Average Return Over all Frequencies)

Rank	Rule	Average Return
1	2MA20b4-NS	0.0550
2	2MA30b4-NS	0.0543
3	2MA20b6-NS	0.0530
4	2MA60b4-NS	0.0527
5	2MA30b6-NS	0.0525
6	2MA10b6-NS	0.0522
7	2MA10b4-NS	0.0519
8	2MA40b6-NS	0.0517
9	2MA50b4-NS	0.0513
10	2MA40b4-NS	0.0512

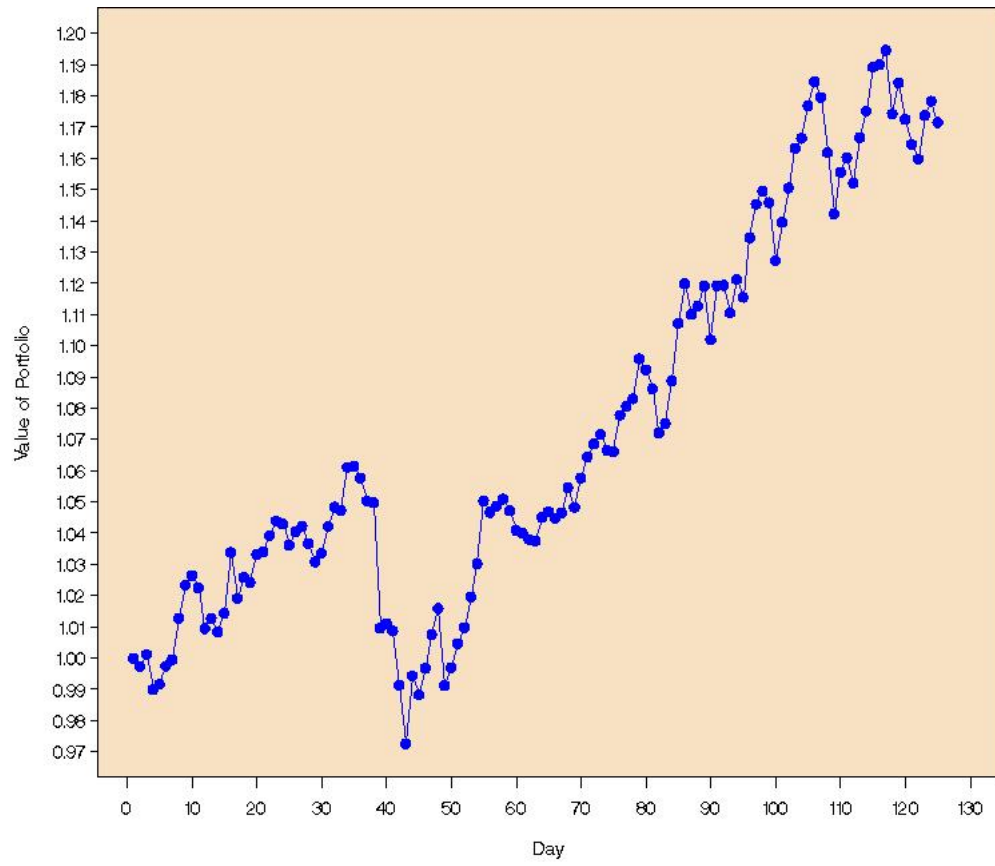


Figure 2.2: Value of Portfolio - out-of-Sample Period (Assuming 1 dollar is invested in the portfolio on day 1)

Table 2.37: Top 10 Rules for Medium Volatility - out-of-Sample (Average Return Over all Frequencies)

Rank	Rule	Average Return
1	2MA60b6-NS	0.0347
2	2MA50b6-NS	0.0343
3	2MA60b4-NS	0.0330
4	2MA50b4-NS	0.0330
5	2MA40b4-NS	.0325
6	KST-4	0.0316
7	2MA40b6-NS	0.0314
8	MA60b6-NS	0.0308
9	2MA30b6-NS	0.0302
10	2MA10b4-NS	0.0300

Main Features of out-of-Sample Results

i) Complex

Both the voting and fractional versions of the rule performed slightly better than average in each volatility range. The result is slightly disappointing given the strong in-sample performance under high volatility. The portfolio spends the vast majority of the out-of-sample period trending upwards, (see Figure 2.2) so it is no surprise that the ‘without shorting’ rules dominate the ‘with shorting’ rules. In fact, the ‘with shorting’ rules are among the worst.

ii) KST

The KST rules perform fairly well in every volatility range. Surprisingly, the best rule under low volatility is KST-4. In general, the KST rules perform roughly the

Table 2.38: Top 10 Rules for Low Volatility - out-of-Sample (Average Return Over all Frequencies)

Rank	Rule	Average Ranking
1	KST-4	0.0171
2	2MA40b6-NS	0.0107
3	2MA30b6-NS	0.0107
4	2MA50b6-NS	0.0107
5	2MA20b6-NS	0.0107
6	2MA60b6-NS	0.0105
7	2MA10b6-NS	0.0094
8	2MA20b4-NS	0.0089
9	2MA40b4-NS	0.0088
10	2MA60b4-NS	0.0087

same in each volatility range. This is slightly counter-intuitive, as high volatility is typically necessary to generate signals. A point that is worth reiterating is that the version of the KST rule used in this study was specifically developed for daily data. Remember, when the rules were tested on daily data in the out-of-sample period, they performed much better under high volatility compared to medium or low. Thus, the counter-intuitive results from the higher frequency data may be a by-product of a sub-optimally parameterized rule.

Similar to the in-sample results, the moderate parameterizations (4, 6, and 7%) outperform the extreme parameterizations (1, 2, and 10%).

iii) Moving Average

- For the moving average rules, in all three volatility ranges, the ‘without shorting’ strategies perform much better than those that allow shorting.
- The rules that have a short MA of 2 perform significantly better than those with a short MA of 1 regardless of the amount of volatility.
- Finally, with reference to the trading bands, it appears that the wider bands (0.4% and 0.6%) excel regardless of the amount of volatility. This differs from the in-sample results where the wider bands were more profitable under high volatility, but the smaller bands excelled under ‘low’. Looking at the trend of the entire portfolio in Figure 2.2 gives some idea as to why this may be the case. Notice that as the portfolio trends upwards, there are many minor corrections downwards followed quickly by a sharp move to the upside. It is likely that the smaller bands sold off on many of these minor corrections and ended up losing money due to the quick recovery to the upside. On the other hand, since the wider bands require a more significant movement in price to trigger a buy/sell signal, they likely sold off on far fewer of the downward corrections.

Frequency

Interestingly, the higher frequency data yields better results for the top trading rules. The 130 and 78 minute frequencies lead the way, with their top rules yielding

returns of 27.7 and 27.1%, respectively. The 30 minute frequency is next (top rule - 21.9%), and the daily frequency performs the worst (top rule - 15.1%). This is the complete opposite of the in-sample results where the ordering was reversed. Thus, there may be some benefit to trading with higher frequency data, although there is no clear answer from this study.

2.6.3 Seasonality

An issue that needs to be addressed is that of seasonality. It has been well documented that opening and closing prices (within, say, 30 minutes of the open and close) are not typical when compared to overall price movements (see, for example, Dacorogna et al. (2001) and Andersen and Bollerslev (1997)). Specifically, volatility tends to be higher near the open and the close. The first step is to see what kind of volatility pattern is exhibited by the data.²² It turns out that each stock has a fairly similar pattern, on average, with high volatility at the open followed by decreasing volatility throughout most of the day, and then an increase in volatility near the close.²³ The volatility at the close is lower than the volatility at the open. If the first and last observations are removed, then the volatility gradually declines in a roughly

²²Trading volume will be used as a measurement of volatility, and the terms ‘volume’ and ‘volatility’ will be used interchangeably in this section.

²³For each stock, the average volume of trading per time period was calculated using the in-sample data, and patterns were discerned from these averages.

linear fashion throughout the day.²⁴ Thus, the seasonality effects seem to have been removed. Now the concern is whether the results are significantly altered by the removal of the data points. Fortunately, the ranking of rules is not significantly altered at any frequency. As should be expected, there are some changes in the rankings, but all of the general results from Sections 2.6.1 and 2.6.2 still hold. As a result, the issue of seasonality can be put to sleep.

2.6.4 Final Thoughts on Results

It is rare to see a randomly chosen portfolio yield an annualized return of roughly 33%.²⁵ One can conclude, merely by visual inspection of Figure 2.2, that it would be very difficult to create a technical trading rule that could beat the buy-and-hold strategy for such a portfolio. The main goal of trading rules is the same as that of investors as a whole; namely, buy during the valleys and sell at the peaks (buy low, sell high). However, when the asset in question primarily trends upwards and there aren't many well pronounced valleys, it is difficult to beat buying and holding. In addition, it is difficult to say anything concrete regarding the rankings in such an

²⁴With the 30 minute data, the first and last observation of the day were removed. In contrast, for the 78 and 130 minute data only the last observation was removed. Remember, the observations are taken from the close of each interval, so for the 78 (130) minute frequency the first observation of the day occurs 78 (130) minutes into the day, and as a result should be unaffected by the volatility near the open.

²⁵To give some perspective, the S&P 500 Index yielded an annualized return of 13.34% over the same period. However, the S&P 500 has a smaller beta than the portfolio used in this study.

atypical environment. The out-of-sample results may very well hold true in other steep bull markets, but the in-sample rankings may be more suitable for a more typical environment.

2.7 Extensions

2.7.1 Should the Choice of Trading Rule Be Dependent on Volatility?

A simple experiment is performed to shed some light on the question as to whether the choice of trading rule should be dependent on volatility. The experiment works as follows. First, fix the frequency. For this example, the 130 minute data is used. Take the best rule under high volatility in the in-sample period, and apply it to the out-of-sample data when the following period is predicted to be highly volatile. Now, do the same thing for the best ‘medium’ and ‘low’ rule in the in-sample data. Calculate the return generated by this combination of rules, call it R_{comb} . Next, take the best rule overall from the in-sample period (still at the same data frequency) and apply it to the entire out-of-sample period, not paying attention to volatility. Calculate its return, call it R_{ind} . If $R_{comb} > R_{ind}$, this indicates that volatility should be taken into consideration when choosing which trading rule to use. In this example, for the 130 minute frequency, $R_{comb} = 26\%$ and $R_{ind} = 16.78\%$.

This procedure is repeated for all four frequencies. The average returns for R_{comb}

and R_{ind} are 9.84% and 8.57% respectively. Thus, the results of the simple experiment indicate that investors should consider the interaction between the profitability of trading rules and volatility when building a technical trading system.

2.7.2 What Can Be Learned by Analyzing the Three ‘Sub-Portfolios’ Separately?

Here, the stocks are divided into three groups depending on which Russell index they were part of (Top 200, Midcap, or 2000), and the trading rules are tested on these ‘sub-portfolios’. This section attempts to answer the following two questions:

- 1) Do certain rules excel when applied to specific sub-portfolios?
- 2) Are the best trading rules out of the universe as a whole more successful on one sub-portfolio relative to the others?

Tables 2.39 and 2.40 show the number of trading rules that beat the benchmark for the entire portfolio, as well as each of the three sub-portfolios, for each frequency. Table 2.39 shows the results for the in-sample data, and Table 2.40 shows the results for the out-of-sample data. First, question 1 is addressed for the in-sample data. Table 2.39 indicates that 9 trading rules beat the benchmark for the Midcap portfolio over all four frequencies, while 11 rules beat the benchmark for the Russell 2000 portfolio.

Table 2.39: Number of Rules That Beat the Benchmark - in-Sample

	Entire Portfolio	Top 200	Midcap	2000
30 min	2	0	3	3
78 min	0	0	3	1
130 min	0	0	0	0
Daily	3	0	3	7

‘2’ in the top-left corner indicates that two trading rules beat the benchmark when applied to the entire portfolio at the 30 minute frequency. Likewise, ‘7’ in the bottom-right corner indicates that 7 trading rules beat the benchmark when applied to the Russell 2000 sub-portfolio at the daily frequency.

For the Midcap portfolio 7 out of the 9 rules were KST, and for the Russell 2000 9 out of 11 were KST. As for the out-of-sample data, Table 2.40 shows that 35 rules beat the benchmark for the Top 200 portfolio over all four frequencies, while four rules beat the benchmark for the Russell 2000 portfolio. 34 of the 35 profitable rules for the Top 200 portfolio are moving average rules, while all four of the profitable rules under the Russell 2000 are KST. Thus, a fairly obvious pattern has emerged; the KST rules appear to be better suited to the Midcap and Russell 2000 portfolios, while the moving average rules excel when applied to the Top 200 portfolio. Remember that, on average, the Russell 2000 sub-portfolio was more volatile than the Midcap, which, in turn, was more volatile than the Top 200. As a result, these results reinforce what was seen earlier; that the KST rules excel under high and medium volatility, while moving average rules excel under low volatility.

Question 2 looks at the performance of the best trading rules in ‘young’ (Russell 2000) vs. ‘mature’ (Russell Top 200) markets. The in-sample data would indicate that, in general, trading rules are more successful on stocks with smaller market capitalizations, since several rules beat the benchmark under each of the Russell Midcap

Table 2.40: Number of Rules That Beat the Benchmark - out-of-Sample

	Entire Portfolio	Top 200	Midcap	2000
30 min	0	0	0	0
78 min	0	7	0	2
130 min	0	25	0	1
Daily	0	3	0	1

‘0’ in the top-left corner indicates that zero trading rules beat the benchmark when applied to the entire portfolio at the 30 minute frequency. Likewise, ‘1’ in the bottom-right corner indicates that one trading rule beat the benchmark when applied to the Russell 2000 sub-portfolio at the daily frequency.

and Russell 2000 portfolios. The markets for large-cap stocks appear to be efficient, since none of the rules applied to the Russell Top 200 were able to beat the benchmark.²⁶ However, the out-of-sample data tells a slightly different story. Four rules beat the benchmark (over all frequencies) for the Russell 2000 portfolio, confirming the results of the in-sample data. On the other hand, 35 rules beat the benchmark for the Russell Top 200 portfolio, contradicting the conclusion from the in-sample data that the markets for large-cap stocks must be efficient.²⁷ There certainly appear to

²⁶The form of efficiency referred to in this section is Weak-form Efficiency, which posits that all publicly available information regarding an asset is already priced-in. The implication is that future prices cannot be predicted by analyzing historical price data, i.e., technical trading rules should not be able to beat buying and holding after properly accounting for data mining bias.

²⁷On the surface, it seems the rules performed much better on the Top 200 portfolio than the Russell 2000 portfolio since there were 35 profitable rules compared to four. First of all, remember that the moving average rules excelled on the Top 200 portfolio, while the KST rules excelled on the Russell 2000 portfolio. Next, note that there were many more parameterizations for the moving average rule (96) than the KST rule (6). Had there been 96 parameterizations for the KST rule (say for example 0.5%, 0.6%, . . . , 9.9%, 10%) there would almost certainly have been many more than four profitable rules for the Russell 2000 portfolio.

be inefficiencies in the markets for the smaller-cap stocks allowing for trading rules to generate significant profits. However, nothing definitive can be said regarding the efficiency of the markets for large-cap stocks.

2.8 Data Mining

The rules that were found to be profitable (i.e. better than the benchmark) in Section 2.6 are subjected to the Step-SPA test in an effort to determine whether they beat the benchmark by chance, or rather, if they have superior predictive ability. The in-sample results are displayed in Table 2.41 and the out-of-sample results in Table 2.42.

Table 2.41: Number of Rules with Superior Predictive Ability - in-Sample

	High	Medium	Low	Total
30 min	0/2	0/86	0/0	0/2
78 min	0/0	30/99	0/0	0/0
130 min	1/1	56/98	0/0	0/0
Daily	0/0	75/96	0/0	3/3

‘0/2’ in the top-left cell indicates that two rules were more profitable than the benchmark under high volatility at the 30 minute frequency, but zero rules were rejected in the Step-SPA test.

In-Sample

For the daily data, all three rules that beat the buy-and-hold overall were rejected in the Step-SPA test, indicating superior predictive ability.²⁸ However, for the 30

²⁸The language may be a bit confusing here. Remember the null hypothesis of the Step-SPA test

Table 2.42: Number of Rules with Superior Predictive Ability - out-of-Sample

	High	Medium	Low	Total
30 min	0/0	0/0	0/0	0/0
78 min	0/0	0/1	0/0	0/0
130 min	0/0	0/18	2/2	0/0
Daily	0/0	0/0	6/6	0/0

'0/1' in the second row indicates that one rule was more profitable than the benchmark under medium volatility at the 78 minute frequency, but zero rules were rejected in the Step-SPA test.

minute data, both rules that beat the benchmark were deemed to have done so by chance. Also notice that for the lower frequency data, many of the profitable rules under medium volatility were rejected. Although, as the frequency of data increases, fewer and fewer rules get rejected, in fact, zero out of 86 profitable rules are rejected for the 30 minute data, compared to 75 out of 96 for daily data.

Out-of-Sample

The out-of-sample results are easy to summarize, as all 8 profitable rules at low volatility were deemed to have superior predictive ability, while all 19 profitable rules under medium volatility were shown to have no predictive ability.

is that the rules are unprofitable, so rejection indicates a profitable rule.

2.9 Conclusions

The in-sample results indicate that the KST and complex rules (especially ‘complex fractional’) are the strongest performers under high volatility, while the KST rules excel under medium, and the moving average rules with narrow trading bands (0 and 0.2%) yield the largest returns under low volatility. The out-of-sample results at the daily frequency, show the KST rules dominating all other rules under high volatility, and also performing very well under ‘medium’ and ‘low’. However, the other frequencies paint a different picture, with the moving average rules (wide trading bands) dominating across the board.

In the in-sample data, several rules beat the benchmark buy-and-hold strategy overall, as well as under high and medium volatility. After facing the scrutiny of the Step-SPA test for data mining, three KST rules were deemed to have superior predictive ability at the daily frequency. Many rules were found to have SPA under medium volatility, at several different frequencies, while one such rule (complex fractional) held up under high volatility (at the 130 minute frequency). There was less success in the out-of-sample data, due in large part to the exceptional buy-and-hold returns. Here, several rules were more profitable than the buy-and-hold under medium and low volatility, but only the low volatility rules passed the data mining test.

There appears to be evidence that the ranking of trading rules in terms of profitability is related to volatility. In most cases, the KST rules performed very well under high and medium volatility, while exponential moving average rules had more success under low volatility. Confirming the results of Hsu and Kuan (2005), the

stocks on the Russell 2000 appear to have inefficient markets, allowing for exploitation by technical trading rules.

2.10 Bibliography

- Andersen T. G. and T. Bollerslev (1997), “Intraday Periodicity and Volatility Persistence in Financial Markets”, *Journal of Empirical Finance*, 4(2-3), 115-158.
- Andersen T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2003), “Modeling and Forecasting Realized Volatility”, *Econometrica*, 71(2), March, 579-625.
- Aronson, D. (2007), *Evidence Based Technical Analysis*, John Wiley & Sons, Inc.
- Brock, W., J. Lakonishok, and B. LeBaron (1992), “Simple Technical Trading Rules and the Stochastic Properties of Stock Returns”, *Journal of Finance* 47, 1731-1764.
- Chang, P. H. K., and C. L. Osler (1999), “Methodological Madness: Technical Analysis and the Irrationality of Exchange-Rate Forecasts”, *Economic Journal* 109, 636-661.
- Curcio, R., C. Goodhart, D. Guillaume, and R. Payne (1997), “Do Technical Trading Rules Generate Profits? Conclusions From the Intra-day Foreign Exchange Market”, *International Journal of Finance*

Economics 2, 263-331.

Dacorogna, M. M., R. Gençay, U. A. Müller, R. B. Olsen, and O. V.

Pictet (2001), “An Introduction to High Frequency Finance”, *Academic Press, San Diego, CA*.

Hansen, P. (2005), “A Test for Superior Predictive Ability”, *Journal of Business and Economic Statistics* 23, 365-380.

Hsu, P. H., Y. C. Hsu, and C. M. Kuan (2010), “Testing the Predictive Ability of Technical Analysis Using A New Stepwise Test without Data Snooping Bias”, *Journal of Empirical Finance*, 17, 471-484.

Hsu, P. H., and C. M. Kuan (2005), “Reexamining the Profitability of Technical Analysis with Data Snooping Checks”, *Journal of Financial Econometrics* 3, 606-628.

Knez, P. and M. Ready (1996), “Estimating the Profits From Trading Strategies”, *Review of Financial Studies* 9, 1121-1164.

Korajczyk, R., and R. Sadka (2004), “Are Momentum Profits Robust to Trading Costs?”, *Journal of Finance* 59, 1039-1082.

Lesmond, D., M. Schill, and C. Zhou (2003), “The Illusory Nature of Momentum Profits”, *Journal of Financial Economics* 71, 349-380.

Lo, A. W., H. Mamaysky, and J. Wang (2000), “Foundations of Technical Analysis: Computational Algorithms, Statistical Inference, and Empirical Implementation”, *Journal of Finance* 55, 1705-1765.

Neely, C. J., and P. Weller (1999), “Technical Trading Rules in the European Monetary System”, *Journal of International Money and Finance* 18, 429-458.

- Neely, C. J., and P. Weller (2002), “The Temporal Pattern of Trading Rule Returns and Central Bank Intervention: Intervention Does Not Generate Technical Trading Rule Profits”, *Journal of International Economics* 58, 211-232.
- Neely, C. J., and P. Weller (2003), “Intraday Technical Trading in the Foreign Exchange Market”, *Journal of International Money and Finance* 22, 223-237.
- Neely, C. J., P. Weller, and R. Dittmar (1997), “Is Technical Analysis in the Foreign Exchange Market Profitable? A Genetic Programming Approach”, *Journal of Finance and Quantitative Analysis* 32, 405-426.
- Ready, M. (2002), “Profits From Technical Trading Rules”, *Financial Management* 31, 43-61.
- Osler, C. (2000), “Support for Resistance: Technical Analysis and Intraday Exchange Rates”, *Economic Policy Review, Federal Reserve Bank of New York* issue July, 53-68.
- Politis, D. N., and J. P. Romano (1994), “The Stationary Bootstrap”, *Journal of the American Statistical Association* 89, 1303-1313.
- Pring, M. (2002), *Technical Analysis Explained, 4th Edition*, McGraw-Hill.
- Romano, J., and M. Wolf (2005), “Stepwise Multiple Testing as Formalized Data Snooping”, *Econometrica* 73, 1237-1282.
- Rouwenhorst, K. G. (1998), “International Momentum Strategies”, *Journal of Finance* 53, 267-284.
- Sullivan, R., A. Timmerman, and H. White (1999), “Data-Snooping,

Technical Trading Rule Performance, and the Bootstrap”, *Journal of Finance* 54, 1647-1692.

Tuttle, K. (2006), *Master Traders*, New York: John Wiley & Sons, Inc.

White, H. (2000), “A Reality Check for Data Snooping”, *Econometrica*, 68(5), 1097-1126.

Chapter 3

The VARLINEX Model Extended: Accurate Value-at-Risk Forecasts During the 2008-09 Market Crash

3.1 Introduction

The recent market crash has made the measurement and forecasting of risk a major issue for both financial institutions and regulators.¹ Under the Basel II accord, banks are required to maintain a specific level of capital which is related to the riskiness of their assets. Specifically, the risk capital of a bank must be sufficient to cover losses on the bank's trading portfolio over a 10-day holding period, 99% of the time (Basel Committee, (1995)). The 10-day loss that is only surpassed 1% of the time is referred to as the Value-at-Risk (VaR). VaR models have been widely accepted as the appropriate method to be used in the determination of required risk capital.²

¹If not otherwise specified, the market crash referred to throughout this paper is the stock market crash that began in September of 2008, initiated by the collapse of Lehman Brothers.

²In this paper, the focus is on market risk, as opposed to other types of risk commonly faced by large financial institutions, such as credit, operational, or liquidity risk.

Banks are allowed to use their own internal VaR models to determine the risk capital charge they must incur. If they choose not to use their own internal model, a default approach will be used, as defined in the Basel II accord. The default approach leads to an overly-conservative capital charge, causing the bank to hold additional capital that earns little or no return.³ In addition, regulators check the ex-post performance of a bank's VaR model, and banks using inaccurate models are penalized by being required to hold more capital.⁴ Thus, it is beneficial for banks to have their own accurate internal VaR model because this will give them the ability to invest the additional capital in a variety of ways instead of being restricted to keeping it in liquid, low-yielding assets. In addition, an accurate VaR model can help to identify exposures that carry an unacceptable amount of downside risk, potentially allowing the portfolio manager to alter the composition of the portfolio before it takes a turn for the worse.

From the perspective of an investment firm, VaR models can be used to keep traders in check by imposing VaR limits on each trader's portfolio. In addition, marginal VaR can be used to evaluate the risk of a new position that is being proposed for inclusion in a portfolio of assets.⁵ In short, VaR models are becoming increasingly important both for banks and for buy-side investment firms. An important question for these financial institutions is 'how reliable are our VaR models during a time of crisis?'. This paper develops a model that provides accurate VaR forecasts in general, but more importantly in a time of crisis.

³The terms 'capital charge' and 'risk capital charge' are used interchangeably.

⁴The details of this penalty can be found on pg. 12 of "Revisions to the Basel II market risk framework".

⁵See Jorion (2007) for a discussion of marginal VaR.

The $\alpha\%$ VaR of a portfolio is obtained using the $1 - \alpha$ quantile of the return distribution of the portfolio. Since the true distribution is unknown, it must be estimated. In particular, the tails of the conditional distribution of returns need to be estimated accurately to get reliable forecasts of VaR.⁶

Most of the existing VaR literature ignores the asymmetric loss of forecasting VaR for different economic agents. The VARLINEX model takes this asymmetry into account by using the Linex loss function, which is an alternative to the traditionally used symmetric loss function. Specifically, bank managers likely feel that the loss of overestimating VaR is greater than the loss of underestimation. This is because, in the event of overestimation, banks will hold more capital than is required by regulation, which will decrease their return on equity. On the other hand, regulators certainly would suffer a greater loss from underestimation relative to overestimation, since underestimation is more likely to result in systemic bank failures. Additionally, a bank manager may wish to err on the side of caution in a time of crisis, in which case his perspective will be similar to that of the regulator.

In this paper, an extension is provided to the VARLINEX model. Two alternative methods of dynamically estimating the asymmetry parameter from the VARLINEX model are introduced. In particular, the first procedure finds the optimal value of the asymmetry parameter a over some in-sample period, and then uses that a value to forecast future VaR.⁷ The estimate of a is updated every 50 days. The second

⁶In the model presented in this paper, the conditioning is on current volatility, skewness, and kurtosis.

⁷A definition of what is meant by ‘optimal’, as well as a detailed description of the estimation method will be provided in Section 3.4.

procedure allows the a parameter to depend on the VIX index. These new estimation methods improve upon the results of the original model (i.e. KSW (2003)),⁸ in addition to outperforming the benchmark Extreme Value Theory model. This paper provides a versatile VaR model capable of yielding accurate forecasts during a time of crisis as well as under calm market conditions.

The remainder of the paper is organized as follows. The model is described in Section 3.2. Section 3.3 reviews the literature. The data is described in Section 3.4. The estimation methods for the a parameter are explained in Section 3.5. In Section 3.6, the results are presented. Section 3.7 concludes.

3.2 VARLINEX Model

Before getting into the VARLINEX model, it may be useful to outline the general structure of the VaR problem.⁹

3.2.1 Initial Structure of VaR Problem

First, a model for the data generating process of the returns needs to be chosen. A common model is $r_t = \mu_t + \epsilon_t$, where ϵ_t is the innovation of the process. The innovation is allowed to be heteroskedastic, but serially uncorrelated with mean zero.

⁸However, the second estimation method only slightly improves upon the original model.

⁹The following subsection outlines one common approach, there are many other ways to set-up the problem.

Let $\epsilon_t = \sigma_t z_t$, where σ_t is a scaling factor, and z_t is an *iid* process, then $r_t = \mu_t + \sigma_t z_t$. Now, the $\alpha\%$ VaR is given by $P(r_t < -VaR_t^\alpha | F_{t-1}) = 1 - \alpha$, where the information set F_{t-1} is a filtration generated by the innovations $\epsilon_1, \dots, \epsilon_{t-1}$, and VaR_t^α denotes the Value-at-Risk for time t and confidence level α . It would be nice to have an expression for VaR_t^α , so let G represent the distribution of the innovation.¹⁰ Then, given a confidence level α and an information set F_{t-1} , the following expression for the Value-at-Risk emerges:

$$VaR_t^\alpha = -(\mu_t + \sigma_t G_{1-\alpha}^{-1}) \quad (3.1)$$

where $G_{1-\alpha}^{-1}$ is the $1 - \alpha$ -quantile of G .

From (3.1), it is clear that in order to model VaR it is necessary to model the mean, μ , the standard deviation, σ , and the quantiles of the distribution of the innovation. It turns out that modelling the mean is not of critical importance,¹¹ so the literature is focused on modelling the volatility process and the quantiles of the distribution of the innovation. This is a general formulation of the VaR problem, there are many variations, for instance, the VARLINEX model also takes into account the skewness and kurtosis.

¹⁰Assume G belongs to the location-scale family of distributions.

¹¹Often times it is modeled using an AR(1) process as in Mittnik and Paoletta (2000).

3.2.2 Value-at-Risk Linear Exponent (VARLINEX) Model

A summary of the derivation of the VARLINEX model is presented in this section (KSW (2003) provide a more comprehensive derivation). A description of the VARLINEX model begins with an introduction to the Linex loss function.

Linex Loss Function

Suppose there exists a random variable y . Let h be a forecast of y with a forecast error given by $u = y - h$. The prediction of y involves the choice of a loss function $L(u)$, and the forecast h is obtained by minimizing $E[L(u)]$. The Linex loss function is defined as

$$L(u) = e^{au} - au - 1,$$

where a is the so called asymmetry parameter referred to earlier. Let V be the value of the bank's trading portfolio, then VaR is defined to be the value V_0 that satisfies

$$F(V \leq V_0) = \alpha_0,$$

where $F(\cdot)$ is the cumulative distribution function of V , and α_0 is the confidence level chosen by the user. Note that V_0 will be a point forecast (h) of a quantile (y) and has a distribution in its own right. V_0 may be calculated parametrically or nonparametrically. In this model it is calculated nonparametrically, following the methods of Cain (1991) who computes optimal forecasts of p_{th} order statistics for an asymmetric loss function. However, note that the loss function used here is not the same as that used in Cain (1991). Zellner (1986) shows that the optimal h is given

by

$$h = \frac{1}{a} \log E[e^{ay}],$$

where $E[\cdot]$ is the expected value over the y distribution. Also, $E[e^{ay}]$ is the moment generating function of y , so the key to Linex forecasting is to find $E[e^{ay}]$.

At this point, a quick note on the a parameter is in order. The reader should note that if a is positive, the function penalizes overestimation more than underestimation, thus there is a greater penalty if the amount of actual capital required is less than the forecast amount. This is the perspective of the bank manager during a bull market. Conversely, if a is negative, underestimation is penalized more than overestimation, thus the penalty is larger if the amount of actual capital required is greater than the forecast amount. This is the perspective of the regulator, or the bank manager in the midst of a crisis.

A key issue in the VARLINEX framework is how to estimate the a parameter. In KSW (2003), a is assumed to be a parameter in an asymmetric utility function. a is estimated using options data and historical prices of the underlying. However, once a has been estimated, the same value is used for the entire forecasting sample. This is certainly a sub-optimal estimation strategy because the desired ratio of forecasted required capital to actual required capital is certain to fluctuate in line with volatility in the underlying. Specifically, the aforementioned ratio would likely increase as volatility increases, and in this scenario, a lower a value would be appropriate.¹² In Section 3.4, two improved dynamic estimation methods are introduced, which is the central contribution of this paper.

¹²When volatility increases, the bank manager would like to be more safe than usual, and as a result, the ratio of forecasted required capital to actual required capital should be higher. This is consistent with a lower a value.

Quantile Distribution

VaR is essentially an estimate of the $1 - \alpha$ quantile of the return distribution of some asset or portfolio of assets, and, as mentioned earlier, it can be described by some distribution, therefore the distributional properties warrant further examination. Assume the existence of a random sample of n observed financial returns r_1, r_2, \dots, r_n from a distribution function $F(r)$. Their joint distribution is given by

$$dG = dF(r_1)dF(r_2) \cdots dF(r_n).$$

A transformation to

$$y_k = \int_{-\infty}^{r_k} dF(t) = F(r_k)$$

gives, for the joint distribution of the y 's

$$dH = dy_1 dy_2 \cdots dy_n \quad (0 \leq y_k \leq 1, \forall k).$$

Since y is a non-decreasing function of r , the y 's and the r 's are in the same order.

Now, re-number the y 's so that y_1 is the smallest, y_2 the second smallest and so on, with

$$0 \leq y_1 \leq y_2 \leq \cdots \leq y_n \leq 1,$$

then the distribution of any ordered y , i.e. y_k , can be derived and is given in Stuart and Ord (1994) as

$$dM = \frac{n!}{(k-1)!(n-k)!} (y_k)^{k-1} (1-y_k)^{n-k} dy_k,$$

which is a beta distribution of the first kind. To obtain the distribution of the k_{th} order statistics $r_{k:n}$, set $F(r_{k:n}) = y_k$, yielding

$$dN = \frac{n!}{(k-1)!(n-k)!} (F(r_{k:n}))^{k-1} (1 - F(r_{k:n}))^{n-k} f(r_{k:n}) dr_{k:n}.$$

Next, the moment generating function (*mgf*) of y_k is derived. The following theorem is Theorem 1 in KSW (2003).

Theorem 3.1. *The mgf of y_k is given by*

$$\begin{aligned} E[e^{ay_k}] &= {}_1F_1(k, k + n - k + 1, a) \\ &= {}_1F_1(k, n + 1, a) \end{aligned}$$

where ${}_1F_1(\cdot)$ is the confluent hypergeometric function.¹³

In general, there is no closed analytical formula for the hypergeometric function, so its value is obtained numerically. The Linex forecast of the k_{th} -ordered percentage y_k is given by

$$\hat{h} = \frac{1}{a} \log[{}_1F_1(k, n + 1, a)].$$

The Linex forecast depends only on the values of k , $n + 1$, and a , not on other parameters which specify the distribution F . The reason is that the original data r_i has been transformed into $F(r_i)$, which follows the standard uniform distribution for any continuous distribution function, according to the Probability Integral Transform Theorem. The Linex forecast is nonparametric in this sense.

Given a $F(r)$, the distribution can be inverted to get the corresponding quantile or VaR

$$\hat{V}_0 = F^{-1}(\hat{h}).$$

¹³See KSW (2003) for the proof.

Since the true distribution of returns (r) is usually unknown, the Cornish-Fisher expansion (Stuart and Ord (1994)) is applied to find the corresponding VaR

$$\hat{V}_0 = k_1 + k_2(z_{\hat{h}} + \frac{1}{6}(z_{\hat{h}}^2 - 1)k_3 + \frac{1}{24}(z_{\hat{h}}^3 - 3z_{\hat{h}})(k_4 - 3) - \frac{1}{36}(2z_{\hat{h}}^3 - 5z_{\hat{h}})k_3^2)$$

where k_1, k_2, k_3, k_4 are the first four cumulants of the distribution $F(r)$, and $z_{\hat{h}}$ is the $\hat{h}\%$ quantile for the standard normal distribution. Since the Normal distribution is used here, the Cornish-Fisher expansion is useful for approximating the quantiles of non-Normal distributions as long as the skewness and kurtosis fall within appropriate ranges, such that the departure from normality is sufficiently small.

3.3 Literature Review

There are two main areas of focus when modelling VaR; namely, *i*) estimation of the quantiles of the return distribution and *ii*) the volatility modelling process. The literature review will be broken down into two sections. Section 3.3.1 will focus on the different quantile estimation methods and Section 3.3.2 examines different volatility modelling processes.

3.3.1 Quantile Estimation

There are three broad categories of quantile estimation techniques: fully parametric, semiparametric, and nonparametric. Fully parametric methods are based on an econometric model for the volatility dynamics (for example, GARCH) and an assumption that the innovation follows a standard probability distribution such as the Normal or the Student-t. This is the approach in the most famous VaR system, Riskmetrics (Riskmetrics (1995)), developed by J.P. Morgan. However, the Riskmetrics approach has fallen out of favour in the VaR literature due to its reliance on the Normal distribution. Financial data are notorious for exhibiting fat tails, and as a result, the Riskmetrics system often underestimates the true VaR.

The semiparametric group features methods based on Extreme Value Theory (EVT) and quantile regression. One EVT approach is to use historical data to model the middle of the return distribution, but use EVT to model the tails (see McNeil and Frey (2000) and Kuester, Mittnik, and Paoletta (2006)). In terms of quantile regression methods, see the class of CAViaR models introduced in Engle and Manganelli (2004).

The most common nonparametric method is historical simulation. In the historical simulation approach, the estimated return distribution of a portfolio is given by the empirical distribution of past returns on the portfolio. This approach is used in numerous papers including Barone-Adesi, Bourgoin, and Giannopoulos (1998).

3.3.2 Volatility Modelling

The bulk of the volatility models used in the VaR literature can be grouped in three general categories: GARCH and variants, models incorporating realized volatility, and stochastic volatility models. By far, the most popular approach to modelling the volatility process in the VaR literature is to use one of the GARCH models. For instance, Giot and Laurent (2004) use a daily-ARCH model to compute one-day-ahead VaR, Kuuster et al. (2006) use the standard GARCH(1,1) as well as a mixed-normal GARCH, and Daal and Yu (2005) compare the performance of various GJR-GARCH models with a mixed GARCH-jump model.

There have been a few attempts to model the volatility process using models of realized volatility in the VaR literature, with varying degrees of success. Giot and Laurent (2004) find that a long memory skewed Student model for the daily realized volatility provides adequate VaR forecasts, but “it does not really improve on the performance of a VaR model based on the skewed Student APARCH model and estimated using daily data only”. Kruse (2006) finds that hybrid models based on realized volatility and filtered historical simulation provide excellent one-day-ahead VaR forecasts.

Stochastic volatility models have not been used very often in a VaR setting. Kruse (2006) shows that a stochastic volatility model combined with a filtered EVT specification performs well; however, it is outperformed by many alternate models, including GARCH and realized volatility models.

3.4 Data

The data consists of daily returns from the Dow Jones Industrial Average (DJI), the NASDAQ Composite Index, and the S&P/TSX Composite Index, from January 4, 2000 to October 22, 2009.¹⁴ The data is divided into an estimation period ranging from January 4, 2000 to December 26, 2003, a ‘pre-crash’ period ranging from December 29, 2003 to October 31, 2007, a ‘during-crash’ period ranging from November 1, 2007 to March 31, 2009, and a ‘post-crash’ period ranging from April 1, 2009 to October 22, 2009.^{15,16} Table 3.1 shows the summary statistics for the full sample and each subsample for the Dow Jones data set.¹⁷ Notice that the data is non-normal in all cases, except for the post-crash subsample.

Table 3.1: DJI - Summary Statistics

	Full Sample	Pre-Crash	Crash	Post-Crash
Mean	0.000069514	0.00033359	-0.0014	0.002
SD	0.0131	0.0069	0.0227	0.0121
Skewness	0.317	-0.2809	0.4445	-0.1555
Kurtosis	15.2096	4.3363	6.843	3.4654
Min	-0.0787	-0.0329	-0.0787	-0.0356
Max	0.1108	0.0251	0.1108	0.0314
Jarque-Bera (p-value)	< 0.001	< 0.001	< 0.001	0.3264

¹⁴Calculated from daily closing prices.

¹⁵The estimation period for the S&P/TSX data ends on April 13, 2004 and the pre-crash period starts on April 14, 2004, otherwise everything else is the same as the other data sets.

¹⁶The transition points marking the beginning and end of the ‘during-crash’ period are meant to roughly correspond to the peak and trough, respectively, of the U.S. indices.

¹⁷Summary statistics for the other two data sets are in the appendix (Table C.57 for NASDAQ and Table C.78 for S&P/TSX).

3.5 Extended VARLINEX Model

This section describes in detail the extension to the original VARLINEX model. The extension was to develop improved methods of estimating the a parameter. In addition, the model is adjusted so that the VARLINEX method is applied to an *iid* residual series rather than the raw data. It makes more sense to discuss the latter adjustment first.

In order to obtain quantile estimates of the return distribution using the VARLINEX model, it is assumed that financial returns are *iid*; however, financial data is rarely *iid*. A solution to this problem is to fit an AR-GARCH model to the raw return series, and then estimate the quantiles of the resulting standardized residual series as is done in McNeil and Frey (2000).¹⁸ Providing the residual series is *iid*, this procedure ensures the VARLINEX method is being utilized appropriately. A logical first step is to start with the simplest specification, which consists of one lag in the conditional mean equation and a GARCH(1,1) with Normal innovations in the conditional variance equation. The AR(1)-GARCH(1,1) model specification used is the following:

$$\begin{aligned}x_t &= \phi x_{t-1} + \epsilon_t \\ \sigma_t^2 &= \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta \sigma_{t-1}^2,\end{aligned}$$

where $X = (x_1, \dots, x_n)$ is the return data, $\epsilon_t = \sigma_t z_t$, $z_t \sim N(0, 1)$, and α_0 , α_1 , β , and ϕ are the parameters to be estimated.

Figure 3.1 shows the autocorrelation function for the raw data series (panels a)

¹⁸Typically, the residual series will be *iid*.

and b)) and the residual series (panels c) and d)) for the DJI data. The plots show that the residual and the absolute value of the residual are at least approximately *iid*, while the absolute value of the raw series is clearly not *iid*.¹⁹ The Ljung-Box test confirms that the residual and absolute residual are in fact *iid* for all three data sets. As a result, there is no need to experiment with different specifications of the AR-GARCH model.

The most obvious weakness of the original VARLINEX method was that the

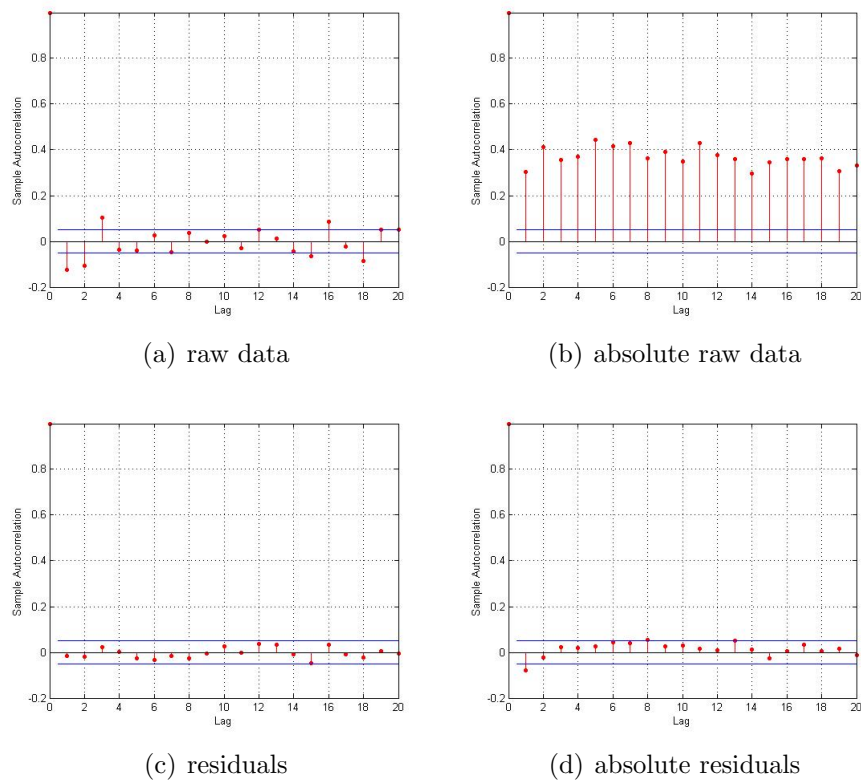


Figure 3.1: Correlograms for the raw data and their absolute values as well as for the residuals and absolute residuals. The horizontal lines are the 95% confidence limits.

¹⁹In all plots and tests concerning the autocorrelation function, up to 20 lagged values were used.

a parameter was estimated only once for the entire sample. In this paper, both of the new estimation procedures estimate a dynamically.²⁰ In addition, the manner in which a is estimated is different from the original model. Now, both of the new estimation procedures will be described.

3.5.1 Estimation method 1

First, a step-by-step look at how a is estimated for the first 50 day window of the out-of-sample period is provided. Suppose the out-of-sample forecasting period begins on day 1051.²¹ Before the VaR for day 1051 can be forecasted, a forecast of a is required. In order to obtain a forecast of a , it is necessary to find the optimal a value from the most recent 50 day window, i.e., from day 1001 to day 1050. What is the ‘optimal a value’? It is the value for a that yields the expected number of VaR violations over the 50 day estimation window.²² For example, for the 95% VaR the expected number of violations over a 50 day period would be $(1 - 0.95) \times 50 = 2.5$. Then, the optimal a value would be the value that yielded 2 or 3 VaR violations (since obtaining a non-integer number of violations is impossible in practice). Given the example just presented, the following is the algorithm used to obtain the optimal a .

i) Days 1 to 1000 are used to estimate the four cumulants of the return distribution

²⁰In the first method, a is re-estimated every 50 days, while in the second method it is re-estimated daily.

²¹Daily data is used exclusively throughout this paper.

²²A violation is said to occur when the return is lower than the VaR forecast.

that are used in the Cornish-Fisher expansion, then starting on day 1001 a VaR forecast is made for some starting value of a .²³

ii) A VaR forecast is made for each day in the initial 50 day window (days 1001 to 1050) for the same a value.

iii) The number of VaR violations is counted.

a) In the example above, if the number of violations is less than 2, increase a by 100 and start again at step *i*).

b) If the number of violations is equal to 2, then store this a value as a_{low} , and start again at step *i*), but this time skip steps *iii*)*a*) and *iii*)*b*).

c) If the number of violations is greater than or equal to 3, then this a value is stored as a_{high} ; proceed to step *iv*). Otherwise, increase a by 100 and start again at step *i*), once again skipping steps *iii*)*a*) and *iii*)*b*).

iv)*a)* If values exist for both a_{low} and a_{high} , then a is calculated as $a = \frac{a_{low} + a_{high}}{2}$.

b) If there is only a value for a_{high} , then $a = a_{high}$.²⁴

The a value from step *iv*) is used in the VaR forecast for days 1051 to 1100. Then the entire process is repeated for each subsequent 50 day window until the end

²³In the paper, a starts with a value of -2000 and then is increased in increments of 100. The VaR forecasts are recalculated each time a is changed until the procedure is complete. Computational issues arise when the value for a becomes sufficiently small (i.e. a negative number that is sufficiently large in absolute value). As a result, in some cases, the VaR forecasts are more aggressive (i.e. smaller in absolute value) than what is optimal.

²⁴This occurs if the lowest a value that is tested (i.e. $a = -2000$) yields a number of violations greater than the expected number.

of the out-of-sample data period.²⁵ The result is an accurate and robust model that produces VaR forecasts that are sensitive to changing market conditions. The shape of the agent's loss function is allowed to change dynamically according to the changing conditions in the market, which is an improvement over the original method, where the shape of the loss function remained constant throughout the forecasting period.

3.5.2 Estimation method 2

The VIX index is a measure of the implied volatility of S&P 500 index options. It is a measure of expected stock market volatility over the next 30 days. As mentioned earlier, the a parameter should be negative when the market is volatile, and positive when the market is calm. As a result, it would be ideal to have large VIX index values correspond to negative a values. A method of mapping VIX index values into a values is needed. First, ranges are defined for both objects. The range for the VIX index is chosen as

$$[VIX_{min}, VIX_{max}],$$

where VIX_{min} is the smallest value in the history of the VIX index up until the beginning of the forecasting period, and likewise, VIX_{max} is the largest value in the history of the VIX index up until the beginning of the forecasting period.²⁶ The range used for a is: $[-2000, 1200]$. The left endpoint is a computational constraint; it is

²⁵The size of the estimation sample remains constant at 1000 days, i.e., a moving-window approach is used.

²⁶The VIX index started January 2, 1990.

very time consuming to compute the confluent hypergeometric function for very small values of a (specifically, less than -2000).²⁷ In the first estimation method presented above, the highest optimal a value is 1200, so this is used as the right endpoint. A simple mapping method would be the following:

1. Let $a = -2000$ when $VIX_t \geq VIX_{max}$
2. Let $a = 1200$ when $VIX_t \leq VIX_{min}$
3. For $VIX_{min} < VIX_t < VIX_{max}$, let $proxy_t = \frac{VIX_{max} - VIX_t}{VIX_{max} - VIX_{min}}$, then

$$a_t = VIX_{min} + (VIX_{max} - VIX_{min}) \cdot proxy_t.$$

This is just a simple linear mapping. One problem with this method is that as VIX_t increases, the VaR forecasts decrease exponentially. Ideally, the VaR forecasts should have a linear relationship with movements in the VIX index. Another problem is that the VaR forecasts change quite a bit from day-to-day, which increases the difficulty of meeting the capital requirements. Fortunately, there are simple fixes to both of these problems. To solve the first problem, take the log of the VIX index, and map these values into a values. The result is a linear relationship between VaR forecasts and VIX values.²⁸ The second problem is solved by taking an n -day moving average of the VIX index and using this value instead of the current daily VIX value.²⁹ This has the effect of smoothing out the day-to-day fluctuation in the VaR forecasts.

²⁷At least this appears to be the case in Matlab; alternative methods of calculating the confluent hypergeometric function were not tested in this study.

²⁸Of course, the relationship is still negative, as before.

²⁹10-day and 50-day moving averages are used.

3.6 Results

The results for the extended VARLINEX model presented in this section are compared to two benchmark models: an EVT model that follows the methodology in Gilli and Kellezi (2006), and the original VARLINEX model (KSW (2003)). The EVT model is estimated in Matlab.³⁰ The original VARLINEX model provides VaR forecasts using Mathematica. Matlab is used for the extended VARLINEX model.

As mentioned earlier, the data consists of daily returns; however, it is not the raw returns that are fed into the VARLINEX forecasting procedure, it is the residual series obtained from an AR(1)-GARCH(1,1) estimation of the raw data series. VaR forecasts are obtained for the residual series, and then these forecasts need to be mapped back to VaR forecasts for the raw data series and compared to the raw returns to determine the number of violations. The residuals are given by

$$z_t = \frac{x_t - \hat{\mu}_t}{\hat{\sigma}_t},$$

where x_t is the raw return at time t , and $\hat{\mu}_t$ and $\hat{\sigma}_t$ are given by

$$\begin{aligned}\hat{\mu}_t &= \hat{\phi}x_{t-1} \\ \hat{\sigma}_t^2 &= \hat{\alpha}_0 + \hat{\alpha}_1(x_{t-1} - \hat{\mu}_{t-1})^2 + \hat{\beta}\hat{\sigma}_{t-1}^2.\end{aligned}$$

Values for $\hat{\phi}$, $\hat{\alpha}_0$, $\hat{\alpha}_1$, and $\hat{\beta}$ are obtained via estimation. The VARLINEX procedure generates VaR forecasts on the series of residuals. These forecasts are then mapped

³⁰The version of Matlab used in this paper is Matlab 7.8.0.

back to give VaR forecasts for the raw series

$$VaR_{raw} = \hat{\mu} + \hat{\sigma}VaR_{residual}.$$

Finally, $VaR_{raw} = (VaR_{raw}^1, \dots, VaR_{raw}^T)$ is compared with the raw series of returns $X = (x^1, \dots, x^T)$; if $x^t < VaR_{raw}^t$, a violation has occurred.³¹

The issue of how to evaluate the accuracy of a VaR model is now addressed. Numerous studies have already addressed this issue, and standard methods are available (see, for example, Kupiec (1995), Christoffersen (1998), and Kuester et al. (2006)). Define $V_t = \mathbf{1}(x^t < VaR^t)$. V_t takes the value 1 when a violation occurs, and 0 otherwise. Christoffersen (1998) introduces the concept of an *efficient* sequence of forecasts. The definition is repeated here, although in a slightly different manner.

Definition 3.1 (Christoffersen (1998)) *The sequence of forecasts $\{VaR^t\}_{t=1}^T$ is said to be efficient with respect to information set F_{t-1} , if $E[V_t|F_{t-1}] = p$, for all t , where p is the theoretical proportion of violations.*

Defining the information set as the sequence of past realizations of V_t , that is, $F_{t-1} = \{V_{t-1}, \dots, V_1\}$, yields the following result.

Lemma 3.1 (Christoffersen (1998)) *Testing $E[V_t|F_{t-1}] = E[V_t|V_{t-1}, \dots, V_1] = p$, for all t , is equivalent to testing that the sequence $\{V_t\}_{t=1}^T$ is iid Bernoulli with parameter p . (See Christoffersen (1998) for proof).*

³¹Note that VaR_{raw}^t is always a negative number in all three data sets.

Several implications of Definition 3.1 are tested to obtain a clear picture of where the model has strength and weakness.

Test of Unconditional Coverage

One implication from Definition 3.1 is that $E[V_t] = p$. This just says that the number of violations generated by the model is equal to the theoretical expected number of violations. $E[V_t] = p$ can be tested as follows:

$$H_0 : E[V_t] = p \quad \text{vs.} \quad H_1 : E[V_t] \neq p.$$

Equivalently, H_0 can be expressed in terms of the likelihood function $L(p; V_1, \dots, V_T) = (1-p)^{n_0}(p)^{n_1}$ and H_1 in terms of $L(\pi; V_1, \dots, V_T) = (1-\pi)^{n_0}(\pi)^{n_1}$. Then, unconditional coverage can be tested for using a likelihood-ratio test

$$LR_{uc} = 2\log L(\hat{\pi}; V_1, \dots, V_T) - 2\log L(p; V_1, \dots, V_T),$$

which is asymptotically $\chi^2(1)$ distributed. $\hat{\pi} = \frac{n_1}{n_0+n_1}$ is the maximum likelihood estimate of π , where n_1 is the number of violations and can be expressed as $n_1 = \sum_{t=1}^T V_t$, n_0 is the number of non-violations, and $n_0 + n_1 = T$ is the total number of days in the forecast sample. Results of the test of unconditional coverage will be presented later.

Test of Independence

Lemma 3.1 stated that the violations should be independently distributed throughout the forecast sample. For example, for the 99% VaR there should be 10 violations

in a 1000 day sample. If one violation occurs every 100 days throughout the sample, the model has performed well. On the other hand, if there are zero violations for the first 900 days, and then all 10 are clustered in the last 100 days, the model is inadequate even though it yielded the expected number of violations. Following Christoffersen (1998) and Kuester et al. (2006), the null hypothesis of independence is tested against an explicit first-order Markov alternative. Specifically, V_t is a binary first-order Markov chain with transition probability matrix

$$\Pi = \begin{bmatrix} 1 - \pi_{01} & \pi_{01} \\ 1 - \pi_{11} & \pi_{11} \end{bmatrix},$$

where $\pi_{ij} = P(V_t = j | V_{t-1} = i)$. The null hypothesis states that the occurrence of a violation at time t does not affect the probability of a violation at time $t - 1$. The approximate joint likelihood, conditional on the V value at time 1, is

$$L(\Pi; V_2, \dots, V_T | V_1) = (1 - \pi_{01})^{n_{00}} \pi_{01}^{n_{01}} (1 - \pi_{11})^{n_{10}} \pi_{11}^{n_{11}},$$

where n_{ij} is the number of transitions from state i to state j

$$n_{ij} = \sum_{t=2}^T \mathbf{1}(V_t = i | V_{t-1} = j),$$

and the maximum likelihood estimators under the alternative hypothesis are

$$\hat{\pi}_{01} = \frac{n_{01}}{n_{00} + n_{01}} \quad \text{and} \quad \hat{\pi}_{11} = \frac{n_{11}}{n_{10} + n_{11}}.$$

Under the null, $\pi_{01} = \pi_{11} \equiv \pi_0$. Then, it follows that the conditional binomial joint likelihood is given by

$$L(\pi_0; V_2, \dots, V_T | V_1) = (1 - \pi_0)^{n_{00} + n_{10}} \pi_0^{n_{01} + n_{11}}.$$

Finally, the likelihood ratio test is given by

$$LR_{ind} = 2\log L(\hat{\Pi}; V_2, \dots, V_T|V_1) - 2\log L(\hat{\pi}_0; V_2, \dots, V_T|V_1),$$

where $\hat{\pi}_0 = \frac{n_{01}+n_{11}}{n_{00}+n_{01}+n_{10}+n_{11}}$ is the maximum likelihood estimate of π_0 .³² LR_{ind} is asymptotically χ^2 distributed with one degree of freedom. In all three data sets, violations are determined to be independently distributed. Results of the independence test are available upon request.

Test of Conditional Coverage

A test for conditional coverage, i.e., $E[V_t|F_{t-1}] = p$, is essentially a joint test of unconditional coverage and independence. In effect, the null of the unconditional coverage test is tested against the alternative of the independence test. The test statistic used in Christoffersen (1998) is

$$LR_{cc} = 2\log L(\hat{\Pi}; V_2, \dots, V_T|V_1) - 2\log L(p; V_2, \dots, V_T|V_1),$$

which is asymptotically χ^2 distributed with two degrees of freedom. Christoffersen (1998) also shows that $LR_{cc} = LR_{uc} + LR_{ind}$. The test of conditional coverage is satisfied in virtually every sub-sample for each data set. In the following section, the new VARLINEX model is compared to the benchmark models in a test of unconditional coverage.³³

³²Note that $\hat{\pi}_0$ is identical to $\hat{\pi}$ in the test of unconditional coverage.

³³The test for unconditional coverage lends itself to comparison with alternative models more naturally than the other tests presented here.

3.6.1 Results for Estimation Method 1

In general, for all three data sets, there is consistent improvement over the original VARLINEX for the pre-crash, during-crash, and full samples, but roughly equivalent results after the crash. Relative to the EVT method, there is significant improvement during the crash and overall, and slight improvement pre- and post-crash.

Table 3.2 displays the number of VaR violations for the new VARLINEX model and the two benchmark models for the full sample. The new VARLINEX is the best model for the 99 and 95% quantiles, while the original VARLINEX is slightly better at the 97.5% quantile. Table 3.3 displays the results for the pre-crisis data. The new VARLINEX model is the best at all three quantiles, save for a tie with the original VARLINEX at the 97.5% quantile. Table 3.4 shows the results during the crisis. The EVT method, which performed fairly well pre-crisis, is apparently unable to adjust to the huge jump in volatility. The new VARLINEX method is the best at all three quantiles. The ‘after-crisis’ results are presented in Table 3.5. All three models perform well, except for the original VARLINEX method at the 95% quantile.

The results for the NASDAQ and S&P/TSX Composite data sets are fairly similar to those presented above for the DJI data. The results can be found in Appendix C.

Figure 3.2 shows the raw returns and the 99% VaR forecasts along with the 95% confidence interval of the forecasts for the DJI data series.

Similar to the results in KSW (2003), confidence intervals are wide under volatile conditions and narrow under calm market conditions. It has been shown that the VaR forecasts were very accurate, even during the market crash; however, if the portfolio manager was worried about suffering substantial losses and not having enough capital

available, he could obtain a more conservative VaR forecast by using the lower confidence interval of the forecast. The more risk averse a portfolio manager is, the more seriously he should consider using the lower confidence interval as the VaR forecast.

3.6.2 Results for Estimation Method 2

In general, for all three data sets, there is consistent improvement over the original VARLINEX during the crash and for the full samples, but inferior results pre-crash and similar results post-crash. Relative to the EVT method, there is significant improvement during the crash and overall, inferior performance pre-crash, and similar results post-crash.

Table 3.6 displays the number of VaR violations for the new VARLINEX model and the two benchmark models for the full sample. The new VARLINEX is the best model for all quantiles. The estimation method used to obtain the results presented here uses a 50-day moving average of the VIX index; the results using a 10-day moving average are also compiled. The results are similar to the 50-day case and are presented in Tables C.21-C.24.³⁴ The longer moving average is recommended since it does a better job of smoothing out the day-to-day fluctuation in capital requirements. Table 3.7 displays the results for the pre-crisis data. The new VARLINEX model is the best at the 95% quantile, the old VARLINEX model is the best at the 97.5% quantile, and the EVT model wins at the 99% quantile. Table 3.8 shows the results

³⁴Results for both the 10- and 50-day cases can be found in Tables C.49-C.56 (NASDAQ) and C.70-C.77 (S&P/TSX).

Table 3.2: DJI - Full sample - number of VaR violations (N=1466)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	135	92	88*	73.3
97.5%	87	43*	45*	36.65
99%	53	25	11*	14.66

‘*’ indicates the model passed the likelihood ratio (LR) test (i.e. H_0 was not rejected) at a 5% significance level.

Table 3.3: DJI - Pre-crisis - number of VaR violations (N=968)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	31	65	59*	48.4
97.5%	16*	29*	29*	24.2
99%	12*	18	8*	9.68

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table 3.4: DJI - Crisis - number of VaR violations (N=355)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	93	26*	23*	17.75
97.5%	70	13*	12*	8.875
99%	41	7*	3*	3.55

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table 3.5: DJI - Post-crisis - number of VaR violations (N=143)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	11*	2	6*	7.15
97.5%	1*	1*	4*	3.575
99%	0*	0*	0*	1.43

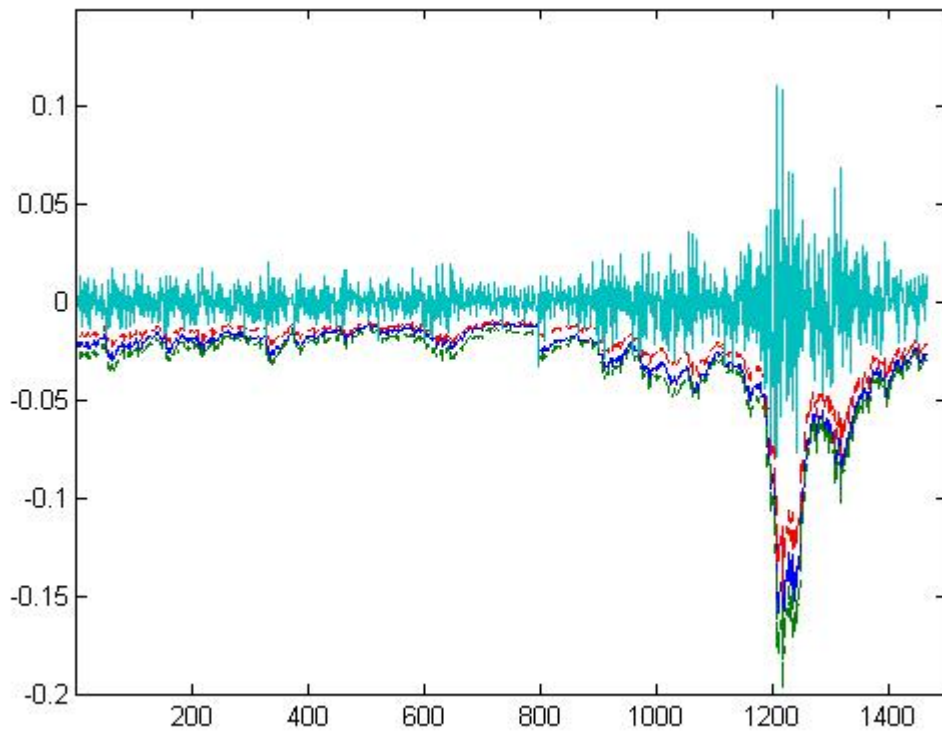
‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

during the crisis. The new VARLINEX method is the best at all three quantiles. Table 3.9 shows the results after the crisis. The new VARLINEX method performs slightly better than the other models overall.

The results for the NASDAQ and S&P/TSX Composite data sets are fairly similar to those presented above for the DJI data.³⁵

Perhaps the most impressive feature of the extended VARLINEX model is its

Figure 3.2: Dow Jones - 95% confidence interval of 99% VaR forecasts



³⁵The results for NASDAQ and S&P/TSX can be found in Appendix C.

ability to predict the crash before the collapse of Lehman Brothers.³⁶ There were several days of significant losses (and several 99% VaR violations) prior to the collapse of Lehman Brothers,³⁷ and as a result the model responded by yielding significantly more conservative VaR forecasts. By the time Lehman collapsed on September 15, 2008, causing a 4.4% drop in the DJI, the VaR forecasts were conservative enough to avoid a violation. In fact, there were no violations at the 99% level during the bulk of the crash (September 15, 2008 to March, 2009).³⁸ This predictive ability is what makes the extended VARLINEX model so useful. It is one thing for a model to be able to yield the correct number of independently distributed VaR violations, but this is not very useful if the model is unable to predict, or at least respond quickly to, significant market downturns.

3.6.3 Which Estimation Method Performs Better?

Prior to the crisis, method 2 gives more aggressive VaR forecasts than method 1, failing the LR test in several cases.³⁹ However, when volatility increases and the crisis occurs, method 2 gives more conservative forecasts, slightly outperforming method 1. After the crisis, method 2 still yields more conservative forecasts, but now it is

³⁶The discussion in this paragraph deals with the DJI data.

³⁷August 25, September 4, and September 9, 2008 all saw losses in excess of 2%, which was substantial for the DJI during the pre-crisis period.

³⁸This is true for the 99% VaR forecasts; however, there are some violations during the bulk of the crash for the 95 and 97.5% VaR.

³⁹This is especially true for the NASDAQ and S&P/TSX data sets.

Table 3.6: DJI - Full sample - number of VaR violations (N=1466)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	135	92	85*	73.3
97.5%	87	43*	39*	36.65
99%	53	25	17*	14.66

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table 3.7: DJI - Pre-crisis - number of VaR violations (N=968)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	31	65	63	48.4
97.5%	16*	29*	31*	24.2
99%	12*	18	14*	9.68

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table 3.8: DJI - Crisis - number of VaR violations (N=355)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	93	26*	18*	17.75
97.5%	70	13*	7*	8.875
99%	41	7*	3*	3.55

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table 3.9: DJI - Post-crisis - number of VaR violations (N=143)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	11*	2	4*	7.15
97.5%	1*	1*	1*	3.575
99%	0*	0*	0*	1.43

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

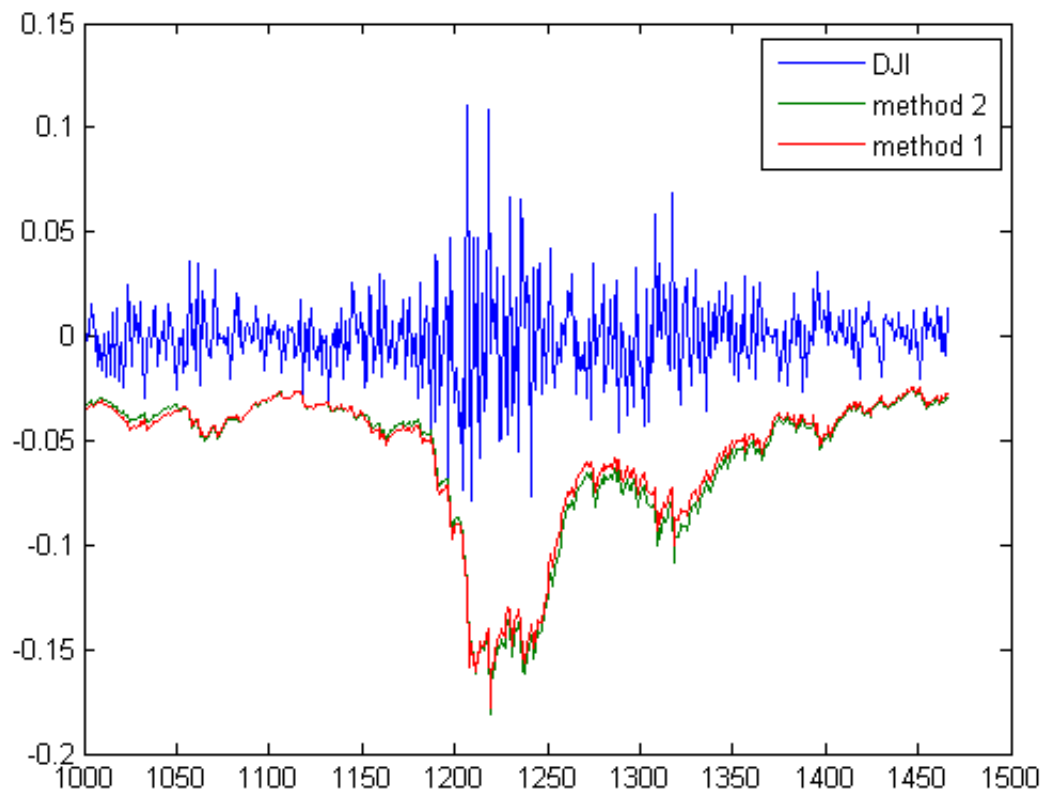
underperforming relative to method 1. Although method 2 performs very well during the crisis, its performance prior to the crisis is inadequate.⁴⁰ Thus, method 1 appears to be the preferred method. Figure 3.3 shows the 99% VaR forecasts for each method, along with the DJI daily returns. Only the final 466 days of the sample are shown to highlight the fact that the VaR forecasts for method 2 are more aggressive before the crisis, but more conservative during the crisis.

3.6.4 Extension - Varying window lengths

In the results presented for method 1, the a parameter was re-optimized every 50 days. This section examines how the results change when different window lengths are used. Window lengths of 10, 25, 75, 100, and 200 are used. The forecasts are expected to be more accurate for shorter windows, since more up-to-date information is being used; however, it will take longer to perform the computation. This is the classic accuracy versus efficiency issue that is often encountered in computational problems. The results for each window length are presented in Tables C.1 through C.20 in the appendix. In general, the forecasts are more accurate for shorter windows; however, there is not much difference in the results for window lengths ranging from 10 to 75 days. The quality of the forecasts begins to deteriorate for window lengths greater than 75.

⁴⁰The pre-crisis results aren't bad under the DJI data, but they are poor under the NASDAQ and S&P/TSX data.

Figure 3.3: Method 1 Vs. Method 2 - 99% VaR forecasts (DJI data)



3.7 Conclusion

In this paper, the VARLINEX model introduced in KSW (2003) is extended resulting in a Value-at-Risk forecasting model capable of producing accurate and robust forecasts before, during, and after the market crash of 2008-09. The extended VARLINEX is more accurate than both the original VARLINEX, and a benchmark model based on Extreme Value Theory. In addition, the results hold for data over three different indices: DJI, NASDAQ, and S&P/TSX Composite.

The original VARLINEX model is extended to accommodate two new estimation procedures for the asymmetry parameter, a , from the Linex loss function. In addition, under both methods, a is estimated dynamically throughout the sample as opposed to only once at the start of the sample as in the original VARLINEX model. The result is a more flexible model that is able to respond quickly to new information.

It is possible to obtain confidence intervals for the VaR estimates in addition to the conventional point estimates. This gives the forecaster a better idea of how reliable the point estimates are. For example, a wide confidence interval indicates that the point estimate may not be very reliable. In such a case, the forecaster may wish to add a risk premia adjustment to make the VaR point forecast more conservative, especially in a down-trending market.

An important feature of the model is its ability to predict the market crash. Portfolio managers utilizing the extended VARLINEX model would have been well-positioned to deal with the significant market risk that ensued. As a result, the

extended VARLINEX model should be a useful addition to any portfolio manager's toolkit.

3.8 Bibliography

- Artzner. P., F. Delbaen, J.-M. Eber, and D. Heath (1999), "Coherent Measures of Risk", *Mathematical Finance*, 9(3), 203-228.
- Bao, Y., T.-H. Lee, and B. Saltoglu (2006), "Evaluating Predictive Performance of Value-at-Risk Models in Emerging Markets: a Reality Check", *Journal of Forecasting*, 25(2), 101-128.
- Barone-Adesi, G., F. Bourgoin, and K. Giannopoulos (1998), "Don't Look Back", *Risk*, 11(8).
- Basel Committee on Banking Supervision (1996), "Overview of the Amendment to the Capital Accord to Incorporate Market Risks", <http://www.bis.org>.
- Basel Committee on Banking Supervision (2009), "Revisions to the Basel II Market Risk Framework", <http://www.bis.org>.
- Berkowitz, J. and J. O'Brien (2002), "How Accurate Are Value-at-Risk Models at Commercial Banks?" *Journal of Finance*, LVII, 1093-1111.
- Cain, M. (1991), "Minimal Quantile Measures of Predictive Cost with a Generalized Cost of Error Function", *Economic Research Papers*, no 40 University of Wales, Aberstwyth.

- Christoffersen, P. (1998) "Evaluating Interval Forecasts", *International Economic Review*, 39, 841-862.
- Daal, E., and J.-S. Yu (2005), "A Comparison of Mixed GARCH-Jump Models with Skewed t-Distribution for Asset Returns", *SSRN Working Paper Series*, i.d. 670502
- Ding, Z., C. W. Granger, and R. F. Engle (1993), "A Long Memory Property of Stock Market Returns and a New Model", *Journal of Empirical Finance*, 1, 83-106.
- Engle, R. F., and S. Manganelli (2004), "CAViaR: Conditional Autoregressive Value at Risk by Regression Quantiles", *Journal of Business and Economic Statistics*, 22(4), 367-381.
- Gilli, M., and E. Kellezi (2006), "An Application of Extreme Value Theory for Measuring Financial Risk", *Computational Economics*, 27, 207-228.
- Giot, P., and S. Laurent (2004), "Modelling Daily Value-at-Risk Using Realized Volatility and ARCH Type Models", *Journal of Empirical Finance*, 11, 379-398.
- Jorion, P. (2007), *Value at Risk: The New Benchmark for Managing Financial Risk*, 3rd Ed. Singapore: McGraw-Hill.
- Knight, J., S. Satchell, and G. Wang (2003), "Value at Risk Linear Exponent (VARLINEX) Forecasts", *Quantitative Finance*, 3, 332-344.
- Kruse, R. (2006), "Can Realized Volatility Improve the Accuracy of Value-at-Risk Forecasts?", *Working Paper*.
- Kuester, K., S. Mittnik, and M. S. Paolella (2006), "Value-at-Risk

- Prediction: A Comparison of Alternative Strategies”, *Journal of Financial Econometrics*, 4(1), 53-89.
- Kupiec, P. (1995), “Techniques for Verifying the Accuracy of Risk Measurement Models”, *Journal of Derivatives*, 2, 173-184.
- McNeil, A. J., and R. Frey (2000), “Estimation of Tail-related Risk Measures for Heteroscedastic Financial Time Series: An Extreme Value Approach”, *Journal of Empirical Finance*, 7, 271-300.
- Mittnik, S., and M. S. Paolella (2000), “Conditional Density and Value-at-Risk Prediction of Asian Currency Exchange Rates”, *Journal of Forecasting*, 19, 313-333.
- RiskMetrics, 1995, *RiskMetrics Technical Document*, 3rd edn. J.P. Morgan.
- Stuart, A., and K. J. Ord (1994), *Kendall's Advanced Theory of Statistics vol 1 Distribution Theory*, (London: Arnold).
- Zellner, A. (1986), “Bayesian Estimation and Prediction Under Asymmetric Loss Functions”, *JASA*, 81, 446-51.

Chapter 4

A Copula-based Goodness-of-fit Test Based on Rosenblatt's Transformation

4.1 Introduction

The concept of a copula can be traced back to Höfding (1940) and Sklar (1959). Suppose there exists a vector of continuous random variables $\mathbf{X} = (X_1, \dots, X_d)$. Sklar's Theorem states that any joint distribution function H can be expressed as

$$H(x_1, \dots, x_d) = C(F_1(x_1), \dots, F_d(x_d)),$$

where F_1, \dots, F_d are uniform marginal distribution functions and C is a unique cumulative distribution function (CDF), known as a copula, with $C : [0, 1]^2 \rightarrow [0, 1]$.¹ Suppose $u_i = F_i(x_i)$, then $x_i = F_i^{-1}(u_i)$, and by a corollary of Sklar's theorem, $C(u_1, \dots, u_d) = H(F_1^{-1}(u_1), \dots, F_d^{-1}(u_d))$. The copula C , it turns out, is simply a CDF with uniformly distributed marginals. Since F_i is the CDF of x_i , the u_i 's are

¹ C is unique as long as the F_i 's are continuous.

$Unif[0, 1]$ by the Probability Integral Transform Theorem. An important feature of using a copula-based approach to multivariate analysis is that it allows for separate specification of the marginal distributions and the dependence structure.

The copula function is a useful tool for modelling the co-movements between variables. One reason the copula is so useful is due to its' ability to model nonlinear types of dependence, whereas more traditional parametric methods rely on a linear dependence assumption. In the context of financial data, the issue of non-normality of asset returns has been well documented, dating back to the seminal work of Mandelbrot (1963). Often times the Normal distribution is used to model financial data due to its' simplicity. Of course the fact that much of the financial data being modelled is non-normal is problematic. In fact, the problem has been further exacerbated in recent years due to the higher levels of volatility witnessed, for instance, during the dot-com build-up and subsequent crash, the crash of 2008, and the summer sell-off's of 2010 and 2011. Not only is the data non-normal, but often times in multivariate analysis there is dependence in the tails, which is not captured by the covariance matrix of the multivariate Normal distribution.² This has created the need for a flexible method of accurately modelling co-movements between assets, markets, risk factors, etc. Copula functions are powerful tools in multivariate analysis due in large part to the flexibility afforded by the decoupling of the specification of the marginal distributions from the specification of the dependence structure of the variables. Copula functions can be used in many different areas of finance, including asset pricing, credit risk analysis, risk management, and portfolio management. The focus of this paper is on copula-based goodness-of-fit tests (hereafter, referred to as 'g-o-f tests' or 'tests'). The tests

²In other words, joint extreme events receive an unrealistically small probability of occurrence under the multivariate Normal based on historical data.

are used to decide which copula family is the most appropriate to model the variables in question. In principal, the tests can be used in any of the previously mentioned applications to finance; however, the present discussion will revolve around two areas where the tests have been widely applied in the literature.

First, the tests are used in the top-down approach of enterprise risk management.³ Enterprise risk management attempts to calculate one firm-wide economic capital value that combines the economic capital requirements resulting from the different types of risk the firm in question is exposed to, such as, credit, market, operational, and liquidity risk. The capital requirements are meant to ensure the firm has enough capital available to absorb large unexpected losses. There are well-known methods to calculate the capital requirement for each type of risk,⁴ but the topic of how to map these individual requirements into a firm-wide requirement is relatively new.⁵ The top-down approach typically uses a copula function to model the dependence structure of the different risk types. Studies that use the top-down approach include Ward and Lee (2002), Dimakos and Aas (2004), Rosenberg and Schuermann (2006), Cech (2006), and Grundke (2010). There are two primary weaknesses with the top-down approach. First, the firm-wide capital requirement tends to be underestimated,⁶ and second, the g-o-f tests used to determine which copula family to use generally lack

³Also known as integrated risk management.

⁴See Jorion (2007) for an overview of such methods.

⁵It should be noted that the Basel II Accord recommends simply adding up the individual requirements to get the firm-wide requirement; however, this approach is undesirable since it ignores the effect of diversification. It assumes perfect positive correlation between each risk type, which is clearly an unrealistic assumption.

⁶Nevertheless, the copula-based estimates are an improvement over those obtained using a multivariate Normal distribution.

power,⁷ meaning it is difficult to tell which copula family should be used to model the data. The present paper addresses the second issue by introducing a new copula-based g-o-f test that has higher power than the alternatives currently available in the literature.

The second area where copula-based g-o-f tests have been widely applied is in the pair-copula decomposition method. The pair-copula decomposition method allows for the expression of a higher-dimensional copula (number of dimensions = $n > 2$) as a function of the $\binom{n}{2}$ combinations of bivariate copulas, as well as conditional and unconditional distribution functions.⁸ The benefit of such a decomposition is additional modelling flexibility; there are a large number of bivariate copulas, but only a few higher-dimensional alternatives. Moreover, a different copula family could potentially be chosen for each bivariate relationship, rather than being restricted to choosing one family to model an n -variate relationship. The pair-copula decomposition method has been the focus of Bedford and Cooke (2002), Kurowicka and Cooke (2006), and Aas, Czado, Frigessi, and Bakken (2009). The difficulty, once again, is how to choose which family to use for each relationship. Goodness-of-fit tests are also employed in this context, and the problem of insufficient power is once again present. As mentioned in the first application, the g-o-f test presented in this paper improves the situation by providing an alternative with higher power.

The contribution of this paper is three-fold; first, using the test statistic denoted $S_n^{(B)}$ in Genest et al. (2009) as a benchmark, a more powerful test is developed

⁷In this context, power refers to the probability of correctly rejecting a false null hypothesis. For example, assume the null hypothesis is that the data is well described by copula x when in fact it is not, then the power is the probability of rejecting this false null hypothesis.

⁸ $\binom{n}{2}$, read ‘ n choose 2’, is equal to $\frac{n!}{(n-2)!2!}$.

by implementing a weighting function in $S_n^{(B)}$,⁹ second, some guidance is given in terms of how to select an appropriate weighting function, and third, a link between the copula-based independence test literature and the copula-based goodness-of-fit test literature is explicitly spelled-out. The idea of using a weighting function in a copula-based test is not new; the inspiration comes from Medovikov (2011) who uses a weighting function in a copula-based test for independence. However, the method used to derive the weighting function is original. The new test statistic can be used by practitioners in any way they see fit, including either of the applications discussed earlier.

Section 4.2 provides more background on the copula function. Section 4.3 describes the existing copula-based g-o-f tests, proposes a new test, spells out the link between the copula-based independence test literature and the copula-based g-o-f literature, and describes the methodology used to derive the weighting function. A power study establishes the improvement of the new test over the benchmark in Section 4.4. In Section 4.5, the new test is implemented in a general application using financial data. Section 4.6 shows how the test can be applied in a pair-copula decomposition setting. Section 4.7 builds on the previous two sections with an application to risk management. Section 4.8 concludes.

⁹ $S_n^{(B)}$ is used as the benchmark because it was found to be the most powerful test in an extensive power study performed in Genest et al. (2009).

4.2 Introduction to Copula

A basic introduction was given in Section 4.1. This section will go into more depth on the topic. Section 4.2.1 looks at some properties of a copula function. Section 4.2.2 introduces several different dependence concepts that are useful in the context of copula functions. Section 4.2.3 describes several parametric families of copula. Section 4.2.4 discusses the issue of copula estimation.

4.2.1 Properties

Some general properties are outlined and then the bounds for a copula are defined.

Proposition 4.1 *A bivariate copula has the following properties:*

1. It is a function $C : [0, 1]^2 \rightarrow [0, 1]$
2. $C(u_1, 0) = C(0, u_2) = 0 \quad \forall (u_1, u_2) \in [0, 1]^2$
3. $C(u_1, 1) = u_1$ and $C(1, u_2) = u_2 \quad \forall (u_1, u_2) \in [0, 1]^2$.

Much like distribution functions, which are bounded by the so called *Fréchet-Höfding* bounds, copulas are also bounded. The bounds are given in the following theorem.

Theorem 4.1 *Let C be a copula. Then for all $(u_1, u_2) \in [0, 1]^2$,*

$$W(u_1, u_2) = \max(u_1 + u_2 - 1, 0) \leq C(u_1, u_2) \leq \min(u_1, u_2) = M(u_1, u_2).$$

A proof can be found in Nelsen (2006). W and M are called the *Fréchet-Höfding* lower bound, and *Fréchet-Höfding* upper bound, respectively, and they are themselves copulas. A final simple copula often of interest is the product (or independence) copula given by $C(u_1, u_2) = u_1 \cdot u_2$. All three copulas are shown in Figures 4.1-4.3 below.

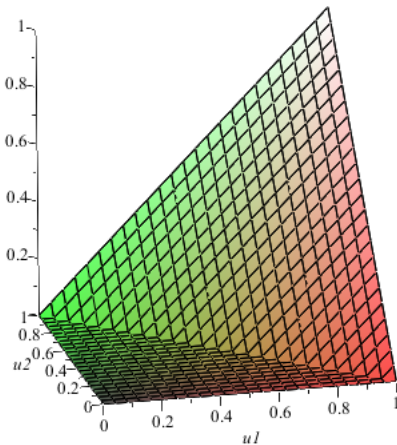


Figure 4.1: Minimum copula

4.2.2 Dependence Concepts

Over the years, a common method of modelling co-movements of several variables has been to assume the variables can be well described by a multivariate distribution (for

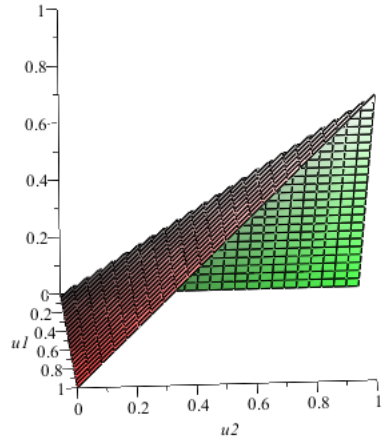


Figure 4.2: Maximum copula

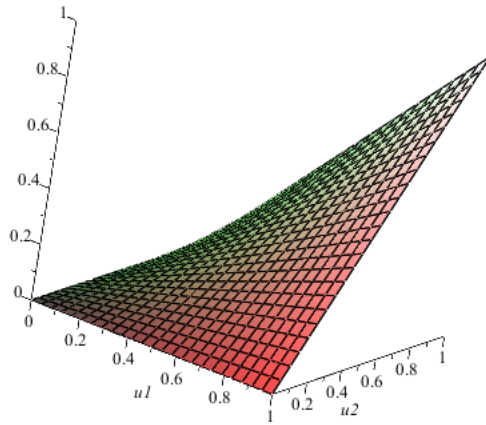


Figure 4.3: Product copula

example, Student-t or Normal). The problem with this method is that it assumes linear correlation between all variables. The copula method is much more flexible because it allows the user to model several different types of dependence, some of which are nonlinear. Several useful dependence concepts that can be utilized within a copula framework are discussed in this section.

Concordance

Concordance concepts attempt to measure the degree to which ‘high’ values of X_1 correspond to ‘high’ values of X_2 , and ‘low’ values of X_1 to ‘low’ values of X_2 . Specifically, observations (x_{1i}, x_{2i}) and (x_{1j}, x_{2j}) from a vector (X_1, X_2) of continuous random variables are concordant if $x_{1i} < x_{1j}$ and $x_{2i} < x_{2j}$, or if $x_{1i} > x_{1j}$ and $x_{2i} > x_{2j}$. (x_{1i}, x_{2i}) and (x_{1j}, x_{2j}) are said to be discordant if $x_{1i} < x_{1j}$ and $x_{2i} > x_{2j}$, or if $x_{1i} > x_{1j}$ and $x_{2i} < x_{2j}$. Two popular concordance measures are Kendall’s τ and Spearman’s ρ .

Kendall’s τ

According to Kruskal (1958), the notion of Kendall’s τ seems to have originated in the work of Fechner around 1897. Following Nelsen (2006), let $\{(x_{11}, x_{21}), (x_{12}, x_{22}), \dots, (x_{1n}, x_{2n})\}$ denote a random sample of n observations from a vector (X_1, X_2) of continuous random variables. There are $\binom{n}{2}$ pairs of observations and each pair has to be either concordant or discordant. Kendall’s τ for the sample

is

$$\tau_n = \frac{c - d}{c + d} = \frac{c - d}{\binom{n}{2}},$$

where c is the number of concordant pairs, and d is the number of discordant pairs. An interpretation of the previous equation is that τ_n is the probability of concordance minus the probability of discordance for some pair of observations (x_{1i}, x_{2i}) and (x_{1j}, x_{2j}) chosen randomly from the sample. Similarly, for *iid* random vectors (X_{11}, X_{21}) and (X_{12}, X_{22}) , the population version of Kendall's τ is given by

$$\tau = P[(X_{11} - X_{12})(X_{21} - X_{22}) > 0] - P[(X_{11} - X_{12})(X_{21} - X_{22}) < 0].$$

Finally, the following definition shows how concordance can be defined in terms of copulas.

Definition 4.1 *Kendall's τ is given by*

$$\tau = 4 \int \int_{I^2} C(u_1, u_2) dC(u_1, u_2) - 1,$$

for any copula C .

Since Kendall's τ is by far the most commonly used dependence concept in the copula finance literature, only brief descriptions of the remaining concepts are provided. A more rigorous treatment can be found in Cherubini, Luciano, and Vecchiato (2004), and Nelsen (2006).

Spearman's ρ

Spearman's ρ represents the correlation of the ranked sample of observations. It is expressed in terms of a copula function in Definition 4.2.

Definition 4.2 *For random variables X_1 and X_2 with copula C , Spearman's ρ is given by*

$$\rho_S = 12 \int \int_{I^2} C(u_1, u_2) du_1 du_2 - 3.$$

Durbin and Stuart (1951) show that there exists a well-defined relationship between Kendall's τ and Spearman's ρ . Specifically, for a given copula

$$\begin{cases} \frac{3}{2}\tau - \frac{1}{2} \leq \rho_S \leq \frac{1}{2} + \tau - \frac{1}{2}\tau^2, & \text{for } \tau \geq 0 \\ -\frac{1}{2} + \tau + \frac{1}{2}\tau^2 \leq \rho_S \leq \frac{3}{2}\tau + \frac{1}{2}, & \text{for } \tau < 0. \end{cases}$$

One benefit of using copulas is that the user is not restricted to using such simplistic measures as the linear correlation coefficient; nevertheless, it is a measure of dependence that may be useful in certain instances.

Linear Correlation Coefficient

It should be noted that the linear correlation coefficient does not qualify as a measure of concordance.

Definition 4.3 For non-degenerate random variables X_1 and X_2 , the linear correlation coefficient is given by

$$\rho = \frac{\text{cov}(X_1, X_2)}{\sqrt{\text{var}(X_1)\text{var}(X_2)}}$$

Tail Dependence

This section refers to tail dependence in the bivariate case only. One way of thinking about tail dependence is that it is a measurement of concordance in extreme values of X_1 and X_2 .

Definition 4.4 Let X_1 and X_2 be continuous random variables with distribution functions F_1 and F_2 , respectively, and let C be the copula of X_1 and X_2 . The upper tail dependence parameter λ_U is given by

$$\lambda_U = \lim_{t \rightarrow 1^-} P[X_2 > F_2^{(-1)}(t) | X_1 > F_1^{(-1)}(t)] = 2 - \lim_{t \rightarrow 1^-} \frac{1 - C(t, t)}{1 - t}$$

if it exists. Similarly, the lower tail dependence parameter λ_L is given by

$$\lambda_L = \lim_{t \rightarrow 0^+} P[X_2 \leq F_2^{(-1)}(t) | X_1 \leq F_1^{(-1)}(t)] = \lim_{t \rightarrow 0^+} \frac{C(t, t)}{t}$$

if it exists.

Later in this study three copula families will be studied; namely, Clayton, Gumbel, and Normal. The tail dependence measure of these three copulas are presented in Table 4.1. Figures 4.4-4.6 show a scatter plot of the three copulas in Table 4.1.

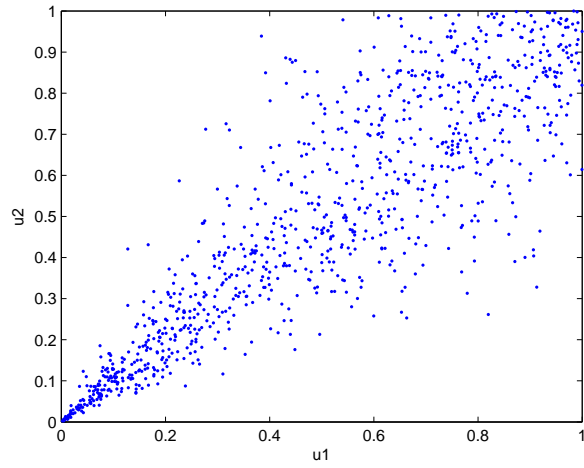


Figure 4.4: Clayton scatter plot $n = 1000$, $\tau = 0.7$

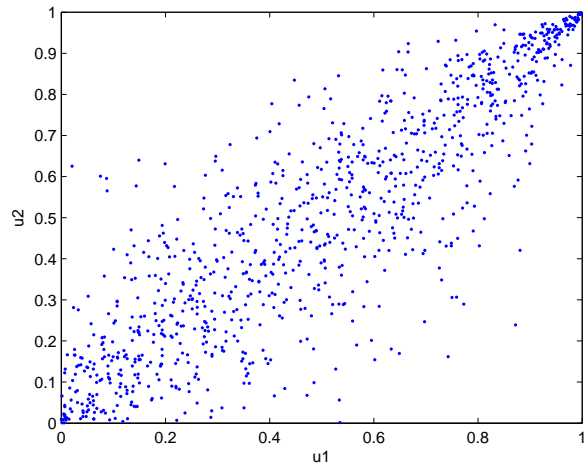
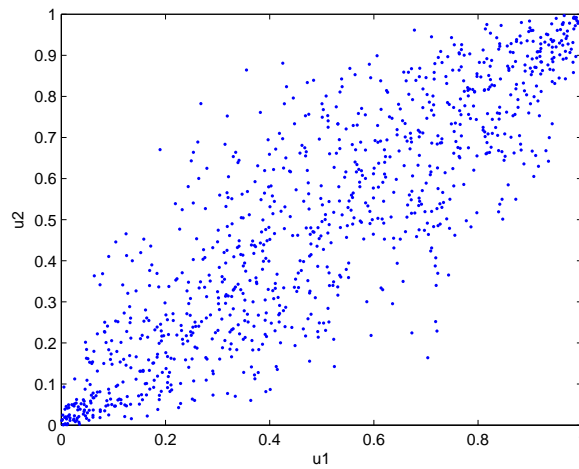


Figure 4.5: Gumbel scatter plot $n = 1000$, $\tau = 0.7$

Table 4.1: Tail dependence parameters for several copulas

Family	λ_L	λ_U
Clayton	$2^{(-1/\theta)}$	0
Gumbel	0	$2 - 2^{(1/\theta)}$
Normal	0	0

Figure 4.6: Normal scatter plot $n = 1000$, $\tau = 0.7$

The dense collection of points in the bottom-left corner of Figure 4.4 is an indication of lower tail dependence. Similarly, the dense collection of points in the upper-right corner of Figure 4.5 is an indication of upper tail dependence. The Normal copula, shown in Figure 4.6, does not have tail dependence; hence, no dense collection of points is visible in the tails.

Quadrant Dependence

The concept of quadrant dependence was introduced by Lehmann (1966), and recast in terms of copulas in Cherubini et al. (2004).

Definition 4.5 *The random variables X_1 and X_2 are positive quadrant dependent iff*

$$C(u_1, u_2) \geq u_1 \cdot u_2 \quad \forall (u_1, u_2) \in [0, 1]^2.$$

Similarly, X_1 and X_2 are negative quadrant dependent iff

$$C(u_1, u_2) \leq u_1 \cdot u_2 \quad \forall (u_1, u_2) \in [0, 1]^2.$$

Of course $u_1 \cdot u_2$ is just the independence copula C_{\perp} , so the condition for positive (negative) quadrant dependence can be re-written as $C(u_1, u_2) \geq (\leq) C_{\perp}$.

4.2.3 Parametric Families of Copulas

Several parametric copula families are introduced in this section, including the Normal, and two Archimedean copulas (Clayton and Gumbel). This section only presents the bivariate version of each copula.

Normal Copula

Definition 4.6 *The normal copula can be defined as follows:*

$$\begin{aligned} C_N(u_1, u_2; \rho) &= \Phi(\Phi^{-1}(u_1), \Phi^{-1}(u_2)) \\ &= \int_{-\infty}^{\Phi^{-1}(u_1)} \int_{-\infty}^{\Phi^{-1}(u_2)} \frac{1}{2\pi} \sqrt{1-\rho^2} \exp \left\{ \frac{-(r^2+s^2-2\rho rs)}{2(1-\rho^2)} \right\} dr ds, \end{aligned}$$

where Φ is the standard Normal CDF.

Under certain circumstances the n -variate Normal copula will yield the n -variate joint standard normal distribution function.

Proposition 4.2 (Cherubini et al. (2004)) *The Normal copula generates the joint Normal standard distribution function, via Sklar's theorem, iff the marginals are standard Normal.*

Remember, one benefit of using copulas is that the dependence structure and the marginals can be modelled separately. Moreover, there is no need for the marginals to be from the same distribution as each other. If at least one of the marginals is not standard Normal, then the Normal copula will not generate the joint standard Normal distribution function. Figure 4.7 shows a scatter plot of the Normal copula with one Student-t-distributed margin and one Beta distributed margin. Histograms of each margin are shown alongside the scatter plot.

One can also speak in terms of the *density* of a copula.

Definition 4.7 The density $c(u_1, u_2)$ of copula $C(u_1, u_2)$ is given by

$$c(u_1, u_2) = \frac{\partial^2 C(u_1, u_2)}{\partial u_1 \partial u_2}$$

Then, the density of the Normal copula is given by

$$c_N(u_1, u_2; \rho) = \frac{1}{\sqrt{1 - \rho^2}} \exp \left\{ \frac{\zeta_1^2 + \zeta_2^2}{2} + \frac{2\rho\zeta_1\zeta_2 - \zeta_1^2 - \zeta_2^2}{2(1 - \rho^2)} \right\},$$

where $\zeta_1 = \Phi^{-1}(u_1)$ and $\zeta_2 = \Phi^{-1}(u_2)$. Figure 4.8 shows the Normal copula density for $\rho = 0.8$.

Archimedean Copulas

Archimedean copulas can be constructed using a so called generator ϕ .

Definition 4.8 (Nelsen (2006)) Let ϕ be a continuous, strictly decreasing function from \mathbf{I} to $[0, \infty]$ such that $\phi(1) = 0$. The pseudo-inverse of ϕ is the function $\phi^{[-1]}$ with $\text{Dom}\phi^{[-1]} = [0, \infty]$ and $\text{Ran}\phi^{[-1]} = \mathbf{I}$ given by

$$\phi^{[-1]} = \begin{cases} \phi^{-1}(t), & 0 \leq t \leq \phi(0) \\ 0, & \phi(0) \leq t \leq \infty \end{cases}$$

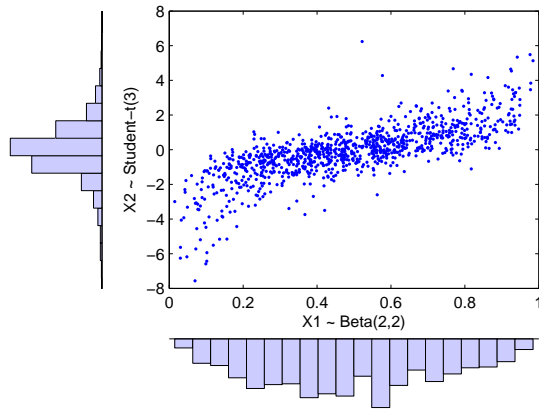


Figure 4.7: Normal copula scatter plot with Student-t(3) and Beta(2,2) marginals

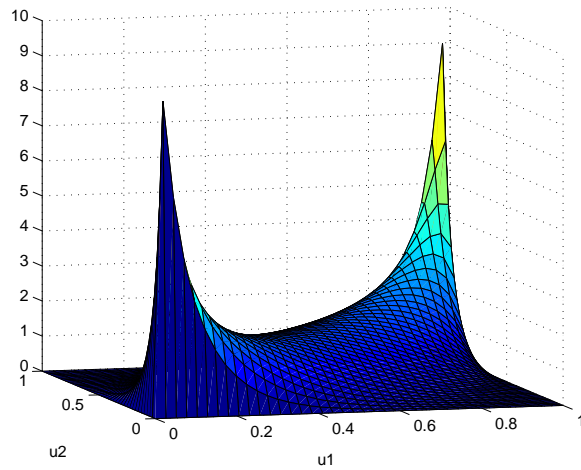


Figure 4.8: Normal copula density ($\rho = 0.8$)

Definition 4.9 (Cherubini et al. (2004)) *Given a generator and its pseudo-inverse, an Archimedean copula C^A is generated as follows:*

$$C^A(u_1, u_2) = \phi^{[-1]}(\phi(u_1) + \phi(u_2))$$

Clayton Copula

Introduced by Clayton (1978), the Clayton copula is derived using the generator $\phi_\theta(t) = \frac{1}{\theta}(t^{-\theta} - 1)$. The expression for the Clayton copula is

$$C_C(u_1, u_2; \theta) = \max\left[\left(u_1^{-\theta} + u_2^{-\theta} - 1\right)^{\frac{-1}{\theta}}, 0\right], \quad \theta \in [-1, \infty) \setminus \{0\}.$$

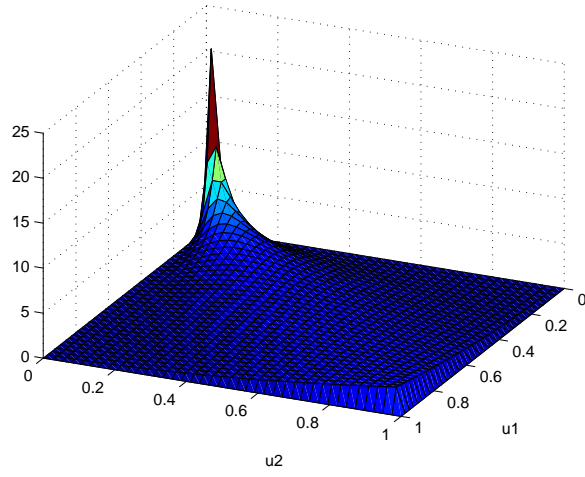
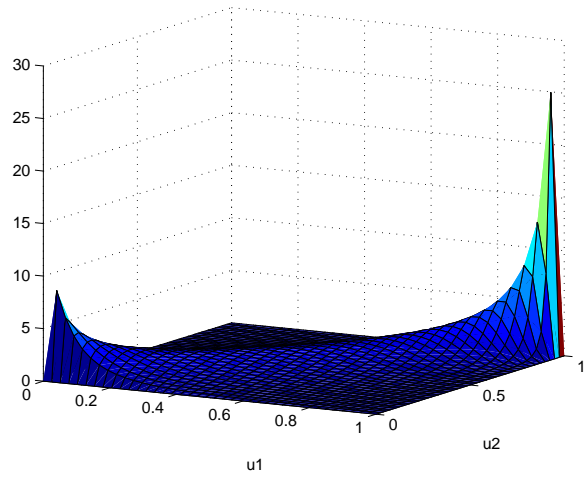
Figure 4.9 shows the density of the Clayton copula for $\theta = 2$. Notice the lower tail dependence, which was shown to exist in Table 4.1.

Gumbel Copula

The Gumbel copula was developed by Gumbel (1960). The generator used to derive the Gumbel copula is $\phi_\theta(t) = -\log(t)^\theta$. The Gumbel copula can be expressed as follows:

$$C_G(u_1, u_2; \theta) = \exp\left\{-\left[(-\log(u_1))^\theta + (-\log(u_2))^\theta\right]^{\frac{1}{\theta}}\right\}, \quad \theta \in [1, \infty).$$

Figure 4.10 shows the density of the Gumbel copula for $\theta = 3$. Notice the upper tail dependence, which was shown to exist in Table 4.1. Many more examples of copula

Figure 4.9: Clayton copula density ($\theta = 2$)Figure 4.10: Gumbel copula density ($\theta = 3$)

functions, Archimedean and otherwise, can be found in Joe (1997).

4.2.4 Parameter Estimation

Two groups of parameters need to be estimated: the parameters for each of the marginal distributions, and the copula parameters. In the context of copulas, there are several well-known and accepted estimation methods which will be outlined in this section.

Exact Maximum Likelihood Method

Let f be the joint probability density function (pdf) of X_1 and X_2 .¹⁰ As shown in, for example, Cherubini et al. (2004), the density of X_1 and X_2 , f , and the copula density c are related as follows:

$$f(x_1, x_2) = c(F_1(x_1), F_2(x_2)) \cdot f_1(x_1) \cdot f_2(x_2),$$

where c is as defined in section 4.2.3 (Definition 4.7). Then, the log-likelihood function is

$$l(\theta) = \sum_{t=1}^T \log(c(F_1(x_{1t}), F_2(x_{2t}))) + \sum_{t=1}^T \log(f_1(x_{1t})) + \log(f_2(x_{2t})),$$

where $\mathbf{X} = \{x_{1t}, x_{2t}\}_{t=1}^T$ is the sample data matrix, and θ is the set of all parameters from both the marginals and the copula. The value for θ that maximizes the function $l(\theta)$ is the maximum likelihood estimator, denoted $\hat{\theta}$.

¹⁰Assume all random variables are continuous in this section.

Inference for the Marginals (IFM) Method

Due to the computational complexity of the exact MLE method, the IFM method (proposed by Joe and Xu (1996)) is introduced as a more computationally efficient alternative. Instead of estimating all of the parameters jointly, the IFM method breaks the problem down into two steps. The first step is to estimate the parameters of the marginals. This is done by performing MLE on the second summation in the $l(\theta)$ equation:

$$\hat{\theta}_1 = \text{ArgMax}_{\theta_1} \sum_{t=1}^T \log(f_1(x_{1t}; \theta_1)) + \log(f_2(x_{2t}; \theta_1)).$$

Then, given the estimates ($\hat{\theta}_1$) from the first step, the copula parameters are estimated

$$\hat{\theta}_2 = \text{ArgMax}_{\theta_2} \sum_{t=1}^T \log(c(F_1(x_{1t}), F_2(x_{2t}); \theta_2, \hat{\theta}_1)).$$

Finally, the IFM estimator is $\hat{\theta}_{IFM} = (\hat{\theta}_1, \hat{\theta}_2)$. Joe (1997) notes that the IFM method is very efficient relative to the MLE method.

Inversion of Kendall's τ

Following Genest et al. (2009), the estimation method implemented in this paper is to solve for the copula parameter(s) in the equation for the sample version of

Kendall's τ given in Section 4.2.2:

$$\tau = 4 \int \int_{I^2} C(u_1, u_2; \theta) dC(u_1, u_2; \theta) - 1. \quad (4.1)$$

The sample τ can be calculated from the data using the equation given earlier:

$$\hat{\tau}_n = \frac{c - d}{\binom{n}{2}},$$

where c is the number of concordant pairs, and d the number of discordant pairs. Substituting $\hat{\tau}_n$ into equation (4.1) and solving for θ yields an estimate $\hat{\theta}$. The marginals are estimated using the empirical distributions, so no parameter estimation is necessary for the marginals.¹¹

4.3 Copula-based Goodness-of-fit Tests

The goodness-of-fit tests discussed in this section test the null hypothesis that the dependence structure between two or more variables in a given data set is well-described by a particular copula family against the complementary alternative. In other words,

$$H_0 : C \in C_0 = \{C_\theta : \theta \in \text{Int}(\Theta)\} \quad \text{vs.} \quad H_1 : C \notin C_0 = \{C_\theta : \theta \in \text{Int}(\Theta)\},$$

where Θ is the parameter space. The central focus in g-o-f testing is how well the copula fits the data, as a result, distributional assumptions on the marginals are unnecessary. Moreover, as discussed in Genest et al. (2009) (pg. 200), modelling

¹¹No assumption is made regarding the parametric form of the marginals when they are estimated by empirical distributions.

the marginals would have the effect of narrowing the null hypothesis. The standard procedure is to use rank data, where R_{ij} is the rank of X_{ij} among X_{1j}, \dots, X_{nj} (where the smallest observation has rank 1) for a given data set $\mathbf{X}_1 = (X_{11}, \dots, X_{1d}), \dots, \mathbf{X}_n = (X_{n1}, \dots, X_{nd})$.¹² Then pseudo-observations are defined as

$$U_j = (U_{1j}, \dots, U_{dj}) = \left(\frac{R_{1j}}{n+1}, \dots, \frac{R_{dj}}{n+1} \right),$$

where $n+1$ is used in the denominator instead of n to avoid numerical problems near the boundary of $[0, 1]^d$. The pseudo-observations can be thought of as a sample from the underlying copula C . Remember the original data $[\mathbf{X}_1, \dots, \mathbf{X}_n]$ was assumed to be independent; however, dependence is now present in the pseudo-observations because of the rank transformation. In addition, the pseudo-observations are only approximately uniform on $[0, 1]$. As a result, a parametric bootstrap procedure is needed to obtain reliable p-values. The bootstrap procedure will be outlined in Section 4.3.2.

4.3.1 Previous Methods

Genest et al. (2009) and Berg (2009) conduct comprehensive studies of copula-based g-o-f tests. Here, the focus will be on tests that are closely related to the new test presented later in this section. Fermanian (2005) and Genest and Rémillard (2008) consider a test of the form

$$\sqrt{n}(C_n - C_{\theta_n}),$$

¹² $[\mathbf{X}_1, \dots, \mathbf{X}_n]$ are assumed to be independent.

where C_n is the so-called empirical copula defined as

$$C_n = \frac{1}{n} \sum_{i=1}^n \mathbf{1}(U_{i1} \leq u_1, \dots, U_{id} \leq u_d),$$

where $(u_1, \dots, u_d) \in [0, 1]^d$, and C_{θ_n} is an estimate of C under H_0 .¹³ Genest and Rémillard (2008) examine rank-based versions of Cramér-von Mises and Kolmogorov-Smirnov statistics.

$$T_1 = n \int_{[0,1]^d} (C_n(\mathbf{u}) - C_{\theta_n}(\mathbf{u}))^2 dC_n(\mathbf{u}) \quad \text{and} \quad T_2 = \sup_{\mathbf{u} \in [0,1]^d} |\sqrt{n}(C_n(\mathbf{u}) - C_{\theta_n}(\mathbf{u}))|.$$

P-values are obtained by way of a parametric bootstrap procedure. The next test relies on Rosenblatt's Transformation, which is now discussed and defined.

Rosenblatt's Transformation

Rosenblatt (1952) introduces a method of transforming a vector of dependent variables into a vector of independent variables that are uniformly distributed on $[0, 1]$, so long as the original vector is from a known distribution. The method is based on the probability integral transform. The formal definition of the Rosenblatt Transformation is presented here in the context of copula functions.

Definition 4.10 (Rosenblatt's Transformation) *Let $\mathbf{u} = (u_1, \dots, u_d) \in [0, 1]^d$ denote a random vector that is distributed as C . Let $u_1^* = u_1$, and for each $i \in \{2, \dots, d\}$,*

$$u_i^* = \frac{\partial^{i-1} C(u_1, \dots, u_i, 1, \dots, 1)}{\partial u_1 \cdots \partial u_{i-1}} \bigg/ \frac{\partial^{i-1} C(u_1, \dots, u_{i-1}, 1, \dots, 1)}{\partial u_1 \cdots \partial u_{i-1}}.$$

¹³It should be noted that although C_n is commonly referred to as an 'empirical copula', it is not, in fact, a copula.

Then, $\mathbf{u}^* = (u_1^*, \dots, u_d^*)$ is an independent $\text{Unif}[0, 1]^d$ vector.^{14,15}

It is clear that \mathbf{u} is distributed as the copula under H_0 , namely C_0 , if and only if the distribution of \mathbf{u}^* is the independence copula

$$C_{\perp}(u_1^*, \dots, u_d^*) = u_1^* \times \dots \times u_d^*, \quad (u_1^*, \dots, u_d^*) \in [0, 1]^d.$$

Thus, a goodness-of-fit test that determines whether \mathbf{u} is well-described by copula family C_0 can be thought of as an independence test that attempts to determine whether the components of \mathbf{u}^* are independent, i.e., well-described by the independence copula C_{\perp} .

The benchmark test used in this study is from Genest et al. (2009). The test statistic takes the form

$$\begin{aligned} T_3 &= n \int_{[0,1]^d} (D_n(\mathbf{u}) - C_{\perp}(\mathbf{u}))^2 d\mathbf{u} \\ &= \frac{n}{3^d} - \frac{1}{2^{d-1}} \sum_{i=1}^n \prod_{k=1}^d (1 - U_{ik}^*)^2 + \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n \prod_{k=1}^d (1 - U_{ik}^* \vee U_{jk}^*), \end{aligned}$$

where $i \vee k = \max(i, k)$, D_n is the empirical copula defined as $D_n(\mathbf{u}) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}(U_i^* \leq \mathbf{u})$, and $U_i^* = (u_{i1}^*, \dots, u_{id}^*)$. In the methods presented earlier (T_1 and T_2), the comparison was between the empirical copula and the estimate of the copula under H_0 using the pseudo-observations. Under this method, the comparison is between the empirical copula and the independence copula using the Rosenblatt transformed pseudo-observations. Remember, the variables should be approximately independent using

¹⁴In practice the components of \mathbf{u}^* are only approximately independently distributed, since the nonparametric marginal distributions introduce dependence.

¹⁵Explicit expressions for u_2^* can be found in the appendix of Aas et al. (2009) for each of the copulas used in this paper (for the bivariate case only).

the transformed observations as long as the data does in fact come from the copula assumed under H_0 . Thus, large values of this test statistic should lead to a rejection of the null hypothesis, indicating that the copula under H_0 does not fit the data well. Similar to the previous test, a parametric bootstrap procedure is used to obtain p-values.

4.3.2 New Test

The proposed test statistic is

$$T_{3W} = n \int_{[0,1]^d} (D_n(\mathbf{u}) - C_{\perp}(\mathbf{u}))^2 w(\mathbf{u}) d\mathbf{u}.$$

The empirical copula $D_n(\mathbf{u})$ is also a function of the estimate $\hat{\theta}_n$ of the copula parameter θ , but this notation will be suppressed for convenience. Medovikov (2011) shows that the asymptotic covariance function of the process $\sqrt{n}(D_n(\mathbf{u}) - C_{\perp}(\mathbf{u}))\sqrt{w(\mathbf{u})}$ is finite as long as $w(\mathbf{u}) < \infty$ for all $\mathbf{u} \in [0, 1]^d$.¹⁶ The weighting functions explored in this paper satisfy this requirement, therefore the asymptotic covariance function will always be finite and the distribution of the test statistic T_{3W} is non-degenerate.¹⁷ Earlier, it was mentioned that the components of \mathbf{u}^* are only approximately independently distributed because the nonparametric marginals induce dependence. Another problem is that small sample estimation error is present from the estimation of the

¹⁶This result is only shown for the two-dimensional case, i.e. $d = 2$.

¹⁷Note that the test statistic in Medovikov (2011) uses pseudo-observations, while T_{3W} uses pseudo-observations which then have Rosenblatt's Transformation applied to them. This transformation does not affect the stated result.

copula parameter under the null hypothesis. For both of these reasons, it is necessary to employ a parametric bootstrap procedure to obtain reliable estimates of the p-value. An outline of the bootstrap procedure is now provided.

Parametric Bootstrap Procedure

1. Obtain sample data (x_1, \dots, x_n) , calculate ranks (R_1, \dots, R_n) , and calculate pseudo-observations (u_1, \dots, u_n) , where $u_i = \frac{R_i}{n+1}$ for all $i \in \{1, \dots, n\}$.
2. Estimate copula parameter(s) θ by $\hat{\theta}$ using the method of inversion of Kendall's τ .¹⁸
3. Compute (u_1^*, \dots, u_n^*) using Rosenblatt's Transformation, assuming the copula under the null hypothesis, $C_{\hat{\theta}}$.
4. Compute the value of the test statistic T_{3W} .
5. For some large integer B , repeat the following steps for every $b \in \{1, \dots, B\}$.
 - a) Generate a random sample $(x_{1,b}, \dots, x_{n,b})$ from the null copula $C_{\hat{\theta}}$, and compute the associated rank vectors $(R_{1,b}, \dots, R_{n,b})$, and pseudo-samples $(u_{1,b}^0, \dots, u_{n,b}^0)$.
 - b) Estimate copula parameter(s) θ by $\hat{\theta}_b^0$ using the method of inversion of Kendall's τ .
 - c) Compute $(u_{1,b}^{0*}, \dots, u_{n,b}^{0*})$ using Rosenblatt's Transformation, assuming the copula under the null hypothesis, $C_{\hat{\theta}_b^0}$.

¹⁸In this case it is the sample Kendall's τ calculated from the pseudo-observations (u_1, \dots, u_n) .

d) Let

$$D_{n,b}^0 = \frac{1}{n} \sum_{i=1}^n \mathbf{1}(u_{i,b}^{0*} \leq \mathbf{u}), \quad \mathbf{u} \in [0, 1]^d.$$

e) Compute

$$T_{3W,b}^0 = n \int_{[0,1]^d} (D_{n,b}^0(\mathbf{u}) - C_{\perp}(\mathbf{u}))^2 w(\mathbf{u}) d\mathbf{u}.$$

6. Then, an approximate p-value for the test is given by

$$\frac{1}{B} \sum_{b=1}^B \mathbf{1}(T_{3W,b}^0 > T_{3W}).$$

Whenever a copula parameter (θ) needed to be estimated, this was accomplished via inversion of Kendall's τ . This amounts to solving for θ in the following equation:

$$4 \int_{[0,1]^2} C_{\theta}(u_1, u_2) dC_{\theta}(u_1, u_2) - 1 = \tau_n,$$

where τ_n is the sample version of τ . There are unique expressions for θ for each of the copula families considered in this paper. For example, in the case of the Normal copula, $\theta = \sin((\pi/2)\tau_n)$.

4.3.3 Connection to Independence Test Literature

First, a brief description of a few copula-based independence tests that are relevant to the current discussion will be described. Kojadinovic and Holmes (2009) develop

a copula-based independence test of the form

$$I_1 = n \int_{[0,1]^d} (D_n(\mathbf{u}) - C_{\perp}(\mathbf{u}))^2 d\mathbf{u}.^{19}$$

This is the same form as the g-o-f test (T_3) from Genest et al. (2009); the difference is that I_1 tests independence of the components of \mathbf{u} rather than \mathbf{u}^* . Medovikov (2011) extends I_1 by implementing a weighting function:

$$I_{1W} = n \int_{[0,1]^d} (D_n(\mathbf{u}) - C_{\perp}(\mathbf{u}))^2 w(\mathbf{u}) d\mathbf{u}.$$

Now, let's review what has been learned. First, for some data set $\mathbf{X} = (X_1, \dots, X_d)$, where $X_i = (X_{1i}, \dots, X_{ni})$ for $i = (1, 2, \dots, d)$, the independence of the components of \mathbf{X} , i.e., (X_1, \dots, X_d) , can be tested using tests I_1 or I_{1W} . In these tests, the null hypothesis is that the components of \mathbf{X} are independent, or in other words, that the distribution of the pseudo-observations (denoted \mathbf{u} earlier) is the independence copula. Second, Rosenblatt's Transformation can be used to convert a vector of dependent variables into a vector of (approximately) independent $Unif[0, 1]^d$ variables (denoted \mathbf{u}^* earlier). Third, using the tests I_1 and I_{1W} to test for the independence of the components of \mathbf{u}^* (i.e., $H_0 : \mathbf{u}^* = (u_1^*, \dots, u_d^*) \sim C_{\perp}(\mathbf{u}^*)$) is equivalent to testing whether the components of \mathbf{u} are well-described by the copula under the null hypothesis (i.e., $H_0 : C(\mathbf{u}) \in C_0 = (C_{\theta} : \theta \in \Theta)$). Specifically, the test statistic presented in Medovikov (2011) (I_{1W}) takes the same form as the new test statistic presented here (T_{3W}), and I_1 from Kojadinovic and Holmes (2009) is the same as test statistic T_3 from Genest et al. (2009). Hence, it is now clear how a simple data transformation

¹⁹Here, an assumption is made that $d_i = 1$ for all $i \in \{1, \dots, p\}$ on page 1140 of Kojadinovic and Holmes (2009).

can convert an independence test into a goodness-of-fit test.

4.3.4 Weighting Function

The choice of weighting function revolves around the idea of exploiting the areas within the unit square where strong dependence is exhibited. Fortunately, as long as the sample size and τ are not too small, strong and exploitable patterns do indeed emerge.²⁰ Here, one sensible method of attempting to incorporate such patterns into a weighting function will be outlined. The method involves the object

$$DEPFUNC = |D_n(\mathbf{u}) - C_{\perp}(\mathbf{u})|. \quad (4.2)$$

Dependence is present where *DEPFUNC* is large, or, in other words, dependence is present when the empirical copula is far away from the independence copula. The weighting function should be constructed such that observations receive more weight where dependence is present and less weight where dependence is absent. The best way to accomplish this is to fit a function to the surface given by *DEPFUNC*. Conveniently, MATLAB has a surface fitting toolbox that is perfect for this problem.²¹ Through trial and error, the polynomial class of functions proved to provide the best fit (in terms of adjusted R^2), of those functions that were integrable.^{22,23} The

²⁰What is meant by ‘not too small’ will be discussed in more detail in Section 4.4.

²¹The version of MATLAB referred to throughout this paper is MATLAB 7.8.0.

²²The method fits a polynomial surface using the least absolute residual robust (LAR) method.

²³Integrability is necessary in order to keep the problem tractable.

weighting function used in this paper takes the form²⁴

$$\begin{aligned}
 w(u_1, u_2) = & p_{00} + p_{10}u_1 + p_{01}u_2 + p_{20}u_1^2 + p_{11}u_1u_2 + p_{02}u_2^2 + p_{30}u_1^3 \\
 & + p_{21}u_1^2u_2 + p_{12}u_1u_2^2 + p_{03}u_2^3 + p_{40}u_1^4 + p_{31}u_1^3u_2 \\
 & + p_{22}u_1^2u_2^2 + p_{13}u_1u_2^3 + p_{04}u_2^4 + p_{50}u_1^5 + p_{41}u_1^4u_2 \\
 & + p_{32}u_1^3u_2^2 + p_{23}u_1^2u_2^3 + p_{14}u_1u_2^4 + p_{05}u_2^5.
 \end{aligned}$$

The adjusted R^2 of the fit is typically in the 0.9 to 0.99 range, indicating that it would be difficult to improve on this method. Next, the method used to choose a specific weighting function is described.

First, it is necessary to define the universe of copula families at the users disposal. This study will be restricted to the use of the Normal, Clayton, and Gumbel copulas. The next step is to search for exploitable patterns in the function *DEPFUNC*. Assume some copula under H_0 , call it $A_0 \in A = \{Clayton, Gumbel, Normal\}$, and assume some true copula $A_T \in A$ comes from the pre-specified universe, then patterns will be exploitable if:

1. The graph of *DEPFUNC* consistently takes a particular form, call it F_0 , when the true copula is the same as the copula under H_0 , i.e., $A_T = A_0$.
2. The graph of *DEPFUNC* consistently takes a particular form for each possible copula pairing when the true copula is not A_0 , where the shapes of *DEPFUNC* in step 2 are different from the shape F_0 in step 1.²⁵

²⁴The polynomial uses five degrees along each dimension, which is the largest number of degrees permissible by the surface fitting toolbox in MATLAB.

²⁵The shapes in step 2 can be the same as each other or different, just as long as they are both different than the shape in step 1.

For clarity, a number of graphs will now be shown. Figure 4.11 shows *DEPFUNC* when the copula under H_0 and the simulated data come from the Clayton family. Note that each figure shows an average *DEPFUNC* over 1000 simulations. In general, each individual *DEPFUNC* may look different than the average shown in the figure, but convergence occurs quickly. For instance, in this case, the average *DEPFUNC* after 50 simulations looks very similar to the average *DEPFUNC* after 1000 simulations.²⁶ High regions in the figures are areas where dependence is present, while

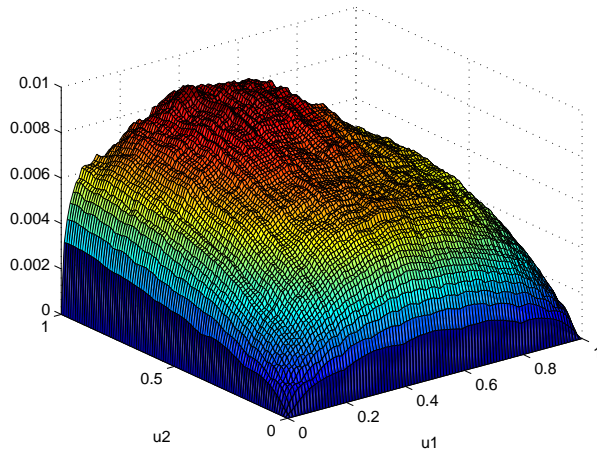


Figure 4.11: Clayton vs. Clayton $n = 500$, $\tau = 0.2$

low regions represent independence. Figure 4.12 shows *DEPFUNC* when the copula under H_0 is Clayton and the simulated data (true process) is Gumbel.

Finally, Figure 4.13 shows *DEPFUNC* when the copula under H_0 is Clayton and the simulated data is from the Normal copula family. The exercise is repeated using

²⁶The stated convergence result holds for most (H_0, H_1) pairings, but not all. Later in the paper, it will be shown that the Normal vs. Gumbel and Gumbel vs. Normal pairings do not converge very quickly.

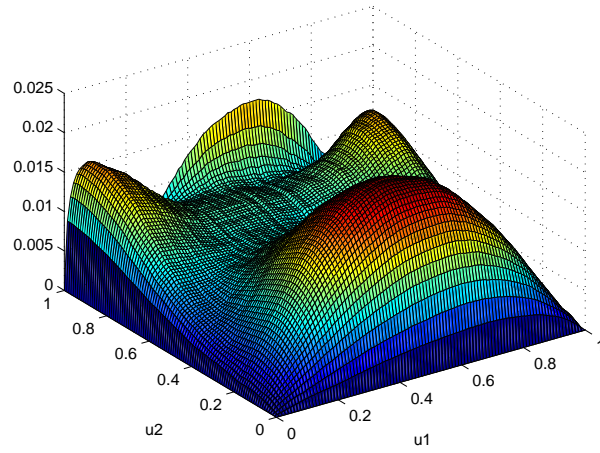


Figure 4.12: Clayton vs. Gumbel $n = 500$, $\tau = 0.2$

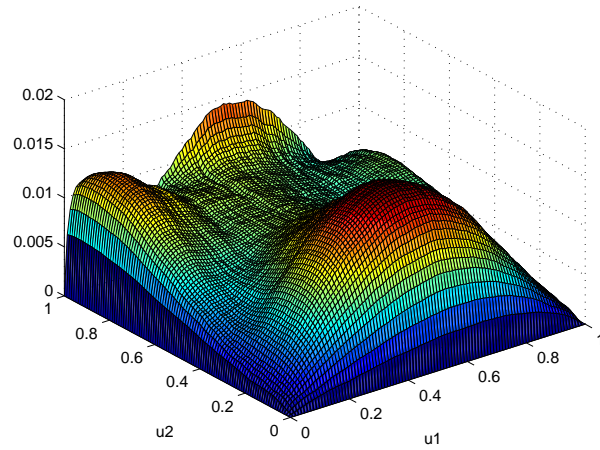


Figure 4.13: Clayton vs. Normal $n = 500$, $\tau = 0.2$

the Gumbel and Normal copulas under H_0 , shown in Figures D.1 through D.6 in the appendix.

One may wonder why any dependence shows up at all in Figures 4.11, D.1, and D.4, when the data comes from the copula under H_0 . In these cases, theoretically, the empirical copula should be equal to the independence copula. However, remember that in practice Rosenblatt's Transformation will only yield an empirical copula that is approximately independent. Then, it is perfectly reasonable to expect to see *some* dependence present in the figures in question. Also, notice that the maximum values of *DEPFUNC* in Figures 4.11, D.1, and D.4 are smaller than that in those cases where the data does not come from the copula under H_0 , indicating that there is, in fact, less dependence. A very interesting observation is that the shape of *DEPFUNC* is essentially identical for all copula families when the data comes from the copula under H_0 . Possible explanations will be deferred to future work on the topic.

The main point to take away from these pictures is that the dependence pattern when H_0 is true is markedly different compared to the various patterns when H_0 is false. Suppose the copula under H_0 is from copula family A_0 , but there is also reason to believe the data could be from copula family $A_1 \in A$, then the weighting function is determined by simulating data from A_1 , applying Rosenblatt's Transformation assuming copula A_0 , and evaluating *DEPFUNC* at each point on a sufficiently fine grid of the unit square, then fitting a function to *DEPFUNC* in the manner described earlier. The sample that was generated may be an outlier, so this procedure is repeated 1000 times and the final weighting function that goes into the test statistic is the average of 1000 weighting functions. When the data does in fact come from A_1 ,

the weighted test statistic should have greater power than the unweighted version.²⁷ When it comes to an application of this method using real data, some method is needed to determine what is the ‘best alternative’ to H_0 ; a commonly used qualitative method is described in Section 4.5. Finally, notice that the weighting function only needs to be calculated once for each parameterization.²⁸ If the same parameterization is used in the future, the same weighting function can be used; however, every unique parameterization requires a unique weighting function. Fortunately, the procedure used to find the appropriate weighting function is not very computationally intensive.

4.4 Power Study

In this section, the new test statistic, T_{3W} , is pitted against the benchmark, T_3 , in a Monte Carlo simulation to determine which has greater power. The possible parameter values are:

copula under $H_0 \in \{\text{Clayton, Gumbel, Normal}\}$

true copula $\in \{\text{Clayton, Gumbel, Normal}\}$

$n \in \{100, 300, 500\}$

$\tau \in \{0.2, 0.4\}$

²⁷This is confirmed in the power study in Section 4.4.

²⁸The parameters are: 1) the copula under H_0 , 2) the ‘best alternative’ copula, 3) n , and 4) τ .

In total there are $3 \times 3 \times 3 \times 2 = 54$ parameterizations. Whenever the copula under H_0 is the same as the true copula, the relevant test is whether the test statistics are able to hold their nominal level, which is set at 5% in this study.²⁹ For $n = 100$, 1000 random samples were generated, and for each random sample 1000 bootstrap samples were generated. For $n = 300$, only 500 random samples and 1000 bootstrap samples were used, and for $n = 500$, only 500 random samples and 500 bootstrap samples were generated due to computational constraints. In each random sample, the choice of whether or not H_0 can be rejected is based on the p-value obtained from the parametric bootstrap procedure presented earlier. The proportion of rejections is an estimate of the power.

Before discussing the results of the power study, a discrepancy between the results in Genest et al. (2009) and Berg (2009) needs to be addressed. Berg (2009) does not use test statistic T_3 , however he uses a similar test statistic of the form

$$T_{3b} = n \int_{[0,1]^d} (D_n(\mathbf{u}) - C_{\perp}(\mathbf{u}))^2 dD_n(\mathbf{u}).$$

This test statistic is also used in Genest et al. (2009). Although the two papers do not present identical parameterizations, it is clear that the results are drastically different. Before coding tests T_3 and T_{3W} , I coded test T_{3b} to see if I could replicate the results of either paper. As shown in Table 4.2 below, my results are very close to those of Berg (2009).³⁰ On the other hand, my code produces much different results from those presented in Genest et al. (2009).³¹

²⁹Remember, power is the probability of rejecting a false null hypothesis. In this case the null is true, so the probability of rejection should be equal to the nominal level.

³⁰Berg (2009) includes the Frank and t copulas in the universe of copulas. I include them as possible choices for the true copula simply to give the comparison more data points.

³¹A comparison table similar to Table 4.2 is available upon request.

Unfortunately, Berg (2009) does not use test T_3 . Adjusting the code to accommodate T_3 rather than T_{3b} is fairly simple, so I have every reason to believe my version of T_3 would have matched up with Berg's version of T_3 , had it been used in his study. I do not wish to make a claim regarding the validity of the results of either Genest et al. (2009) or Berg (2009); however, it seems clear that the appropriate comparison is between the results obtained from T_{3W} and the results obtained from my coded version of T_3 , rather than between the results obtained from T_{3W} and the results obtained from T_3 that are presented in Genest et al. (2009). Tables 4.3-4.5 shows the results of the size and power study.

Summary of Results

In general, the weighted test outperforms the unweighted test; however, one feature that stands out is that for low $\{n, \tau\}$ combinations there is not much of a difference between the two tests. The reason for this result is made clear by analyzing Figures 4.14 and 4.15.

Remember from earlier that the condition for an improvement in power is that the shape of *DEPFUNC* when the copula under H_0 is the true copula needs to be different from the shape of *DEPFUNC* when the copula under H_0 is not the true copula. Figures 4.14 and 4.15 are fairly similar, indicating there is no exploitable

Table 4.2: Percentage of rejection of H_0 - comparison between my T_{3b} and Berg's T_{3b}

	Copula under H_0	True copula	My T_{3b}	Berg's T_{3b}
$n = 100, \tau = 0.2$				
	Clayton	Frank	0.068	0.064
		Normal	0.0704	0.066
		Gumbel	0.1407	0.13
		t	0.095	0.084
	Normal	Clayton	0.178	0.196
		Frank	0.072	0.074
		Gumbel	0.044	0.036
		t	0.086	0.088
	Gumbel	Clayton	0.2826	0.311
		Frank	0.0886	0.08
		Normal	0.0655	0.073
		t	0.088	0.098
$n = 100, \tau = 0.4$				
	Clayton	Frank	0.2258	0.242
		Normal	0.0704	0.066
		Gumbel	0.3788	0.412
		t	0.164	0.16
	Normal	Clayton	0.572	0.596
		Frank	0.092	0.116
		Gumbel	0.062	0.05
		t	0.108	0.116
	Gumbel	Clayton	0.7166	0.754
		Frank	0.0962	0.094
		Normal	0.0804	0.089
		t	0.143	0.152

Table 4.3: Percentage of rejection of H_0 - weighted vs. unweighted ($n=100$)

	Copula under H_0	True copula	T_{3W}	T_3	T_{3b}
$n = 100, \tau = 0.2$					
	Clayton	Clayton	0.046	0.043	0.052
		Normal	0.106	0.115	0.066
		Gumbel	0.219	0.218	0.13
	Normal	Clayton	0.085	0.086	0.196
		Normal	0.044	0.049	0.046
		Gumbel	0.066	0.07	0.036
	Gumbel	Clayton	0.217	0.204	0.311
		Normal	0.039	0.04	0.073
		Gumbel	0.049	0.051	0.049
$n = 100, \tau = 0.4$					
	Clayton	Clayton	0.046	0.051	0.054
		Normal	0.42	0.4	0.264
		Gumbel	0.664	0.576	0.412
	Normal	Clayton	0.472	0.428	0.596
		Normal	0.038	0.04	0.054
		Gumbel	0.094	0.096	0.05
	Gumbel	Clayton	0.716	0.63	0.754
		Normal	0.052	0.054	0.089
		Gumbel	0.062	0.06	0.051

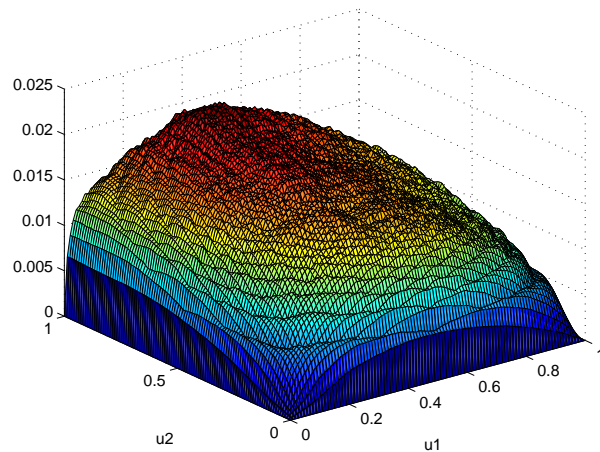
Figure 4.14: Clayton vs. Clayton $n = 100, \tau = 0.2$

Table 4.4: Percentage of rejection of H_0 - weighted vs. unweighted ($n=300$)

	Copula under H_0	True copula	T_{3W}	T_3
$n = 300, \tau = 0.2$				
	Clayton	Clayton	0.058	0.062
		Normal	0.384	0.352
		Gumbel	0.726	0.636
	Normal	Clayton	0.428	0.372
		Normal	0.054	0.052
		Gumbel	0.184	0.182
Gumbel	Clayton	0.82	0.75	
	Normal	0.158	0.16	
	Gumbel	0.058	0.058	
$n = 300, \tau = 0.4$				
	Clayton	Clayton	0.046	0.048
		Normal	0.944	0.916
		Gumbel	0.998	0.99
	Normal	Clayton	0.992	0.976
		Normal	0.058	0.052
		Gumbel	0.194	0.2
Gumbel	Clayton	1	0.996	
	Normal	0.19	0.186	
	Gumbel	0.054	0.056	

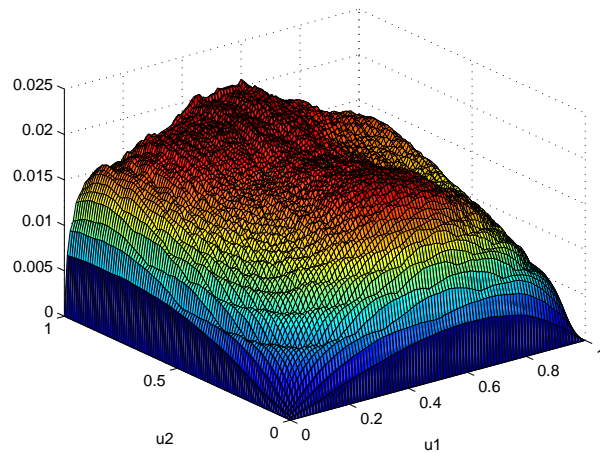
Figure 4.15: Clayton vs. Normal $n = 100, \tau = 0.2$

Table 4.5: Percentage of rejection of H_0 - weighted vs. unweighted (n=500)

	Copula under H_0	True copula	T_{3W}	T_3	T_{3b}
$n = 500, \tau = 0.2$					
	Clayton	Clayton	0.044	0.046	0.048
		Normal	0.66	0.604	0.448
		Gumbel	0.944	0.902	0.849
	Normal	Clayton	0.592	0.54	0.713
		Normal	0.052	0.056	0.048
		Gumbel	0.286	0.252	0.085
	Gumbel	Clayton	1	1	0.985
		Normal	0.272	0.268	0.239
		Gumbel	0.052	0.056	0.048
$n = 500, \tau = 0.4$					
	Clayton	Clayton	0.054	0.056	0.052
		Normal	1	0.982	0.995
		Gumbel	1	1	1
	Normal	Clayton	1	0.994	1
		Normal	0.05	0.046	0.054
		Gumbel	0.392	0.364	0.189
	Gumbel	Clayton	1	1	1
		Normal	0.316	0.294	0.384
		Gumbel	0.042	0.042	0.055

difference that could yield a more powerful test.³²

At the other end of the $\{n, \tau\}$ spectrum ($\{300, 0.4\}$ and $\{500, 0.4\}$), the power converges to one, as one would expect in a consistent test. This convergence only occurs for certain copula pairings. For $n = 500$, a power advantage remains for the weighted test in the copula pairings that have yet to converge (6.8% on average), but, oddly enough, not for $n = 300$. Thus, for $\{n, \tau\}$ pairs that are either ‘too small’ or ‘too large’ the weighted version of the test has virtually no advantage.³³ For the other $\{n, \tau\}$ pairs ($\{100, 0.4\}$, $\{300, 0.2\}$, and $\{500, 0.2\}$) the weighted test has a clear advantage over the unweighted test. The magnitude of the advantage does not appear to depend on either the null copula or the true copula. The average advantage over the unweighted test across all copula pairings is 5.2% for $\{100, 0.4\}$, 7.4% for $\{300, 0.2\}$, and 6% for $\{500, 0.2\}$.³⁴ For certain combinations of H_0 and H_1 the power advantage can reach as high as 15%. Finally, note that the nominal level is controlled well (bold values in Tables 4.3-4.5).

For two of the six (H_0, H_1) pairings (Normal vs. Gumbel and Gumbel vs. Normal), the weighted test slightly underperforms the unweighted test for $n = 100$ and $n = 300$.³⁵ A clue as to why the weighted test underperforms for these two pairings is given in the *DEPFUNC* plots. For example, compare Figures 4.16 and 4.17 which show *DEPFUNC* for Gumbel vs. Normal and Gumbel vs. Clayton, respectively.

³²The shape of *DEPFUNC* does not change much across all possible copula pairings for $n = 100$, $\tau = 0.2$.

³³The advantage for the weighted test, averaged across all copula pairings, is -1.7% for $n = 100$, $\tau = 0.2$, 0.2% for $n = 300$, $\tau = 0.4$, and 2.6% for $n = 500$, $\tau = 0.4$, for an average advantage across these three $\{n, \tau\}$ combinations of $\frac{-1.7+0.2+2.6}{3} = 0.4\%$.

³⁴Averaged across the three $\{n, \tau\}$ combinations, the advantage for the weighted test is 6.2% .

³⁵The difference is insignificant.

Figure 4.17 has nice smooth curves, while the curves in Figure 4.16 are a little rough. Remember, these *DEPFUNC* plots are an average of 1000 *DEPFUNC*s, so the roughness in Figure 4.16 is an indication that there is more volatility from sample-to-sample in the Gumbel vs. Normal case relative to the Gumbel vs. Clayton case. The weighting function is calculated based on the *average DEPFUNC* in each case, and a randomly chosen sample in the Gumbel vs. Normal case appears to be further away from its average *DEPFUNC* than in the Gumbel vs. Clayton case, thus decreasing the effectiveness of the weighting function.³⁶ A similar argument can be made to explain the underperformance of the weighted test in the Normal vs. Gumbel pairing.

A comparison to Berg's test T_{3b} is also included in Tables 4.3 and 4.5.³⁷ Overall, the weighted test performs better, but T_{3b} has higher power for certain parameterizations. For instance, T_{3b} tends to have higher power for Gumbel vs. Normal and Normal vs. Clayton. Also, T_{3b} has relatively better performance for smaller n . T_{3b} has similar power to the weighted test for $n = 100$, but significantly lower power for $n = 500$. Now that the advantage to using the weighted test has been established through simulation, examples of its practical use are shown in the following three sections.

³⁶The most effective weighting function would be a case where *DEPFUNC* is the same for every randomly chosen sample from a given pairing. Any variation decreases the effectiveness of the weighting function.

³⁷Berg does not display results for $n = 300$, so there is no such comparison in Table 4.4.

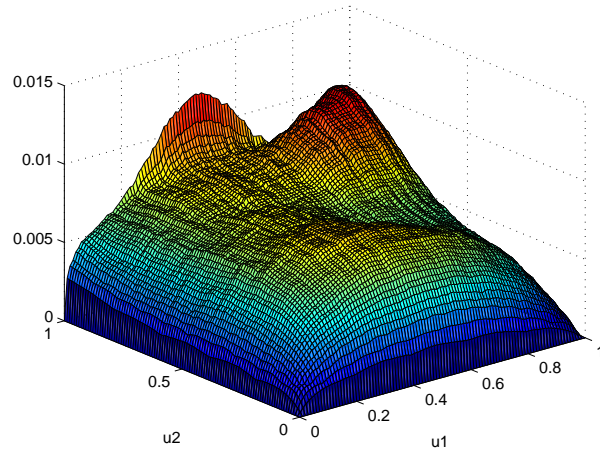


Figure 4.16: Gumbel vs. Normal $n = 500$, $\tau = 0.2$

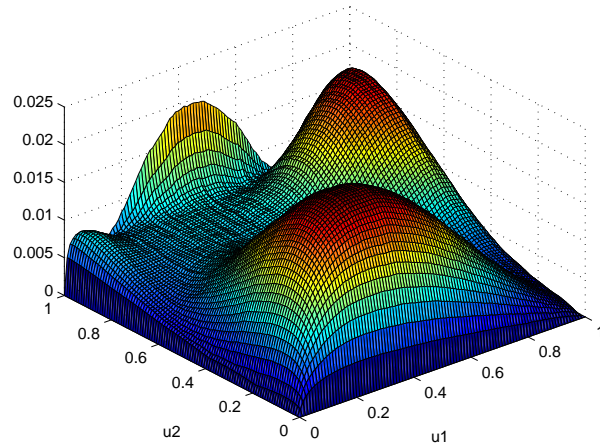


Figure 4.17: Gumbel vs. Clayton $n = 500$, $\tau = 0.2$

4.5 General Application

This section examines an application to financial data that shows how the test proposed in Section 4.3.2 can be utilized in practice. The time series used are daily log returns of MSFT, CERN, and DELL from January 3, 2001 to Dec. 31, 2010.³⁸ Each series has 2514 observations. Table 4.6 shows the summary statistics of the log returns. Significant autocorrelation is present in all three return series as shown by the low p-values in the Ljung-Box Q test, and the data is non-normal according to the Jarque-Bera test.

In order to get rid of the autocorrelation, an AR-GARCH model is used to

Table 4.6: Summary Statistics

	MSFT	CERN	DELL
Mean	0.00008201	0.00025572	-0.000044192
SD	0.0087	0.0124	0.0109
Skewness	0.2126	-3.2071	0.0021
Kurtosis	9.4891	84.5242	7.7156
Ljung-Box - $Q(20)$ (p-value)	< 0.001	0.0126	0.0019
Ljung-Box - $Q^2(20)$ (p-value)	< 0.001	< 0.001	< 0.001
Jarque-Bera (p-value)	< 0.001	< 0.001	< 0.001

pre-filter the data. The resulting series of standardized residuals from the estimation should be *iid*. A logical first step is to start with the simplest specification, which consists of one lag in the conditional mean equation and a GARCH(1,1) with Normal innovations in the conditional variance equation. The AR(1)-GARCH(1,1) model

³⁸The returns are calculated using closing quotes.

specification used is the following:

$$\begin{aligned}x_t &= \phi x_{t-1} + \epsilon_t \\ \sigma_t^2 &= \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta \sigma_{t-1}^2\end{aligned}$$

where $X = (x_1, \dots, x_n)$ is the log return data, $\epsilon_t = \sigma_t z_t$, $z_t \sim N(0, 1)$, and α_0 , α_1 , β , and ϕ are the parameters to be estimated.

Figure 4.18 shows the autocorrelation function for the raw data series (panels a) and b)) and the residual series (panels c) and d)) for the MSFT data. The plots show that the residual and the absolute value of the residual are at least approximately *iid*, while the absolute value of the raw series is clearly not *iid*.³⁹ The Ljung-Box test confirms that the residual and absolute residual are in fact *iid* for all three stocks. As a result, there is no need to experiment with different specifications of the AR-GARCH model.

In practice, the appropriate copula for any application is typically chosen by a combination of graphical and formal methods.⁴⁰ The graphical method used in this section was introduced in Genest and Rivest (1993). The idea is to visually compare the lambda function for several copula families to the empirical lambda function to get a sense of which family provides the closest fit to the data. The lambda function is defined as

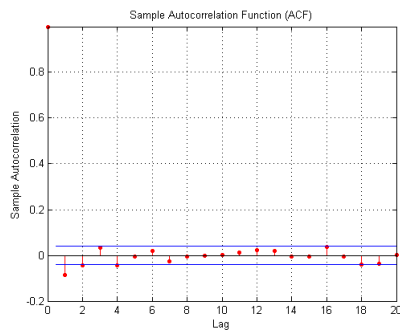
$$\lambda(v) = v - K(v),$$

where $K(v)$ is the copula distribution function given by

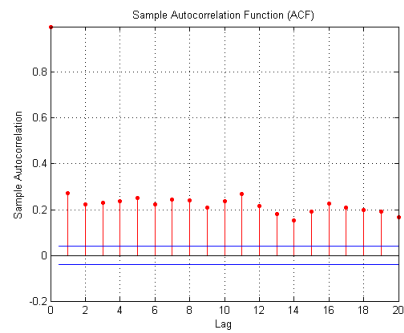
$$K(v) = P(C(u1, u2) \leq v), \quad v \in [0, 1].$$

³⁹In all plots and tests concerning the autocorrelation function, up to 20 lagged values were used.

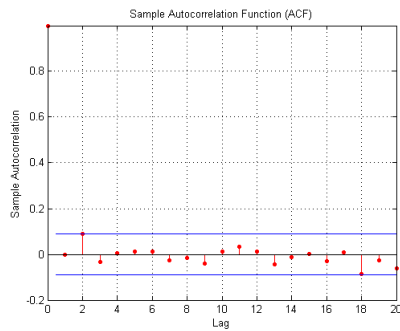
⁴⁰A goodness-of-fit test is considered a formal method.



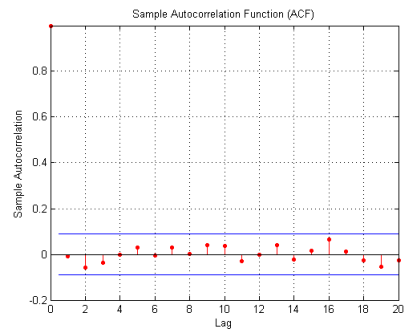
(a) raw data



(b) absolute raw data



(c) residuals



(d) absolute residuals

Figure 4.18: Correlograms for the raw MSFT data and their absolute values as well as for the residuals and absolute residuals. The horizontal lines are the 95% confidence limits.

There is an explicit expression for $K(v)$ for each of the Archimedean copula families (in this paper, Gumbel and Clayton are in the class of Archimedean copulas), but numerical methods are required for elliptical copulas such as the Normal.

To test the ability of the g-o-f test to identify the correct copula family, two cases are examined. First, situations where the graphical tool seems to clearly identify one copula family as the best fit, and second, situations where the graphical tool is inconclusive. In the first case, a powerful g-o-f test should be able to reject copula families providing a poor fit in favour of the closest fitting family. In the second case, rejection in a statistical sense may not occur, but the magnitude of the p-values may still provide guidance regarding which copula provides the best fit to the data.

Case 1

In this experiment, the same n values are used as in the power study; namely, $n \in \{100, 300, 500\}$. The simulation study showed the power converging to 1 for most parameterizations such that $n = 500$, so the cases where $n > 500$ are not as interesting. Since actual data is being used here, there is no guarantee that convergence will occur for the same n value as in the simulation study, but the results of *Case 1* show that the power is indeed very high for $n = 500$, and little of interest is likely to be discovered for larger sample sizes. Conversely, the power tends to be very low for $n < 100$ and these cases are avoided as well. The g-o-f test has difficulty distinguishing one copula family from another, so such parameterizations are less interesting. The procedure used to identify situations where there is one copula that is clearly a better fit than one or more of the other copulas is now described.

1. For each pair of stocks, start at day 1 in the sample (Jan. 3, 2001) and plot all of the lambda functions (empirical, Gumbel, Clayton, and Normal) on the same graph.
2. If one copula is clearly a better fit than any of the others, take note of it and move forward 100 days.⁴¹
3. Repeat steps 1 and 2 until the whole sample has been searched for each of the three pairs of variables ((MSFT, CERN), (MSFT, DELL), and (CERN, DELL)).
4. For each case that was noted in step 2: a) apply the g-o-f test using the poorly fitting copula as the copula under H_0 and the close fitting copula as the ‘best alternative’. A powerful g-o-f test will reject H_0 a high percentage of the time. b) Apply the g-o-f test using the close fitting copula as H_0 and the poorly fitting copula as the ‘best alternative’. A properly functioning g-o-f test will fail to reject H_0 a high percentage of the time.

The application of the g-o-f test is considered a success if the conditions in both step 4)a) and 4)b) are satisfied. For $n = 100$ there were 16 cases where one copula clearly fit the data better than one of the others based on an (admittedly subjective) analysis using the graphical tool. The lambda functions for one such case are shown in Figure 4.19. There were only 4 successes in the 16 cases, reflecting the low power for $n = 100$ in the simulation study. The sample τ ranged from 0.14 to 0.58 over the 16 cases.

⁴¹The reason for moving forward 100 days at a time through a much larger sample (rather than just looking at a single 100 day sample) is simply to obtain a sufficiently large number of cases to analyze.

For $n = 300$ there were 18 cases where one copula clearly fit the data better than one of the others, with 12 successes. The sample τ ranged from 0.22 to 0.54. Finally, for $n = 500$ there were 16 successes in 20 cases, and the sample τ ranged from 0.18 to 0.46. The increased reliability of the test as n increases is expected given the results of the simulation study.

Using Figure 4.19 as a guide, let's walk through an example step-by-step. First, notice that the Gumbel copula fits the data very closely, while the Clayton copula does not. If the copula under H_0 is Clayton, the g-o-f test presented in this paper should reject H_0 . In this particular example, the p-value of the test is 0.002, indicating rejection of H_0 , as expected.⁴² Next, if the copula under H_0 is Gumbel, the test should fail to reject H_0 , which it does in this example, with a p-value of 0.643. Since the conditions from 4)a) and 4)b) are both met, the conclusion is that the g-o-f test confirms the graphical evidence of a strong fit from the Gumbel copula.

An important observation to take away from the exercise performed above is that even if only one of 4)a) and 4)b) are satisfied, information regarding which copula provides the best fit can still be extracted from the p-values. For instance, had the p-value in the first step of the previous example been 0.07 instead of 0.002, the null hypothesis that Clayton fit the data could not have been rejected. However, it was much *closer* to rejection than the null hypothesis that Gumbel fit the data, and as a result, the practitioner is better off choosing the Gumbel to model the data based on the large p-value discrepancy (0.07 vs. 0.643). This observation will come in handy in *Case 2* when the graphical tool gives little guidance in terms of choosing the best fitting copula. In addition, looking back at all of the 'failed' cases in this section, the

⁴²The critical value is 0.05, so H_0 is rejected for any p-value less than 0.05.

g-o-f test still identified the correct copula in every case based on a relative p-value comparison.⁴³

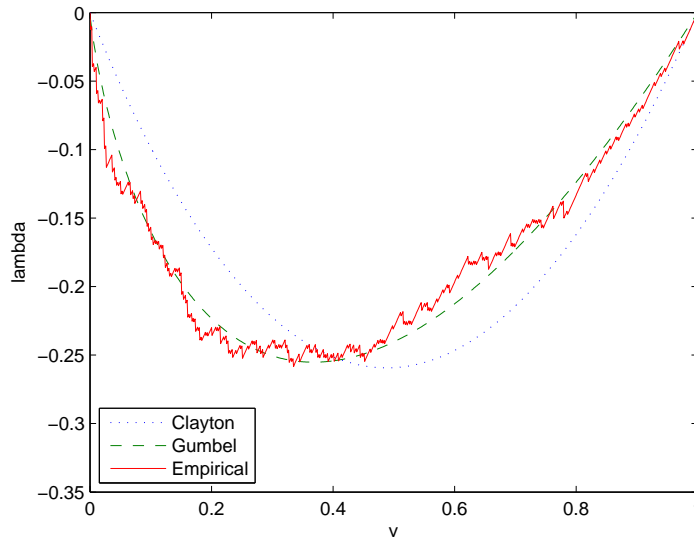


Figure 4.19: Lambda functions ($n = 300$)

Case 2

Case 2 examines situations where the graphical tool does not provide sufficient clarity regarding the best fitting copula. There are many cases where the g-o-f test is

⁴³By ‘correct copula’, it is meant the copula that the lambda plots have identified as the best copula using the author’s subjective judgement. Referring back to Figure 4.19 as a typical example, hopefully it is clear that there shouldn’t be much argument regarding which copula is the best fit in each of the examples in *Case 1* according to the lambda plots, even though there is no objective criterion with which to make the judgement.

unable to reject the null hypothesis; however, as mentioned in *Case 1*, the magnitude of the p-values still gives information regarding goodness-of-fit. From the standpoint of a practitioner, regardless of whether or not the g-o-f test rejects any copula in a statistical sense, a decision still needs to be made regarding which copula to use. Step 4 from the procedure described earlier can still be applied in this case, but instead of worrying about how large the p-values in 4)a) and 4)b) are relative to the critical value, the relevant issue is how large they are relative to each other. An example will be illustrative. Figure 4.20 shows the lambda functions for an example where the best fitting copula cannot be determined visually. To simplify matters, assume the copula universe only consists of Gumbel and Clayton. First, let Clayton be the copula under H_0 , and Gumbel the ‘best alternative’. In this example, the p-value from the g-o-f test is 0.098, meaning Clayton cannot be rejected at the 0.05 significance level. Next, let Gumbel be the copula under H_0 and Clayton the ‘best alternative’. The p-value is 0.304, so Gumbel cannot be rejected; however, Clayton was clearly *closer* to being rejected than was Gumbel, thus, the practitioner should choose Gumbel as the copula that provides the best fit to the data. In the data sample used in this study, the p-value disparity is not always as convincing as in the previous example, but often times it is, even for smaller sample sizes ($n = 100$). This highlights a common problem when applying this test to real data; namely, joint acceptance of the models under both H_0 and H_1 . This is a common problem when dealing with non-nested hypotheses. According to Gouriéroux and Monfort (1994), joint acceptance is often an indication that the data is not sufficiently informative, which certainly appears to be the case for small sample sizes.

It is interesting to look at p-value disparity across copula pairings. Graphically,

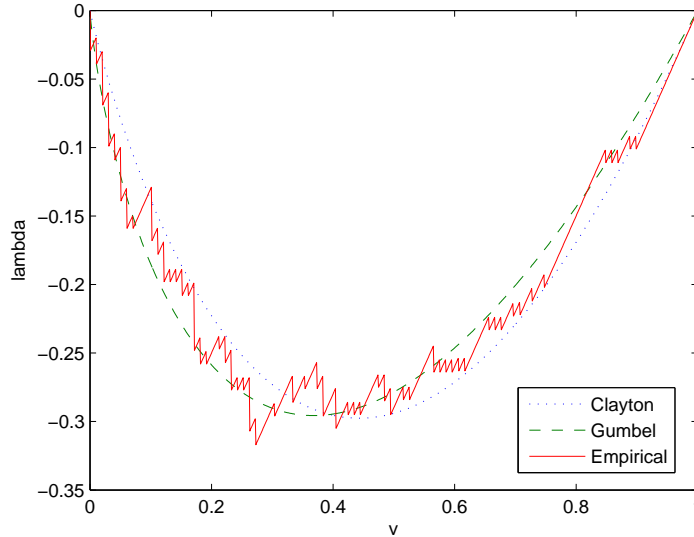


Figure 4.20: Lambda functions ($n = 100$)

the Normal lambda function is typically bounded by the Gumbel and Clayton, as shown in Figure 4.21.⁴⁴ As a result, it is more difficult to distinguish between a Normal copula and a Gumbel or a Clayton, than it is to distinguish between a Gumbel and a Clayton. This is reflected in the p-value disparity from the g-o-f test. The disparity is greater for the (Gumbel, Clayton) case than for either of the (Normal, Clayton) or (Normal, Gumbel) cases. Still, in the majority of (Normal, Gumbel) and (Normal, Clayton) cases there is a substantial difference in p-values and a clear victor emerges. The g-o-f test presented in this paper appears to be a useful tool in determining which copula family to use to model financial data.

⁴⁴Figure 4.21 is a fairly representative picture for the data set used in this study. For a handful of sub-samples the Normal lambda function will fall below both functions far large parts of the domain, but typically it is bounded throughout the majority of the domain as in Figure 4.21.

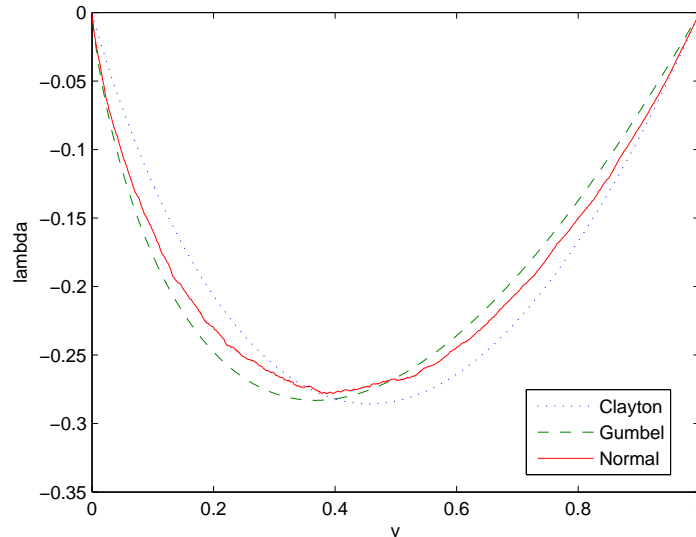


Figure 4.21: Lambda functions

Extending the problem to the case of a universe with n copulas adds some complexity. The first step would be to use the graphical tool to identify as many copula families as possible that provide a poor fit. Then, using the remaining m ($\leq n$) copulas, perform $\binom{m}{2}$ pairwise g-o-f tests and determine a winner in each pairwise comparison. The copula with the best win/loss record can be considered the best fit.

A potential problem in performing the analysis in the manner described in the four-step procedure above is that some copula pairings had significantly more power than others in the simulation study. For instance, the pairings in Table 4.7 all tended to have relatively higher power than those in Table 4.8. A problem could occur if the following condition is met.

Table 4.7: Pairings with High Power

Copula Under H_0	Best Alternative
Clayton	Gumbel
Clayton	Normal
Gumbel	Clayton
Normal	Clayton

Table 4.8: Pairings with Low Power

Copula Under H_0	Best Alternative
Gumbel	Normal
Normal	Gumbel

Condition 4.1

When copula A_1 is under the null and copula A_2 is the best alternative, the power (from the simulation study) is high, but when copula A_2 is under the null and copula A_1 is the best alternative, the power is relatively lower.

If Condition 4.1 is met, then the p-value comparison is an invalid method of choosing the best copula since the p-values will be systemically different. Fortunately, Tables 4.3-4.5 seem to indicate that Condition 4.1 does not hold. As a result, this issue is not problematic in the present study. As a final caution, keep in mind that if the universe of copulas was expanded, then Condition 4.1 should be re-tested to ensure there is no issue with p-value comparability.

4.6 Application to Pair-copula Decomposition

One of the weaknesses of the pair-copula decomposition method is the non-existence of sufficiently powerful g-o-f tests. The test proposed in this paper goes part of the way to correcting this problem, but further improvements are necessary in order for g-o-f tests to be considered reliable stand-alone tools in copula selection. This section shows how the test proposed in this paper can be applied to copula selection within the context of the pair-copula decomposition method. First, the reader will likely find a brief introduction to the decomposition method useful.

The motivation for the pair-copula decomposition method comes from the limited selection of multivariate (as opposed to bivariate) copulas. The decomposition expresses a multivariate copula as a function of several bivariate copulas, thus allowing for a greater degree of flexibility in the modelling process. Let $\mathbf{X} = (X_1, \dots, X_n)$ be a vector of random variables with the joint density function $f(x_1, \dots, x_n)$. The density can be decomposed as

$$f(x_1, \dots, x_n) = f_n(x_n) \cdot f(x_{n-1}|x_n) \cdot f(x_{n-2}|x_{n-1}, x_n) \cdots f(x_1|x_2, \dots, x_n).$$

Introducing copulas into the mix, the density can be re-written in terms of copula density c using the chain rule as follows:

$$f(x_1, \dots, x_n) = c_{1\dots n}(F_1(x_1), \dots, F_n(x_n)) \cdot f_1(x_1) \cdots f_n(x_n),$$

where $F_1(x_1), \dots, F_n(x_n)$ are marginals with multivariate distribution function F such that $F(x_1, \dots, x_n) = C(F_1(x_1), \dots, F_n(x_n))$. In the bivariate case, the density can be expressed as

$$f(x_1, x_2) = c_{12}(F_1(x_1), F_2(x_2)) \cdot f_1(x_1) \cdot f_2(x_2),$$

where c_{12} is the bivariate copula density for the pair of transformed variables $F_1(x_1)$ and $F_2(x_2)$. A conditional density can be written as

$$f(x_1|x_2) = c_{12}(F_1(x_1), F_2(x_2)) \cdot f_1(x_1). \quad (4.3)$$

Extending to the trivariate case yields

$$\begin{aligned} f(x_1|x_2, x_3) &= c_{12|3}(F(x_1|x_3), F(x_2|x_3)) \cdot f(x_1|x_3) \\ &= c_{13|2}(F(x_1|x_2), F(x_3|x_2)) \cdot f(x_1|x_2). \end{aligned} \quad (4.4)$$

Substituting (4.3) into (4.4) yields

$$f(x_1|x_2, x_3) = c_{13|2}(F(x_1|x_2), F(x_3|x_2)) \cdot c_{12}(F_1(x_1), F_2(x_2)) \cdot f_1(x_1).$$

Now, it is clear that a full trivariate decomposition can be expressed as

$$\begin{aligned} f(x_1, x_2, x_3) &= f_1(x_1) \cdot f_2(x_2) \cdot f_3(x_3) \\ &\quad \cdot c_{12}(F_1(x_1), F_2(x_2)) \cdot c_{23}(F_2(x_2), F_3(x_3)) \\ &\quad \cdot c_{13|2}(F(x_1|x_2), F(x_3|x_2)). \end{aligned} \quad (4.5)$$

However, equation (4.4) would seem to indicate that the decomposition can be done in several different, although mathematically equivalent, ways. For instance, instead of using variable pairs (1, 2) and (2, 3) in equation (4.5), the pairs (1, 2) and (1, 3) or (1, 3) and (2, 3) could have been used alternatively, with the appropriate conditional

copula in each case. The strategy behind the choice regarding which specific decomposition to use is discussed later in this section. For more than three variables, the problem becomes more complicated because the number of different decompositions increases quickly with the dimension. To simplify the treatment in this section, the interested reader is referred to the works of Bedford and Cooke (2001, 2002), Kurowicka and Cooke (2004), and Aas et al. (2009) for in-depth analysis of higher dimension decompositions. Joe (1996) shows that, for every j , the conditional marginals in equation (4.5) can be written as

$$F(x|\mathbf{v}) = \frac{\partial C_{x,v_j|\mathbf{v}_{-j}}(F(x|\mathbf{v}_{-j}), F(v_j|\mathbf{v}_{-j}))}{\partial F(v_j|\mathbf{v}_{-j})},$$

where \mathbf{v} is the set of conditioning variables. When v is univariate this is simply $F(x|v) = \frac{\partial C_{xv}(F(x), F(v))}{\partial F(v)}$. Aas et al. (2009) provide explicit expressions for the conditional marginal distributions of several copula families.

Data

The data set used in this section is daily log returns of the DEX Universe Bond Index, Barrick Gold Corporation, and the S&P/TSX Index from Oct. 1, 2007 to Oct. 1, 2009, which gives 500 observations. The summary statistics for the data set are shown in Table 4.9.

As was the case with the previous data set, the data used in this section also exhibits serial correlation. Once again, an AR-GARCH is used to model the serial correlation in the conditional mean and conditional variance. As before, it is logical to

Table 4.9: Summary Statistics

	DEX	Barrick Gold	S&P/TSX
Mean	0.00010802	0.000181	-0.00018161
SD	0.0014	0.0177	0.009
Skewness	-0.2029	0.5231	-0.4701
Kurtosis	4.4776	9.9354	6.9526
Ljung-Box - $Q(20)$ (p-value)	0.1248	< 0.001	< 0.001
Ljung-Box - $Q^2(20)$ (p-value)	0.0015	< 0.001	< 0.001
Jarque-Bera (p-value)	< 0.001	< 0.001	< 0.001

start with the simplest specification, which consists of one lag in the conditional mean equation and a GARCH(1,1) with Normal innovations in the conditional variance equation. The AR(1)-GARCH(1,1) model specification used is the same as before:

$$x_t = \phi x_{t-1} + \epsilon_t$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta \sigma_{t-1}^2,$$

where $X = (x_1, \dots, x_n)$ is the log return data, $\epsilon_t = \sigma_t z_t$, $z_t \sim N(0, 1)$, and α_0 , α_1 , β , and ϕ are the parameters to be estimated.

Figure 4.22 shows the autocorrelation function for the raw data series (panels a) and b)) and the residual series (panels c) and d)) for the Barrick Gold data. The plots show that the residual and the absolute value of the residual are at least approximately *iid*, while the raw series and the absolute value of the raw series are clearly not *iid*. The Ljung-Box test confirms that the residual and absolute residual do not exhibit serial correlation for all three assets. As a result, there is no need to experiment with different specifications of the AR-GARCH model, and analysis can proceed with the standardized residuals.

As discussed in Aas et al. (2009), there are three central components to model

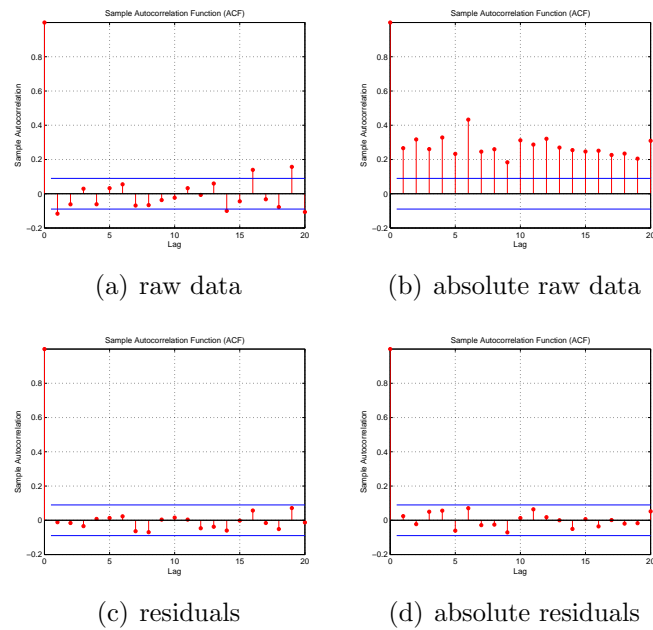


Figure 4.22: Correlograms for the raw Barrick Gold data and their absolute values as well as for the residuals and absolute residuals. The horizontal lines are the 95% confidence limits.

selection. First, one needs to determine the best way to perform the decomposition. Once the specific form of the decomposition has been defined, a copula family needs to be chosen for each bivariate copula density in the decomposition. Finally, the parameters of the chosen copulas need to be estimated.

Pair-copula Decomposition Selection

In the current problem with three variables, two of the three possible *pairs* of variables need to be modelled (see equation (4.5)).⁴⁵ Remember, one benefit of modelling data via copula functions is the flexibility in modelling the dependence structure between variables. As a result, it is logical to choose to model the pairs which exhibit the greatest degree of dependence. This only requires calculating Kendall's τ for each pair. The two pairs with the largest absolute value of Kendall's τ are chosen to be modelled. In each subset of the data set that is used in this study, Kendall's (absolute) τ is highest for (DEX, S&P/TSX) and (Barrick, S&P/TSX).⁴⁶ Then, the decomposition is

$$\begin{aligned}
 f(x_1, x_2, x_3) = & f_1(x_1) \cdot f_2(x_2) \cdot f_3(x_3) \\
 & \cdot c_{13}(F_1(x_1), F_3(x_3)) \cdot c_{23}(F_2(x_2), F_3(x_3)) \\
 & \cdot c_{12|3}(F(x_1|x_3), F(x_2|x_3)), \tag{4.6}
 \end{aligned}$$

where '1' refers to DEX, '2' refers to Barrick, and '3' refers to S&P/TSX. As mentioned in Aas et al. (2009), when the marginals are unknown (which is typically the case in practice), the normalized ranks of the data are used. The transformed data are approximately uniformly distributed, so in equation (4.6), $F_i(x_i) = x_i$ and $f_i(x_i) = 1$

⁴⁵To be clear, the three possible pairs are (DEX, Barrick), (DEX, S&P/TSX), and (Barrick, S&P/TSX).

⁴⁶For example, for (DEX, Barrick) $\tau = 0.0463$, for (DEX, S&P/TSX) $\tau = -0.1812$, and for (Barrick, S&P/TSX) $\tau = 0.3001$ for the full sample.

for $i = \{1, 2, 3\}$. Then, the decomposition can be written as

$$f(x_1, x_2, x_3) = c_{13}(x_1, x_3) \cdot c_{23}(x_2, x_3) \cdot c_{12|3}(F(x_1|x_3), F(x_2|x_3)) \quad (4.7)$$

Copula Selection and Parameter Estimation

The g-o-f test proposed in this paper comes into play in copula selection. The method of selecting appropriate copula functions that was described and utilized in Section 4.5 is used in the same way here. Matters are complicated slightly by the need to choose a copula for the conditional copula density $c_{12|3}(F(x_1|x_3), F(x_2|x_3))$. The problem is that the conditional marginal densities $F(x_1|x_3)$ and $F(x_2|x_3)$ depend on the choice of copula family for the two unconditional copula densities $c_{13}(x_1, x_3)$ and $c_{23}(x_2, x_3)$.^{47,48} The solution is to split the problem of copula selection and parameter estimation into four steps. In step 1, copula families are chosen for the two unconditional copula densities based on the ‘relative p-value’ methodology used in Section 4.5. In step 2, the parameter in each copula needs to be estimated.⁴⁹ Typically this is done via maximum likelihood estimation (MLE) or inversion of Kendall’s τ . To keep in line with the methodology presented in Aas et al. (2009), as well as to illustrate an alternative method of copula parameter estimation in practice, MLE will

⁴⁷To see this, remember that $F(x_1|x_3) = \frac{\partial C_{13}(x_1, x_3)}{\partial x_3}$.

⁴⁸If the unconditional densities are measured with error, this error will be passed on to the conditional density in the next step. An n -dimensional problem has $n - 1$ such steps, so the problem of compounding errors could be significant, particularly in high dimensions.

⁴⁹The copula families used in this paper only have one parameter; however, in general, copulas may have multiple parameters.

be used in this application. The log likelihood equation would look like

$$\sum_{t=1}^T \log(c_{13}(x_{1,t}, x_{3,t}, \Theta_{11})) + \log(c_{23}(x_{2,t}, x_{3,t}, \Theta_{12})).$$

The parameters $(\hat{\Theta}_{11}, \hat{\Theta}_{12})$ that maximize the log likelihood are the maximum likelihood estimates. Then, in step 3, $F(x_1|x_3)$ and $F(x_2|x_3)$ can be calculated using the copula families and corresponding MLE estimates from steps 1 and 2, respectively, and the copula for $c_{12|3}$ can be chosen (using the g-o-f test once again). Finally, in step 4, the parameters of the full log likelihood equation

$$\begin{aligned} \sum_{t=1}^T \log(c_{13}(x_{1,t}, x_{3,t}, \Theta_{11})) + \log(c_{23}(x_{2,t}, x_{3,t}, \Theta_{12})) \\ + \log(c_{12|3}(F(x_{1,t}|x_{3,t}), F(x_{2,t}|x_{3,t}), \Theta_{21})) \end{aligned}$$

are estimated using MLE.

Next, the sample is broken up into two sub-samples; Oct. 1, 2007 to Sept. 12, 2008 (pre-crash), and Sept. 15, 2008 to Oct. 1, 2009 (post-crash).⁵⁰ First, the full sample is analyzed. The first step is to choose the specific form of the decomposition. It was mentioned earlier that for both the full sample and both sub-samples the two most important variable pairs are (DEX, S&P/TSX) and (Barrick, S&P/TSX), so this step is already complete and the form of the decomposition was given in equation (4.7). Next, copula families need to be chosen for the two unconditional copula densities. According to the g-o-f test, the best fit for c_{13} is the Normal copula. The lambda functions can be looked to for confirmation. Figure 4.23 shows that the Normal copula is clearly the best fit for the data. The reader should note that, by definition, the

⁵⁰North American financial markets suffered significant losses on Sept. 15, 2008 in response to a news release regarding Lehman Brothers precarious financial predicament.

Gumbel and Clayton copulas are unable to represent a negative dependence structure. Since the Kendall's τ value is negative for the variable pair (DEX, S&P/TSX), the Clayton and Gumbel copulas would not be expected to outperform the Normal.⁵¹ The Clayton copula is deemed the best fit for c_{23} . The lambda plot (Figure 4.24) does not paint a clear picture for c_{23} , so the g-o-f test is quite useful in this instance.

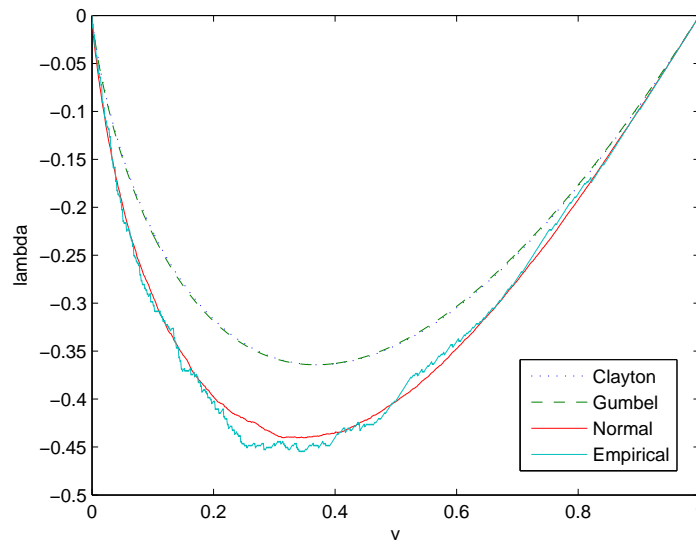


Figure 4.23: Lambda functions

The conditional copula $c_{12|3}$ depends on the copula family and corresponding parameter values from the previous step, i.e. $c_{13}(\Theta_{11})$ and $c_{23}(\Theta_{12})$, so the parameters $(\Theta_{11}, \Theta_{12})$ are estimated using MLE. Now, $c_{12|3}(\Theta_{21})$ is found to be best approximated by a Clayton copula. Now that all three copula families have been chosen, the

⁵¹Rotated Gumbel and Clayton copulas (see Fantazzini (2004)) could have been added to the universe to handle such cases with negative correlation.

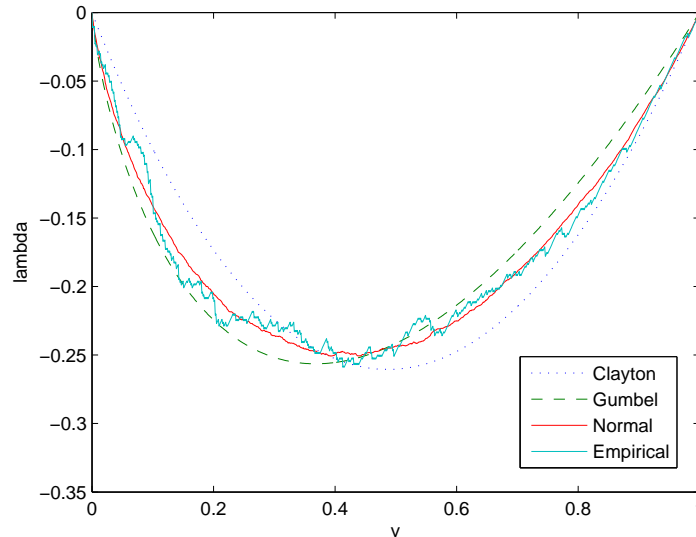


Figure 4.24: Lambda functions

parameters $(\Theta_{11}, \Theta_{12}, \Theta_{21})$ of the entire decomposition are estimated via MLE and the corresponding likelihood value is stored. For the full sample, the likelihood value is 79.7. Now, if the procedure is repeated with the second best copula (according to the g-o-f test) used in each case, the likelihood is 58.0.⁵²

For the pre-crash sub-sample, the copula families chosen are Normal (c_{13}), Gumbel (c_{23}), and Clayton ($c_{12|3}$). The likelihood value is 27.0. For the model using second best copulas, the likelihood value is 16.5 and the chosen copulas are Clayton (c_{13}), Normal (c_{23}), and Normal ($c_{12|3}$). Finally, for the post-crash sub-sample, the chosen copulas are Normal (c_{13}), Clayton (c_{23}), and Normal ($c_{12|3}$). Interestingly, the copulas used to model c_{23} and $c_{12|3}$ change going from the pre-crash data to post-crash data. The likelihood value is 65.6 compared to a value of 52.7 for the model composed of

⁵²Higher likelihood values are representative of better fitting models.

second best copulas. Now that the details of the decomposition have been decided upon, it can be applied to various fields in quantitative finance, such as asset pricing, asset allocation, and Value-at-Risk (VaR) forecasting. VaR forecasting will be the topic of the following section.

4.7 Value-at-Risk Application

This section implements the methodology from the two previous sections in a VaR forecasting framework. The same assets that were used in Section 4.6 are used here. In this section, the data set runs from Sept. 29, 2005 to April 20, 2009. The first 500 days (Sept. 29, 2005 to Dec. 24, 2007) are used as an estimation period for both the time series parameters and the g-o-f test.⁵³ The final 300 days (Dec. 27, 2007 to April 20, 2009) are used to forecast the portfolio VaR. The portfolio is equally-weighted amongst the three assets (DEX, Barrick, and TSX). In general, a more sophisticated method of choosing portfolio weights could be implemented, but such an approach is ignored at the present for simplicity of exposition. The following procedure, similar to the one used in Aas and Berg (2011), is used to obtain one-day VaR forecasts.

1. Obtain the decomposition using the methodology from Section 4.6. This is repeated every 50 days throughout the forecast sample, so that the decomposition does not fall victim to the problem of stale data.⁵⁴ A moving-window approach

⁵³The data was filtered by an AR-GARCH as in the previous section, with the standardized residuals displaying the desired characteristic of no serial correlation.

⁵⁴The value '50' was chosen after taking into consideration the standard efficiency vs. accuracy tradeoff.

is implemented so that the size of the estimation sample is always 500.

2. Compute mean and standard deviation forecasts, $\hat{\mu}_{i,t}$ and $\hat{\sigma}_{i,t}$, for each asset $i = 1, 2, 3$, given information up to time $t - 1$.
3. For each $k = 1, \dots, 1000$:
 - Simulate values u_1, u_2, u_3 from the decomposition in step 1.⁵⁵
 - Convert u_1, u_2, u_3 to $N(0,1)$ variables z_1, z_2, z_3 via the inverse Normal CDF, i.e., $z_i = \Phi^{-1}(u_i)$.
 - Calculate the log-return $r_{i,t} = \hat{\mu}_{i,t} + \hat{\sigma}_{i,t}z_i$ for each asset $i = 1, 2, 3$.
 - Calculate the portfolio log-return $r_{p,t} = \frac{1}{3} \sum_{i=1}^3 r_{i,t}$.
4. For significance levels $q \in \{0.005, 0.01, 0.05\}$:
 - Compute the one-day VaR_t^q as the $q\%$ quantile of the distribution of $r_{p,t}$.
 - If $r_{p,t} < VaR_t^q$, a violation has occurred.

In step 1 of the above procedure, it turns out that the copula families chosen to model the data and the data pairs to be modelled in the initial step of the decomposition remain constant throughout the forecast sample. As in Section 4.6, (DEX, S&P/TSX) and (Barrick, S&P/TSX) are the variable pairs to be modelled in the initial step of the decomposition since they have the highest absolute τ values. The Normal copula is chosen by the g-o-f test to model both data pairs. For the conditional copula in the second step of the decomposition, $c_{12|3}$, the Gumbel copula is

⁵⁵This is done using algorithm 2 in Aas et al. (2009).

chosen. On the other hand, the parameter estimates do change when the decomposition is re-estimated every 50 days. The values are given in Table 4.10.⁵⁶

Table 4.11 shows how the number of violations yielded by the pair-copula decom-

Table 4.10: Copula Parameter Values

	θ_{13}	θ_{23}	$\theta_{12 3}$
Days 1-50	-0.1097	0.6613	1.087
Days 51-100	-0.1719	0.6964	1.0811
Days 101-150	-0.1564	0.6374	1.0695
Days 151-200	-0.1874	0.6252	1.087
Days 201-250	-0.2181	0.5878	1.0753
Days 251-300	-0.2028	0.575	1.0582

position VaR model compares to the expected number of violations. There are several tests that can be used to determine the accuracy of a VaR model (see Christoffersen (1998)). The standard test to determine whether there is a significant difference between the number of violations generated by the VaR model and the expected number of violations is a likelihood ratio (LR) test.⁵⁷ The null hypothesis is that the expected proportion of violations is equal to α . Under the null, the LR test statistic takes the form

$$LR = 2\log\left(\left(\frac{x}{N}\right)^x \left(1 - \frac{x}{N}\right)^{N-x}\right) - 2\log(\alpha^x(1 - \alpha^{N-x})),$$

where x is the number of violations and N is the number of periods in the forecasting sample. LR is asymptotically $\chi^2(1)$.⁵⁸ The VaR model based on the pair-copula

⁵⁶The parameter for the Normal copula is defined on $[-1, 1]$, and the parameter for the Gumbel copula is defined on $[1, \infty)$.

⁵⁷This test is used in many VaR-based studies, for example, Kupiec (1995) and Christoffersen (1998).

⁵⁸This test is often referred to as a test of unconditional coverage.

decomposition does a good job of forecasting based on this test of unconditional coverage, as shown by Table 4.11. Finally, Figure 4.25 shows the log-returns for the portfolio over the forecasting sample along with the VaR forecasts at each of the three significance levels. The top line is the 95% VaR, the middle line is 99%, and the bottom line is 99.5%. The arrows identify where the two 99.5% violations took place. The simple example presented in this section has shown how the new g-o-f test developed in this paper and the pair-copula decomposition method can be applied in practice in a quantitative finance setting.

Table 4.11: Number of Violations

	VaR model	Expected
95%	10*	15
99%	5*	3
99.5%	2*	1.5

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

4.8 Conclusion

A new copula-based goodness-of-fit test is proposed. The new test incorporates a weighting function into one of the Cramer-von Mises type tests based on Rosenblatt’s Transformation that was introduced in Genest et al. (2009). A connection between the copula-based independence test literature and the copula-based g-o-f literature is clearly explained. The central contribution is to develop the methodology used to find an appropriate weighting function. A simulation study shows that the new test has more power than the benchmark (the original unweighted test from Genest et

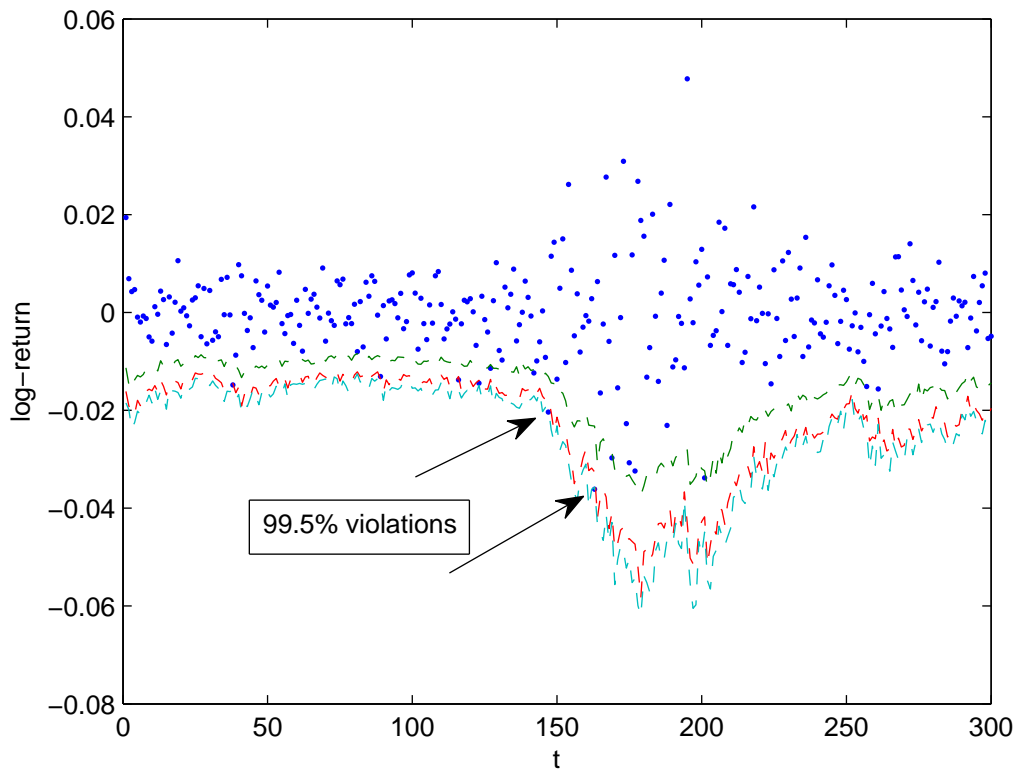


Figure 4.25: VaR forecasts along with log-returns from portfolio

al. (2009)). However, the improvement is marginal and the original problem of low power in small sample sizes still remains. Further significant improvements need to be made before g-o-f tests can be counted on as a reliable, stand-alone tool. In the meantime, supplementing the analysis with graphical tools such as the method based on the lambda function used in Sections 4.5 and 4.6 is advisable. An application to financial data shows how the test can be used to determine which copula family provides the best fit to a given data set, specifically in cases where graphical tools are inconclusive. Finally, an application to the pair-copula decomposition method shows that the test is able to identify the correct copulas based on likelihood values, and an application to Value-at-Risk forecasting displays how the decomposition method may be used in practice in a quantitative finance setting.

4.9 Bibliography

- Aas, K. and D. Berg (2011), *Dependence Modeling: Vine Copula Handbook*, Singapore: World Scientific Publishing Co.
- Aas, K., Czado, C., Frigessi, A., and H. Bakken (2009), “Pair-Copula Construction of Multiple Dependence”, *Insurance: Mathematics and Economics*, 44, 182-198.
- Bedford, T. and R. M. Cooke (2001), “Probability Density Decomposition for Conditionally Dependent Random Variables Modeled by Vines”, *Annals of Mathematics and Artificial Intelligence*, 32, 245-268.

- Bedford, T. and R. M. Cooke (2002), "Vines - a New Graphical Model for Dependent Random Variables", *Annals of Statistics*, 30, 1031-1068.
- Berg, D. (2009), "Copula Goodness-of-Fit Testing: An Overview and Power Comparison", *The European Journal of Finance*, 15(7-8), 675-701.
- Cech, C. (2006), "Copula-Based Top-Down Approaches in Financial Risk Aggregation", *Working Paper*.
- Cherubini, U., Luciano, E., and W. Vecchiato (2004), *Copula Methods in Finance*, West Sussex: Wiley.
- Christoffersen, P. (1998) "Evaluating Interval Forecasts", *International Economic Review*, 39, 841-862.
- Clayton, D. (1978), "A Model for Association in Bivariate Life Tables and its application in Epidemiological Studies of Familial Tendency in Chronic Disease Incidence", *Biometrika*, 65, 141-151.
- Dimakos, X. K., and K. Aas (2004), "Integrated Risk Modeling", *Statistical Modeling*, 4(4), 265-277.
- Durbin, J. and A. Stuart (1951), "Inversions and Rank Correlations", *Journal of the Royal Statistical Society, Series B*, 13(2), 303-309.
- Fantazzini, D. (2004), "Copula's conditional dependence measures for portfolio management and value at risk", Working paper, University of Konstanz.
- Fermanian, J. (2005), "Goodness of fit tests for copulas", *Journal of Multivariate Analysis*, 95, 1191-1252.
- Genest, C. and B. Rémillard (2008), "Validity of the Parametric Bootstrap for Goodness-of-Fit Testing in Semiparametric Models", *Annales de*

- l'Institut Henri Poincaré - Probabilités et Statistiques*, 44(6), 1096-1127.
- Genest, C., Rémillard, B., Beaudoin, D. (2009), “Goodness-of-Fit Tests for Copulas: A Review and Power Study”, *Insurance: Mathematics and Economics*, 44, 199-213.
- Genest, C. and L.-P. Rivest (1993), “Statistical Inference Procedures for Bivariate Archimedean Copulas”, *Journal of the American Statistical Association*, 88, 1034-1043.
- Gourieroux, C. and A. Monfort (1994), “Testing Nonnested Hypothesis”, In *Handbook of Econometrics*, vol. 4, ed. Engle, R. and D. McFadden. Amsterdam: North-Holland .
- Grundke, P. (2010), “Top-Down Approaches for Integrated Risk Management: How Accurate are They?”, *European Journal of Operational Research*, 203(3), 662-672.
- Gumbel, E. (1960), “Bivariate Exponential Distributions”, *Journal of the American Statistical Association*, 55, 698-707.
- Höfding, W. (1940), “Maßstabinvariante Korrelationstheorie”, *PhD Thesis*.
- Joe, H. (1997), *Multivariate Models and Dependence Concepts*, London: Chapman & Hall.
- Joe, H., and J. Xu (1996), *The Estimation Method of Inference Functions for Margins for Multivariate Models*, Dept. of Statistics University of British Columbia, Tech. Rept. 166.
- Jorion, P. (2007), *Value at Risk: The New Benchmark for Managing Financial Risk*, 3rd Ed. Singapore: McGraw-Hill.
- Kojadinovic, I and M. Holmes (2009), “Tests of Independence Among

- Continuous Random Vectors Based on Cramér-von Mises Functionals of the Empirical Copula Process”, *Journal of Multivariate Analysis*, 100, 1137-1154.
- Kruskal, W. (1958), “Ordinal Measures of Association”, *Journal of the American Statistical Association*, 53(284), 814-861.
- Kupiec, P. (1995), “Techniques for Verifying the Accuracy of Risk Measurement Models”, *Journal of Derivatives*, 2, 173-184.
- Kurowicka, D. and R. M. Cooke (2006), *Uncertainty Analysis with High Dimensional Dependence Modeling*, New York: Wiley.
- Mandelbrot, B. (1963), “Long-Run Linearity, Locally Gaussian Process, H-Spectra and Infinite Variances”, *International Economic Review*, 10(1), 82-111.
- Medovikov, I. (2011), “A Test of Independence in Econometric Models”, *Working Paper*.
- Nelsen, R. (2006), *An Introduction to Copulas*, 2nd Ed. New York: Springer.
- Rosenberg, J. V., and T. Schuermann (2006), “A General Approach to Integrated Risk Management with Skewed, Fat-Tailed Risks”, *Journal of Financial Economics*, 79(3), 569-614.
- Rosenblatt, M. (1952), “Remarks on a Multivariate transformation”, *Annals of Mathematical Statistics*, 23, 470-472.
- Sklar, A. (1959), “Fonctions de Répartition à n dimension et leurs Marges”, *Publications de l’Institut de Statistique de l’Université de Paris*, 8, 229-331.

Ward, L. S., and D. H. Lee (2002), “Practical Application of the Risk-Adjusted Return on Capital Framework”, *Working Paper*.

Chapter 5

Conclusion

This thesis explored various topics relevant to a portfolio manager. The chapter on technical analysis alerted portfolio managers that the reliability of trading signals is dependent on the level of volatility in the asset being traded. This is useful information for both quantitative traders and portfolio managers trying to decide when the best time is to enter/exit a position. Four classes of trading rules were studied in this chapter, including exponential moving average, the KST indicator, and two classes of complex rules that were composed of simple moving average rules. The KST indicator (with overbought/oversold levels used to generate trading signals) was found to be relatively more profitable during periods of medium to high volatility, while exponential moving average rules excelled under low volatility. Moreover, the profitability of trading rules was tested on intradaily equity data. This appears to be the first study of its kind using intradaily data for equities. Unfortunately, a consensus could not be reached in determining whether it is more profitable to trade on intradaily data relative to daily data. Finally, confirming the findings of Hsu and Kuan (2005), the small-cap stocks on the Russell 2000 appear to have inefficient markets, allowing for exploitation by technical trading rules.

The chapter on Value-at-Risk provided a useful risk management tool for portfolio managers. A VaR model previously introduced in KSW (2003) was extended so that it produced accurate forecasts before, during, and after the 2008-09 market crash. New estimation procedures for a key parameter in the model are developed that lead to a flexible model capable of responding quickly to changing market conditions. The new model outperforms a benchmark model that had been successful prior to the crash, as well as the original model from KSW (2003). Risk management is a critical component of portfolio management; an accurate VaR model can help to identify exposures that carry an unacceptable amount of downside risk, potentially allowing the portfolio manager to make changes to the portfolio before disaster strikes. An important feature of the model is its ability to predict the market crash. Portfolio managers utilizing the extended VARLINEX model would not have run into capital availability problems due to market risk as long as they altered either the amount of capital reserves on hand, or the exposures of their portfolio, to keep in line with the VaR forecasts.

The final chapter addresses the issue of bivariate data modelling, which is another relevant aspect of portfolio management. Copula functions provide a flexible method of modelling multivariate data since the specification of the marginals is done separately from the specification of the dependence structure. This is a useful feature in financial data modelling because correct specification of the dependence structure is very important. Many different copula families are available to model the data in question, and a reliable method of choosing the best one is needed. This paper develops a copula-based goodness-of-fit test that has more power than a benchmark which had previously been one of the strongest available in the literature. The new test

developed in this paper implements a weighting function in one of the test statistics proposed in Genest et al. (2009), which was also used as the benchmark. A method for deriving a well-functioning weighting function is explained in detail. An application involving equity data shows that the test is useful in selecting the best fitting copula function in situations where the graphical tool fails to provide any clues. In addition, applications to the pair-copula decomposition method and Value-at-Risk forecasting illustrate how the g-o-f tests can be utilized in practice.

5.1 Bibliography

- Genest, C., Rémillard, B., Beaudoin, D. (2009), “Goodness-of-Fit Tests for Copulas: A Review and Power Study”, *Insurance: Mathematics and Economics*, 44, 199-213.
- Knight, J., S. Satchell, and G. Wang (2003), “Value at Risk Linear Exponent (VARLINEX) Forecasts”, *Quantitative Finance*, 3, 332-344.

Appendix A

List of Stocks

	Stock	Ticker Symbol	Exchange	Russell Index
1	Art Technology Group, Inc.	ARTG	NASDAQ	2000
2	Blue Coat Systems, Inc.	BCSI	NASDAQ	2000
3	Boyd Gaming Corp.	BYD	NYSE	Midcap
4	CNET Networks, Inc.	CNET	NASDAQ	2000
5	Entegris, Inc.	ENTG	NASDAQ	2000
6	The First Marblehead Corporation	FMD	NYSE	2000
7	Goldman Sachs Group, Inc.	GS	NYSE	Top 200
8	Graphtech International Ltd.	GTI	NYSE	Midcap
9	Johnson & Johnson	JNJ	NYSE	Top 200
10	Microsoft Corp.	MSFT	NASDAQ	Top 200
11	Manitowoc Company, Inc.	MTW	NYSE	Midcap
12	Northern Trust Corporation	NTRS	NASDAQ	Midcap
13	Patterson Companies, Inc.	PDCO	NASDAQ	Midcap
14	Pacific Sunwear of California, Inc.	PSUN	NASDAQ	2000
15	RF Micro Devices, Inc.	RFMD	NASDAQ	2000
16	Schering Plough Corp.	SGP	NYSE	Top 200
17	Synopsys, Inc.	SNPS	NASDAQ	Midcap
18	UDR, Inc.	UDR	NYSE	Midcap
19	Exxon Mobil Corp.	XOM	NYSE	Top 200
20	XTO Energy Inc.	XTO	NYSE	Top 200

Appendix B

Legend for Tables

1) KST rules

The KST rules are denoted by $KST - n$, where n is the level of the overbought/oversold line. n takes the values 1, 2, 4, 6, 7, and 10%, thus there are six KST rules. Shorting is allowed for all KST rules.

2) Complex rules

The complex rules are $CF - x$, for complex fractional, and $CV - x$ for complex voting, where $x \in \{NS, S\}$. 'NS' indicates no shorting, and 'S' indicates shorting. Thus, there are two voting and two fractional rules for a total of four complex rules.

3) Moving average rules

An example of the notation for a moving average rule is $2MA60b4 - NS$. The initial '2' indicates that the short MA is calculated using two periods of data. The absence of an '2' (for example, $MA60b4 - NS$) indicates that the short MA is just the price. '60' indicates the number of periods

in the long EMA. In this position, we can have the values 10, 20, 30, 40, 50, and 60. 'b4' indicates that the trading band is 0.4%. In this position we can have b0 (no band), b2 (0.2%), b4, and b6 (0.6%). Finally, 'NS' indicates no shorting. We can have either 'S' (indicating shorting) or 'NS' in this position. Thus, we have $2 * 6 * 4 * 2 = 96$ moving average rules.

Appendix C

DJI, NASDAQ, & S&P/TSX

Results

Varying window lengths

Window length = 10

Table C.1: DJI - Full sample - number of VaR violations (N=1466)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	135	92	71*	73.3
97.5%	87	43*	33*	36.65
99%	53	25	15*	14.66

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table C.2: DJI - Pre-crisis - number of VaR violations (N=968)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	31	65	43*	48.4
97.5%	16*	29*	23*	24.2
99%	12*	18	10*	9.68

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table C.3: DJI - Crisis - number of VaR violations (N=355)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	93	26*	22*	17.75
97.5%	70	13*	9*	8.875
99%	41	7*	5*	3.55

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table C.4: DJI - Post-crisis - number of VaR violations (N=143)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	11*	2	6*	7.15
97.5%	1*	1*	1*	3.575
99%	0*	0*	0*	1.43

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Window length = 25

Table C.5: DJI - Full sample - number of VaR violations (N=1466)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	135	92	87*	73.3
97.5%	87	43*	30*	36.65
99%	53	25	16*	14.66

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table C.6: DJI - Pre-crisis - number of VaR violations (N=968)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	31	65	54*	48.4
97.5%	16*	29*	20*	24.2
99%	12*	18	11*	9.68

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table C.7: DJI - Crisis - number of VaR violations (N=355)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	93	26*	27*	17.75
97.5%	70	13*	9*	8.875
99%	41	7*	5*	3.55

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table C.8: DJI - Post-crisis - number of VaR violations (N=143)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	11*	2	6*	7.15
97.5%	1*	1*	1*	3.575
99%	0*	0*	0*	1.43

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Window length = 75

Table C.9: DJI - Full sample - number of VaR violations (N=1466)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	135	92	82*	73.3
97.5%	87	43*	39*	36.65
99%	53	25	13*	14.66

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table C.10: DJI - Pre-crisis - number of VaR violations (N=968)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	31	65	57*	48.4
97.5%	16*	29*	27*	24.2
99%	12*	18	10*	9.68

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table C.11: DJI - Crisis - number of VaR violations (N=355)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	93	26*	25*	17.75
97.5%	70	13*	10*	8.875
99%	41	7*	3*	3.55

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table C.12: DJI - Post-crisis - number of VaR violations (N=143)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	11*	2	5*	7.15
97.5%	1*	1*	2*	3.575
99%	0*	0*	0*	1.43

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Window length = 100

Table C.13: DJI - Full sample - number of VaR violations (N=1466)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	135	92	87*	73.3
97.5%	87	43*	48*	36.65
99%	53	25	18*	14.66

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table C.14: DJI - Pre-crisis - number of VaR violations (N=968)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	31	65	61	48.4
97.5%	16*	29*	34*	24.2
99%	12*	18	12*	9.68

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table C.15: DJI - Crisis - number of VaR violations (N=355)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	93	26*	21*	17.75
97.5%	70	13*	11*	8.875
99%	41	7*	6*	3.55

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table C.16: DJI - Post-crisis - number of VaR violations (N=143)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	11*	2	5*	7.15
97.5%	1*	1*	3*	3.575
99%	0*	0*	0*	1.43

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Window length = 200

Table C.17: DJI - Full sample - number of VaR violations (N=1466)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	135	92	102	73.3
97.5%	87	43*	49	36.65
99%	53	25	21*	14.66

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table C.18: DJI - Pre-crisis - number of VaR violations (N=968)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	31	65	67	48.4
97.5%	16*	29*	37	24.2
99%	12*	18	17	9.68

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table C.19: DJI - Crisis - number of VaR violations (N=355)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	93	26*	30	17.75
97.5%	70	13*	11*	8.875
99%	41	7*	4*	3.55

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table C.20: DJI - Post-crisis - number of VaR violations (N=143)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	11*	2	5*	7.15
97.5%	1*	1*	1*	3.575
99%	0*	0*	0*	1.43

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

‘a’ related to VIX

10 day moving average

Table C.21: DJI - Full sample - number of VaR violations (N=1466)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	135	92	88*	73.3
97.5%	87	43*	40*	36.65
99%	53	25	14*	14.66

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table C.22: DJI - Pre-crisis - number of VaR violations (N=968)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	31	65	64	48.4
97.5%	16*	29*	33*	24.2
99%	12*	18	12*	9.68

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table C.23: DJI - Crisis - number of VaR violations (N=355)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	93	26*	20*	17.75
97.5%	70	13*	6*	8.875
99%	41	7*	2*	3.55

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table C.24: DJI - Post-crisis - number of VaR violations (N=143)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	11*	2	4*	7.15
97.5%	1*	1*	1*	3.575
99%	0*	0*	0*	1.43

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table C.25: NASDAQ - Full sample - number of VaR violations (N=1466)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	112	91	83*	73.3
97.5%	74	54	49	36.65
99%	46	23	14*	14.66

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table C.26: NASDAQ - Pre-crisis - number of VaR violations (N=968)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	20	64	57*	48.4
97.5%	8	38	32*	24.2
99%	5*	18	10*	9.68

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table C.27: NASDAQ - Crisis - number of VaR violations (N=355)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	82	24*	20*	17.75
97.5%	64	14*	11*	8.875
99%	41	4*	3*	3.55

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table C.28: NASDAQ - Post-crisis - number of VaR violations (N=143)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	10*	3	6*	7.15
97.5%	2*	1*	6*	3.575
99%	0*	1*	1*	1.43

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Varying window lengths

Window length = 10

Table C.29: NASDAQ - Full sample - number of VaR violations (N=1466)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	112	91	58*	73.3
97.5%	74	54	34*	36.65
99%	46	23	22*	14.66

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table C.30: NASDAQ - Pre-crisis - number of VaR violations (N=968)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	20	64	41*	48.4
97.5%	8	38	21*	24.2
99%	5*	18	10*	9.68

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table C.31: NASDAQ - Crisis - number of VaR violations (N=355)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	82	24*	14*	17.75
97.5%	64	14*	10*	8.875
99%	41	4*	9	3.55

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table C.32: NASDAQ - Post-crisis - number of VaR violations (N=143)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	10*	3	3*	7.15
97.5%	2*	1*	3*	3.575
99%	0*	1*	3*	1.43

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Window length = 25

Table C.33: NASDAQ - Full sample - number of VaR violations (N=1466)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	112	91	87*	73.3
97.5%	74	54	38*	36.65
99%	46	23	15*	14.66

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table C.34: NASDAQ - Pre-crisis - number of VaR violations (N=968)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	20	64	60*	48.4
97.5%	8	38	25*	24.2
99%	5*	18	11*	9.68

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table C.35: NASDAQ - Crisis - number of VaR violations (N=355)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	82	24*	20*	17.75
97.5%	64	14*	11*	8.875
99%	41	4*	3*	3.55

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table C.36: NASDAQ - Post-crisis - number of VaR violations (N=143)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	10*	3	6*	7.15
97.5%	2*	1*	2*	3.575
99%	0*	1*	1*	1.43

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Window length = 75

Table C.37: NASDAQ - Full sample - number of VaR violations (N=1466)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	112	91	77*	73.3
97.5%	74	54	37*	36.65
99%	46	23	15*	14.66

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table C.38: NASDAQ - Pre-crisis - number of VaR violations (N=968)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	20	64	51*	48.4
97.5%	8	38	24*	24.2
99%	5*	18	11*	9.68

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table C.39: NASDAQ - Crisis - number of VaR violations (N=355)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	82	24*	21*	17.75
97.5%	64	14*	9*	8.875
99%	41	4*	3*	3.55

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table C.40: NASDAQ - Post-crisis - number of VaR violations (N=143)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	10*	3	5*	7.15
97.5%	2*	1*	4*	3.575
99%	0*	1*	1*	1.43

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Window length = 100

Table C.41: NASDAQ - Full sample - number of VaR violations (N=1466)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	112	91	80*	73.3
97.5%	74	54	40*	36.65
99%	46	23	15*	14.66

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table C.42: NASDAQ - Pre-crisis - number of VaR violations (N=968)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	20	64	54*	48.4
97.5%	8	38	26*	24.2
99%	5*	18	11*	9.68

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table C.43: NASDAQ - Crisis - number of VaR violations (N=355)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	82	24*	21*	17.75
97.5%	64	14*	10*	8.875
99%	41	4*	3*	3.55

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table C.44: NASDAQ - Post-crisis - number of VaR violations (N=143)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	10*	3	5*	7.15
97.5%	2*	1*	4*	3.575
99%	0*	1*	1*	1.43

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Window length = 200

Table C.45: NASDAQ - Full sample - number of VaR violations (N=1466)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	112	91	87*	73.3
97.5%	74	54	45*	36.65
99%	46	23	20*	14.66

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table C.46: NASDAQ - Pre-crisis - number of VaR violations (N=968)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	20	64	61*	48.4
97.5%	8	38	30*	24.2
99%	5*	18	15*	9.68

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table C.47: NASDAQ - Crisis - number of VaR violations (N=355)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	82	24*	22*	17.75
97.5%	64	14*	12*	8.875
99%	41	4*	4*	3.55

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table C.48: NASDAQ - Post-crisis - number of VaR violations (N=143)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	10*	3	4*	7.15
97.5%	2*	1*	3*	3.575
99%	0*	1*	1*	1.43

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

10 day moving average

Table C.49: NASDAQ - Full sample - number of VaR violations (N=1466)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	112	91	92	73.3
97.5%	74	54	49	36.65
99%	46	23	15*	14.66

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table C.50: NASDAQ - Pre-crisis - number of VaR violations (N=968)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	20	64	72	48.4
97.5%	8	38	40	24.2
99%	5*	18	12*	9.68

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table C.51: NASDAQ - Crisis - number of VaR violations (N=355)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	82	24*	17*	17.75
97.5%	64	14*	8*	8.875
99%	41	4*	3*	3.55

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table C.52: NASDAQ - Post-crisis - number of VaR violations (N=143)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	10*	3	3*	7.15
97.5%	2*	1*	1*	3.575
99%	0*	1*	0*	1.43

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

50 day moving average

Table C.53: NASDAQ - Full sample - number of VaR violations (N=1466)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	112	91	91	73.3
97.5%	74	54	52	36.65
99%	46	23	18*	14.66

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table C.54: NASDAQ - Pre-crisis - number of VaR violations (N=968)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	20	64	70	48.4
97.5%	8	38	43	24.2
99%	5*	18	15*	9.68

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table C.55: NASDAQ - Crisis - number of VaR violations (N=355)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	82	24*	18*	17.75
97.5%	64	14*	8*	8.875
99%	41	4*	3*	3.55

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table C.56: NASDAQ - Post-crisis - number of VaR violations (N=143)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	10*	3	3*	7.15
97.5%	2*	1*	1*	3.575
99%	0*	1*	0*	1.43

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table C.57: NASDAQ - Summary Statistics

	Full Sample	Pre-Crash	Crash	Post-Crash
Mean	0.00017912	0.00042717	-0.0014	0.0025
SD	0.0152	0.0094	0.0252	0.014
Skewness	0.0764	-0.1498	0.2513	-0.1839
Kurtosis	10.6474	3.4222	5.5939	3.7005
Min	-0.0914	-0.0386	-0.0914	-0.0388
Max	0.1181	0.0302	0.1181	0.0389
Jarque-Bera (p-value)	< 0.001	< 0.001	< 0.001	0.1075

Table C.58: S&P/TSX Composite - Full sample - number of VaR violations (N=1350)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	155	95	78*	67.5
97.5%	97	46	41*	33.75
99%	43	26	15*	13.5

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table C.59: S&P/TSX Composite - Pre-crisis - number of VaR violations (N=859)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	60	56*	57	42.95
97.5%	32	30*	32	21.475
99%	14*	18	16	8.59

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table C.60: S&P/TSX Composite - Crisis - number of VaR violations (N=350)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	83	35	19*	17.5
97.5%	62	15*	8*	8.75
99%	29	8	0	3.5

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table C.61: S&P/TSX Composite - Post-crisis - number of VaR violations (N=141)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	12*	4*	2	7.05
97.5%	3*	2*	1*	3.525
99%	0*	0*	0*	1.41

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Varying window lengths

Window length = 10

Table C.62: S&P/TSX Composite - Full sample - number of VaR violations (N=1350)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	155	95	54*	67.5
97.5%	97	46	31*	33.75
99%	43	26	12*	13.5

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table C.63: S&P/TSX Composite - Pre-crisis - number of VaR violations (N=859)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	60	56*	32*	42.95
97.5%	32	30*	20*	21.475
99%	14*	18	10*	8.59

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table C.64: S&P/TSX Composite - Crisis - number of VaR violations (N=350)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	83	35	18*	17.5
97.5%	62	15*	9*	8.75
99%	29	8	2*	3.5

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table C.65: S&P/TSX Composite - Post-crisis - number of VaR violations (N=141)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	12*	4*	4*	7.05
97.5%	3*	2*	2*	3.525
99%	0*	0*	0*	1.41

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Window length = 25

Table C.66: S&P/TSX Composite - Full sample - number of VaR violations (N=1350)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	155	95	84	67.5
97.5%	97	46	37*	33.75
99%	43	26	15*	13.5

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table C.67: S&P/TSX Composite - Pre-crisis - number of VaR violations (N=859)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	60	56*	52*	42.95
97.5%	32	30*	25*	21.475
99%	14*	18	12*	8.59

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table C.68: S&P/TSX Composite - Crisis - number of VaR violations (N=350)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	83	35	26*	17.5
97.5%	62	15*	10*	8.75
99%	29	8	3*	3.5

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table C.69: S&P/TSX Composite - Post-crisis - number of VaR violations (N=141)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	12*	4*	6*	7.05
97.5%	3*	2*	2*	3.525
99%	0*	0*	0*	1.41

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

10 day moving average

Table C.70: S&P/TSX Composite - Full sample - number of VaR violations (N=1350)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	155	95	90	67.5
97.5%	97	46	46	33.75
99%	43	26	23	13.5

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table C.71: S&P/TSX Composite - Pre-crisis - number of VaR violations (N=859)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	60	56*	62	42.95
97.5%	32	30*	35	21.475
99%	14*	18	21	8.59

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table C.72: S&P/TSX Composite - Crisis - number of VaR violations (N=350)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	83	35	26*	17.5
97.5%	62	15*	10*	8.75
99%	29	8	2*	3.5

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table C.73: S&P/TSX Composite - Post-crisis - number of VaR violations (N=141)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	12*	4*	2	7.05
97.5%	3*	2*	1*	3.525
99%	0*	0*	0*	1.41

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

50 day moving average

Table C.74: S&P/TSX Composite - Full sample - number of VaR violations (N=1350)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	155	95	93	67.5
97.5%	97	46	45*	33.75
99%	43	26	22	13.5

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table C.75: S&P/TSX Composite - Pre-crisis - number of VaR violations (N=859)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	60	56*	64	42.95
97.5%	32	30*	34	21.475
99%	14*	18	20	8.59

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table C.76: S&P/TSX Composite - Crisis - number of VaR violations (N=350)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	83	35	27	17.5
97.5%	62	15*	10*	8.75
99%	29	8	2*	3.5

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table C.77: S&P/TSX Composite - Post-crisis - number of VaR violations (N=141)

	EVT	Original VARLINEX	New VARLINEX	Target
95%	12*	4*	2*	7.05
97.5%	3*	2*	1*	3.525
99%	0*	0*	0*	1.41

‘*’ indicates the model passed the LR test (i.e. H_0 was not rejected) at a 5% significance level.

Table C.78: S&P/TSX - Summary Statistics

	Full Sample	Pre-Crash	Crash	Post-Crash
Mean	0.00030455	0.00062662	-0.0012	0.0021
SD	0.0142	0.0078	0.0231	0.015
Skewness	-0.4702	-0.4952	-0.1785	-0.448
Kurtosis	11.6228	4.1401	5.9415	3.1879
Min	-0.0932	-0.0352	-0.0932	-0.0441
Max	0.0982	0.024	0.0982	0.0393
Jarque-Bera (p-value)	< 0.001	< 0.001	< 0.001	0.0631

Appendix D

Dependence Functions

(*DEPFUNC*)

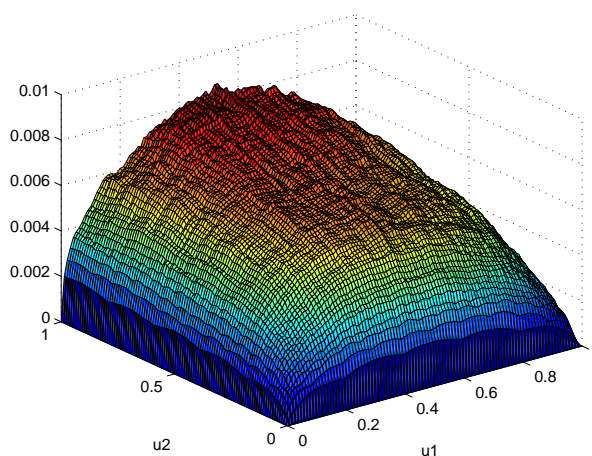


Figure D.1: Gumbel vs. Gumbel $n = 500$, $\tau = 0.2$

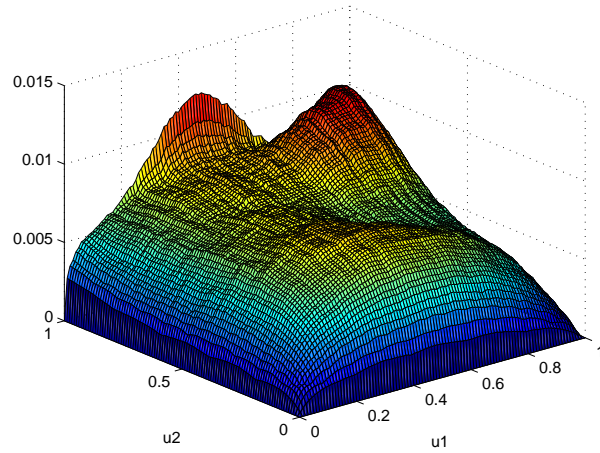


Figure D.2: Gumbel vs. Normal $n = 500$, $\tau = 0.2$

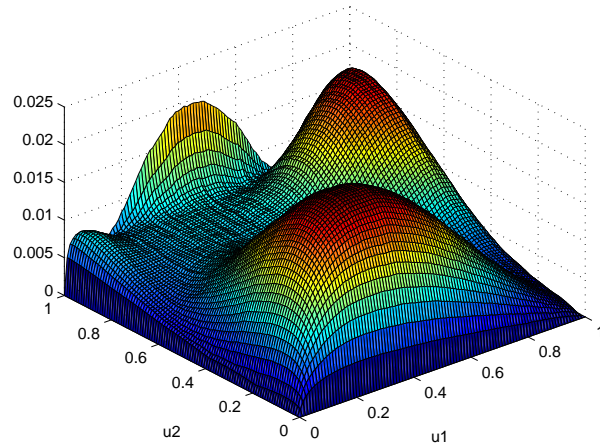


Figure D.3: Gumbel vs. Clayton $n = 500$, $\tau = 0.2$

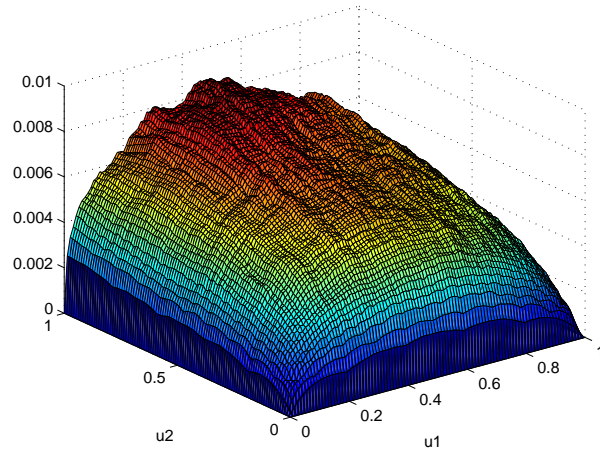


Figure D.4: Normal vs. Normal $n = 500$, $\tau = 0.2$

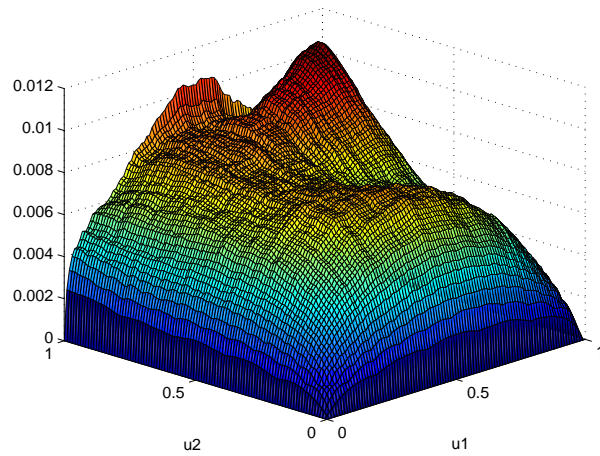


Figure D.5: Normal vs. Gumbel $n = 500$, $\tau = 0.2$

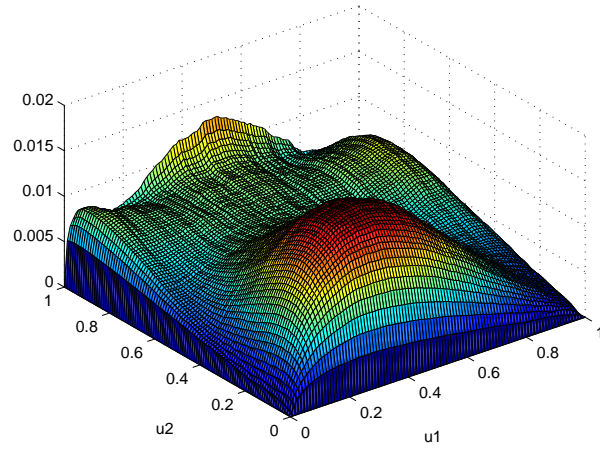


Figure D.6: Normal vs. Clayton $n = 500$, $\tau = 0.2$

Currirculum Vitae

Name: Mike McCausland

Place of Birth: Owen Sound, Ontario, Canada

Year of Birth: 1983

Post-Secondary Education and Degrees:

Brock University
St. Catharines, Ontario, Canada
2002–2006 BSc (Mathematics)

The University of Western Ontario
London, Ontario, Canada
2006–2007 M.A. (Economics)

The University of Western Ontario
London, Ontario, Canada
2007–2011 Ph.D. (Economics)

Honors and Awards:

Ontario Graduate Student Scholarship
The University of Western Ontario
2009–2010

Brock University Student Research Award
Brock University
2004–2005

Related Work Experience:

Teaching Assistant
The University of Western Ontario
2006–2011