DIGITAL COMMONS © UNIVERSITY OF SOUTH FLORIDA

University of South Florida Digital Commons @ University of South Florida

USF Tampa Graduate Theses and Dissertations

USF Graduate Theses and Dissertations

April 2019

Essays on Time Series and Machine Learning Techniques for Risk Management

Michael Kotarinos University of South Florida, wkotarinos@gmail.com

Follow this and additional works at: https://digitalcommons.usf.edu/etd

Part of the Finance and Financial Management Commons, Operational Research Commons, and the Statistics and Probability Commons

Scholar Commons Citation

Kotarinos, Michael, "Essays on Time Series and Machine Learning Techniques for Risk Management" (2019). USF Tampa Graduate Theses and Dissertations. https://digitalcommons.usf.edu/etd/8378

This Dissertation is brought to you for free and open access by the USF Graduate Theses and Dissertations at Digital Commons @ University of South Florida. It has been accepted for inclusion in USF Tampa Graduate Theses and Dissertations by an authorized administrator of Digital Commons @ University of South Florida. For more information, please contact digitalcommons@usf.edu.

Essays on Time Series and Machine Learning Techniques for Risk Management

by

Michael Kotarinos

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy Department of Mathematics and Statistics College of Arts and Sciences University of South Florida

Major Professor: Chris P. Tsokos, Ph.D. Lu Lu, Ph.D. Kandethody Ramachandaran, Ph.D. Yicheng Tu, Ph.D.

> Date of Approval: April 2, 2019

Keywords: K-th Moving Average GARCH, MACBETH, Multi-Level Time Series Clustering, Asset Allocation

Copyright © 2019, Michael Kotarinos

Dedication

This dissertation is dedicated to my mother Patricia for her constant and unconditional support over many years.

Acknowledgements

I would like to thank Professor Chris Tsokos for his support over many years. I am grateful for his many years of insights and cannot think of a better or more supportive mentor. I am grateful for all that he has done to support both my academic and personal development and cannot put into words how grateful I am.

I would also like to thank Professor Lu, Professor Ramachandaran and Professor Tu for serving on my dissertation committee and their support during my time at USF.

I would also like to acknowledge Professor Kamp for volunteering to chair my dissertation and also for his support over my time at USF. I had the pleasure of taking four different economics courses with him during my undergraduate studies. He was undoubtedly one of the best instructors I have had during my time at USF and am grateful he found time to serve as the chair of my defense.

I would like to acknowledge Vishruti for her work as a research assistant over the last year. Her help was invaluable in the development of several promising research applications.

TABLE OF CONTENTS

List of Tables	iii
List of Figures	iv
Abstract	vi
Chapter One: Introduction	1
1.1 Traditional Techniques for Risk Management	
1.2 Development of a Modular Framework for Customizable Risk Management	
1.2 Development of New Approaches for Assessing Risk and Return of Individual Assets	4
Chapter Two: Overview of Traditional Financial and Time Series Methods	4
2.1 Time Series Techniques for Risk Management and Assessment	4
2.2 Capital Asset Pricing Model and Markowitz	8
Chapter Three: Development of an Alternative Framework for Asset Allocation	
3.1 Introduction	11
3.2 Overview of New Asset Management Framework	12
3.3 Multi-Level Time Series Clustering	
3.4 Examples of Multi-Level Time Series Clustering	20
3.5 Multi-Criteria Decision Analysis in an Integrated Framework	21
3.6 A Note on Suitability	
3.7 Example of Portfolio Selection for a Value Based Investor	27
3.7.1 Selection Criteria	27
3.7.1 MCDA Process Example	
3.7.1 Time Series Methods for Chosen Tickers	32
3.7.1 Note on the Risk Free Rate and Riskless Assets in the Sharpe Computation	34
3.7.1 Comparison to Benchmark	35
3.8 Contributions	38
Chapter Four: Forecasting Single Asset Return in ETF Driven Markets: Revisiting the K-th	
Weighted ARIMA Model	40
4.1 Introduction	
4.2 Market Sectors and Exchange Traded Funds	
4.3 K-th Moving Average ARIMA Model	
4.4 K-th Weighted Moving Average ARIMA Model	
4.5 Example: Model for JNK	
4.6 Model Performance	
4.6.2 iShares MSCI Emerging Markets ETF (EEM)	
4.6.2 iShares China Large-Cap ETF (FXI)	
4.6.2 iShares Russell 2000 ETF (IWM)	
4.6.2 SPDR Bloomberg Barclays High Yield Bond ETF (JNK)	

4.6.2 SPDR S&P 500 ETF (SPY)	JT.
4.6.2 Financial Select Sector SPDR ETF (XLF)	52
4.7 Conclusions	
4.8 Contributions	
Chapter Five: New Moving Average Approaches for Forecasting and Characterizing Volatility	56
5.1 Introduction	
5.2 K-th Weighted Moving Average GARCH Model	70
5.3 Exchange Traded Fund Selection	
5.4 Data Collection	
5.5 Example: SPY Ticker Model Fitting	
5.6 Example: Results	
5.7 Example: Periodicity Modeling with Supervised Models	
5.8 Theoretical Implications of Volatility Periodicity Estimation	
5.9 Contributions	
Charten Star Paters Descent	22
Chapter Six: Future Research	
6.1 Applications of the Combined MLTS-MACBETH Approach Across Disciplines	13
6.2 Nonlinear Multi-Criteria Decision Analysis Topological Characterization of	22
Preferences	13
6.3 Theoretical Exploration of Non-Transitive Preferences and its Implications on	2.4
Selections	
6.4 Applications of Moving Average Procedures to Various GARCH Class Models) 4
6.5 Combination of a Decision Making Environment for Panel Data with a Bayesian	
Hierarchical Framework	
6.6 Determination of Optimal Rebalancing Periods) 5
References	€
Appendices)1
Appendices Appendix A: Brief Definitions of Selected Terms	

List of Tables

Table 3.1	Clusters and Numbers of Members in Each Cluster	21
Table 3.2	Number of Companies Chosen from Each Block	33
Table 3.3	Chosen Tickers	33
Table 3.4	Comparison of Generated Portfolios to Benchmark	36
Table 4.1	EEM residuals	51
Table 4.2	FXI residuals	52
Table 4.3	IWM residuals	53
Table 4.4	JNK residuals	55
Table 4.5	SPY residuals	56
Table 4.6	XLF residuals	57
Table 5.1	Unfiltered GARCH model parameters	68
Table 5.2	Fifth Moving Average GARCH model fit	68
Table 5.3	Fifth Weighted Moving Average GARCH model fit	70
Table 5.4	AIC values for modeling approaches	70
Table 5.5	AIC values of AAAFF GARCH volatility model during different business cycle periods	75
Table 5.6	AIC values of AAAFF 5-day moving average GARCH volatility model during different business cycle periods	75

List of Figures

Figure 3.1	Overview of clustering technique	14
Figure 3.2	MACBETH criteria weights	28
Figure 3.3	Box-and-Whisker plots of company utility	31
Figure 3.4	Two-Dimensional visualization	32
Figure 3.5	Performance over time	37
Figure 4.1	Daily adjusted JNK price	45
Figure 4.2	First order differenced JNK price	46
Figure 4.3	Auto-correlation of differenced JNK prices	46
Figure 4.4	Partial Auto-correlation of differenced JNK prices	47
Figure 4.5	Daily adjusted fifth MA JNK price	48
Figure 4.6	First order differenced fifth MA JNK price	49
Figure 4.7	Daily weighted MA adjusted JNK price	50
Figure 4.8	Daily differenced weighted MA adjusted JNK price	50
Figure 4.9	EEM residuals over time	52
Figure 4.10	FXI residuals over time	53
Figure 4.11	IWM residuals over time	54
Figure 4.12	JNK residuals over time	55
Figure 4.13	SPY residuals over time	56
Figure 4.14	XLF residuals over time	57

Figure 5.1	Time series of IWM returns; unmodified	63
Figure 5.2	Time series of IWM returns; modified with K-th moving average	64
Figure 5.3	Time series of IWM returns; modified iwth K-th weighted moving average	65
Figure 5.4	Time series of AAAFF prices	72
Figure 5.5	Region of AAAFF corresponding to March 2001 peak	72
Figure 5.6	Region of AAAFF corresponding to November 2001 trough	73
Figure 5.7	Region of AAAFF corresponding to December 2007 peak	73
Figure 5.8	Region of AAAFF corresponding to June 2009 trough	73
Figure 5.9	Price of AAAFF during March 2001 peak	74
Figure 5.10	Price of AAAFF during November 2001 trough	74
Figure 5.11	Price of AAAFF during December 2007 peak	75
Figure 5.12	Price of AAAFF during June 2009 trough	75

Abstract

The Capital Asset Pricing Model combined with the Sharpe ratio is a standard method for choosing assets for selection in a portfolio. However, this method has many structural issues and was designed for a time when high dimensional computing was in its infancy. An alternative to these methods using a mix of Multi-Level Time Series Clustering, the MACBETH algorithm and traditional time series techniques was constructed that minimized data loss and allow for customized portfolio construction for investors with different risk profiles and specialized investment needs. It was shown that these methods are adaptable to cloud computing environments and allow for modular customization as needed, while also being quite powerful and adaptable as developed. These approaches extend the risk-return foundation of finance into a risk-return-suitability framework that is more in line with the methods being used by most financial practitioners and regulators.

In addition to methods for portfolios, new techniques for the selection, screening, and analysis of individual assets were developed. The K-th Moving Average approach to ARIMA forecasting (2007) was extended for volatility forecasting to allow for estimation in a GARCH environment. New analysis techniques to combine these methods with machine learning methods were also developed and shown to yield unique insights into the decomposition of signal performance.

The new proposed approaches are data-driven analytical characterizations that combine theory and practice to generate new characterizations and insights into the workings of financial markets and systems by using a mix of existing methods with newly developed techniques. These new methods develop allow for a more sophisticated approach to the understanding of risk and volatility on multiple levels with broad applications.

1 Introduction

1.1 Traditional Techniques for Risk Management

One of the major challenges that has always faced managers is the management of risk. New opportunities, projects, investments, or ventures involve both risk and uncertainty and determining ways to curtail, assess, and manage both of these aspects are a critical part of any organization's success.

The first chapter of this dissertation focuses on traditional methods developed in finance to assess the risk of a pool of assets in a portfolio. Early techniques developed in this field had numerous issues, but also were a parsimonious way of characterizing and modeling risk. Over time, these methods retained their popularity due to ease of use and an understanding of the issues these methods face.

Since the time these methods were developed new techniques and methodologies are now available and older techniques which were intriguing but not implementable at the time are now viable options. The development of cloud based computing, advances in GPU technology and the increased adoption of distributed computing networks means that new algorithms and techniques can be deployed to characterize and assess risk.

The first chapter discusses older techniques in the context of newer advances made in the quantitative sciences. After contextualizing these older results, the chapter discusses the issues that these techniques faced and the necessary issues that any new framework for understanding risk must address. This sets the foundation for the discussion of new methods to characterize and assess risk and the criteria under which these techniques will be evaluated and interpreted.

1.2 Development of a Modular Framework for Customizable Risk Management

One of the major principles of coding and development in data science has been the advent of modular programming. In a modular programming framework, complex coding problems are broken down into smaller, manageable pieces. Each of these pieces is coded and often reused throughout a problem. This approach allows for a systematic approach for testing and allows for quick customization in the process.

The new risk management framework proposed in Chapter 3 is modular in its construction. This approach to risk management is a broad series of tools that are combined together to perform a high dimensional characterization of risk, assess the most important factors to a given investor in the choice of investments, and select an appropriate portfolio that maintains the relevant spatio-temporal information in the data throughout the decision making process. The modularity of this framework means that the individual pieces and components of the approach can be further customized and refined to address unique needs or to allow for decision making with complex preference structures.

This chapter's approach on categorizing and assessing risk extends the ideas of modern portfolio theory and traditional techniques into a high dimensional computing environment and allows for methods that combine the complexity of modern day data science techniques with the ease of execution and implementation that existed in earlier methods. These portfolios are shown to need relatively little re-balancing and are an attractive option for a wide class of investors.

1.3 Development of New Approaches for Assessing Risk and Return of Individual Assets

Much of the focus in recent years has been on the performance not of individual assets but of a portfolio as a whole. In modern portfolio theory, the focus is on managing a pool of assets rather than trying to identify individual assets for selection. That being said, there is still a great deal of emphasis on choosing and understanding individual assets.

In many asset management environments, mathematical and statistical techniques are used to create a shortlist of potential assets for further consideration in a larger portfolio. Before one of these assets can be included its behavior, suitability, risk, and various other factors have to be considered. After determining that it may be appropriate, it must then be assessed whether it is an appropriate fit for the remaining assets. As such, a great deal of analysis is still often performed at the individual security level.

In addition to the focus on the security on the portfolio as a whole, there is still some focus given to the attractiveness of a given asset. For example, a company might be considered undervalued, have strong growth prospects, or have relatively little risk. Thus, there is still a significant amount of desirability for understanding the risk, reward, and desirability of individual assets.

Chapter 3 focused on developing a framework for understanding the movement of a portfolio. Chapters 4 and 5 focus on the development of techniques for analyzing the risk and return of individual assets. Chapter 4 introduces the K-th moving average and weighted moving average approach to analyzing a stock. This method is generalized to volatility in Chapter 5. Chapter 5 also extends the discussion of volatility by developing a new technique to deconstruct a volatility measure and analyze its performance using labeled data. This approach serves as a supervised learning alternative to the unsupervised spectral decomposition that is traditionally used. The use of this supervised method establishes a new way of thinking about the decomposition of variance using insights from the distinction in machine learning between supervised and unsupervised learning.

2 Overview of Traditional Financial and Time Series Methods

In this chapter, a brief overview will be given of some foundational work regarding traditional time series techniques and their applications to managing, assessing, and categorizing risk. After discussing these applications, the chapter will proceed with a brief review of traditional risk management methods.

2.1 Time Series Techniques for Risk Management and Assessment

Financial services has a long and storied history of focusing not just on the expected return of a set of assets but also on the risks and uncertainties centered around these cash flows. While later parts of this chapter will focus on the impact of these returns on the nature of markets, this particular chapter focuses on the development of time series methods specifically focusing on handling risk and uncertainty.

One of the primary tools developed in the 1980s to predict and assess the structural nature of volatility is Generalized Autoregressive Conditional Heteroskedastictity (GARCH), which is a generalization of the similarly named ARCH process. Suppose that one were to believe that the volatility of a current signal was related in some way to prior signals in a process. One might then be lead to believe that an appropriate model may be

$$y_t = a_1 y_{t-1} + a_2 y_{t-2} + \dots + a_q y_{t-q} + \epsilon_t$$

where $\epsilon_t \sim N(0, \sigma_t^2)$. This model is commonly referred to as an AR(q) model, which informally can be considered to be a process that uses the past behavior of y_t to charac-

terize its behavior. Note that in this particular formulation the variance is allowed to be hetereskedastictic over time¹. Similarly, one could instead use a model such as

$$y_t = \epsilon_t x_{t-1}$$

where x_t is an exogenous representation of variance. This model may seem appealing as it is a rather straightforward process, but the structural implications about convergence are quite strange. The assumption that a distribution could be constant in terms of mean over time but not in terms of variance seems to be quite limiting, and does not allow for the co-evolution and natural change of both moments. Granger and Anderson (Granger & Anderson, 1978) proposed the model

$$y_t = \epsilon_t y_{t-1}$$

which removed the need for an exogenous x_t variable. However, the model was shown to have issues with convergence on unconditional variances. The standard ARCH model addresses these convergence issues with the modification of using

$$y_t = \epsilon_t h_t^{1/2}$$

$$h_t = \alpha_0 + \alpha_1 y_{y-1}^2,$$

with $V(\epsilon_t = 1)$ (Engle, 1982). The standard ARCH model is said to be of order p in terms of the number of α_i terms used in the model. Engle noted caution with use of the model, noting that covariance stationarity is equivalent to the characteristic equation of the parameters has no roots inside the unit circle.

¹The second moment is not assumed to be constant over time. Note that a non-constant variance over time would be a violation of strong stationarity, but not necessarily a violation of weak stationarity.

The ARCH model was extended to allow for past conditional variances in the estimate of the current conditional variance. For parameters p and q, the GARCH model (Bollerslev, 1986) extends the ARCH specification by redefining

$$h_t = lpha_0 + \sum_{i=1}^q lpha_i \epsilon_i^2 + \sum_{i=1}^p eta_i h_{t-1}.$$

One informal way of characterizing the GARCH process is considering it to be the equivalent of an extending from an autoregressive approach to an autoregressive moving average approach. One particularly common explanation favored by Engle is that "the best predictor of the variance in the next period is a weighted average of the long run average variance, the variance predicted for this period and the new information this period (Engle, 2008)." Engle draws the conclusion that perhaps one could think of this updating process as being similar to the traditional Bayesian updating process².

Numerous competing versions of GARCH have evolved over the years, such as IGARCH and EGARCH. It is not uncommon for new GARCH techniques or specifications to appear in the literature, although the actual level of improvement over standard GARCH estimates is debatable³

. One particularly interesting observation about the nature of the estimation process is that it is entirely focused on the sigma parameter, which is a measure of risk. There has been a significant amount of debate in recent years regarding not just risk but also

²The Bayesian connection is debatable. Traditional Bayesian methods do not look or observe any data before making statements about parameterization, while GARCH processes are implemented as canned algorithms with a maximum likelihood estimation procedure. It could perhaps be said that the Empirical Bayes framework is a better representation of the nature of a GARCH process than a Bayesian update.

³It is not uncommon to hear practitioners in quantitative finance using GARCH(1,1) as a workhorse model, similar to how many statisticians use ordinary least squares (OLS) as a workhorse for many applied problems. Although much of the literature seems to focus on new applications or ways to implement GARCH based frameworks, most applied quantitative analysts do not seem to believe these models to be significantly different enough from GARCH to warrant investigation. Quantitative analysts often seem wary to deviate from other parameterizations of GARCH other than GARCH(1,1), and as such the idea of exploring alternative structures of parameters seems unlikely in a field that is unwilling to even consider different parameter orders. A thorough study exploring alternative GARCH models found none of them convincingly outperformed GARCH(1,1), so perhaps the applied financiers have some insights the theorists do not (Hansen & Lunde, 2005).

uncertainty as well. Taleb argued that "black swan" events based on uncertainty draw on the fallacy of assuming risk is the only real measure of risk and can lead to catastrophic market failures (Taleb, 2007). Among the many issues Taleb raised, he noted that:

- Past behavior does not always describe future behavior, especially after a change has been implemented⁴.
- 2. The over-reliance on risk measurements does not properly characterize the nature of uncertainty, especially when uncertainty is not related to the measure of risk.
- 3. Faulty statistical and mathematical assumptions can have serious implications.

In particular, the assumption of default rates in housing prices being relatively independent events over time was found to be true when the default rate was relatively low, but far from true when the rate rise, creating numerous financial issues. While much of time series analysis has focused on and admitted issues relating to long-term forecasting, not as much attention has necessarily been paid to uncertainty modeling. While game theory has shown some intriguing applications for modeling uncertainty, its statistical applications at times have been dubious⁵. While military applications found great use given the great deal of uncertainty and because of the important role of the players in the game (namely the United States and the U.S.S.R. but not exclusively), such game theoretic frameworks remain popular to this day in this domain. In statistics, game theory is mostly used as a framework for parameter estimation, and as such has not been as popular as a tool for assessing risk⁶.

⁴such an observation is usually the basis for a differences-in-differences regression procedure

⁵Due to Von Neumann and Morgenstern's seminal book arriving in the 1940s and Nash's work in the 1950s arriving during the cold war, game theory has always been a topic of military strategists and theorists and was a major focus of the RAND corporation at the time. Long (Long, 2008) has a summary of some of the topics that formed the basis for the military's focus during these times and their game theoretic relations.

⁶Some private equity firms and sophisticated quantitative finance groups have at times used game theory in their assessments. Such models are typically proprietary in nature and as such are difficult to discuss at great length.

2.2 Capital Asset Pricing Model and Markowitz

One of the most well known and utilized models in the financial literature is the Capital Asset Pricing Model (CAPM). CAPM can be considered as a framework for evaluating risk and return in portfolio construction and is often used as a simplistic approach for making decisions regarding assets in a portfolio.

The CAPM approach builds on the earlier work of Markowitz, who argued that (Markowitz, 1952) :

- Portfolio selection begins with observations and experience forming beliefs⁷ about future performance of securities.
- 2. Beliefs about future performance of securities lead to the selection of securities.
- 3. Investors will desire higher expected returns for a portfolio, and will attempt to avoid portfolios with high variances relative to returns.
- 4. In order to achieve (3), an investor will use diversification⁸.

Markowitz then proceeds with the construction of a frontier of portfolios. The portfolios on the edge of the frontier can be considered possible allocations with differing levels of Von Neumann-Morgenstern utility, as they represent trade-offs between risk and return (Von Neumann & Morgenstern, 1953). This characterization leads to dominated portfolios, establishing a Pareto frontier across asset mixes (Markowitz, 1959).

This approach was not without its detractors. Borch noted that for a preference over a family of distributions $F(x, m_1, m_2, ..., m_n)$ with first moments of some n distributions that there exists a utility function

⁷Markowitz's work could be considered quite Bayesian in belief about reality, but is structured to be completely frequentist. The possibility of a Bayesian alternative construction is an appealing area for future discussion.

⁸This is set as an alternative to merely focusing on expected discounted cash flows and emphasizes parts of what is now called modern portfolio theory.

$$U(m_1...m_n) = \int_{\mathbb{R}} u(x) dF(x, m_1...m_n)$$

which exists if and only if u(x) is a polynomial of degree n (Borch, 1953). If this utility representation is indeed a utility of money as described by Markowitz (Markowitz, 1952) and u'(x) > 0 and u''(x) < 0, then u(x) cannot be a polynomial⁹. This leads to Borch's conclusion that one of the following two are violated in the Markowitz framework:

- 1. The consistency conditions of Von-Neumann Morgenstern Utility,
- 2. The usual preferences over the utility of money.

Borch's paper was not the only detractor. It was later shown that the convex downward nature of the Pareto frontier that is assumed in the Markowitz model for a two parameter distribution is only true for a select few distributions (primarily normal), and that there is no theoretical basis for this assumption (Feldstein, 1969). Hakansson noted that generalizations to multiple period cases produce similar inconsistencies to Borch's criticisms of the single period case (Hakansson, 1971). These should not be considered the only objections to this approach. In particular, Fama and French's objections lead to the five factor model commonly used by financiers in assessing portfolios.

While there are many criticisms of this approach, in many ways its popularity is due to the later work by Sharpe, Linter, and Mossin in the development of a mean-variance equilibrium model (Sharpe, 1964). This model used the risk-return framework to create a theoretical price of assets given by

$$E(R_i) = R_f + \beta_i E(R_m - R_f)$$

which equates the return of some asset i in terms of the risk free rate, its beta, and the return on the market. Assets that are greater than the expected return are said to be

⁹This is a fairly reasonable characterization of the derivative. This merely means the utility of money is increasing but the marginal utility of money is decreasing.

performing above the benchmark line. This amount is known as alpha, and as such the assets are said to have a positive alpha.

Similar to the Markowitz framework, this model also has numerous issues. The first is that as it is based on the Markowitz framework, all of the complaints previously mentioned still stand. Some of the more well known papers criticizing specifically the CAPM framework are (Roll, 1977) (Fama & French, 1993) (Lai, 2015).

Despite its many criticisms and complaints, CAPM is still widely used. The model is mathematically straightforward, easy to use, and provides a simple framework for inexperienced financiers to evaluate portfolios, select assets, and describe risk and return. While there are many criticisms of this framework, the most common issue is that alternatives are usually either too mathematically complex for many financiers to grasp or often have their own theoretical issues in construction.

3 Development of an Alternative Framework for Asset Allocation

3.1 Introduction

At the time of the introduction of CAPM and the Sharpe ratio, computing was limited to large servers at universities and research institutes using data punch cards to perform calculations. Because of this, it was not uncommon for some matrix calculations to be performed by hand. The Capital Asset Pricing Model's simplistic approach when used in combination with the Sharpe ratio may have been flawed but was feasible to implement on a large scale at many financial institutions.

Much of the current discussion and implementation of machine learning and Bayesian methods in the last several years has been possible thanks to the rapid acceleration and development of computing hardware and the relatively inexpensive availability of cloud computing. In addition to advances in Central Processing Unit (CPU) speed, recent advances in Graphical Processing Unit (GPU) architecture have lead to most individual researchers having access to computing power in a single desktop equivalent to that of an entire server lab ten years ago. In the case that more computing power is needed, corporate clients can now contract for computing time on an as needed basis from companies such as Microsoft, Oracle, Amazon, and Google. The result is that many new techniques can now be applied and implemented that were not feasible at the time of Markowitz.

This section develops a new framework for assessing portfolios not just in terms of risk and return but also in terms of suitability. After discussing multi-level time series (MLTS) clustering, this chapter continues with a discussion of Multi-Criteria Decision Analysis (MCDA) models and their applications. An example is then shown of MCDA techniques applied to the MLTS framework as an alternative approach to selecting assets. Investor suitability and its importance is briefly discussed, which is covered in more detail in later chapters. This combination of MCDA techniques with a MLTS approach is shown to produce a basket of possible portfolios that are suitable to unique investor needs, which can be adjusted in terms of risk versus return using traditional techniques such as the Sharpe ratio. The suitability extends the risk-return theoretical foundation of finance into a risk-return-suitability theoretical foundation, which resolves many of the theoretical two-parameter issues that have plagued mathematical characterizations in finance. It is noted that many of these approaches are modular in nature, and can be adjusted as needed as the field of asset management continues to evolve. These approaches are datadriven in their construction, using a mix of currently developed algorithms with newly developed approaches to understand complex financial systems.

This chapter concludes with a note on the contributions of this work to the literature.

3.2 Overview of New Asset Management Framework

The proposed new asset management framework involves a six step process:

- 1. Construct a database of stocks. Run a MLTS clustering algorithm on each sector, breaking off stocks into individual groups.
- 2. In each of these groups, sub-cluster based off the selection criteria.
- 3. Select the stocks from the sector matching the user's preferences (such as value driven) and add them to a portfolio.
- 4. Weight the stocks so that each stock's representation is equivalent to desired weight¹.
- 5. Use GARCH and ARIMA models to estimate the portfolio's risk and return over time.

¹Portfolio managers often have desire to "overweight" or "underweight" representation of an asset class relative to the market. In this case, one can take an equal weight to that asset class representation in the S&P 500, adjusted for sectors that are not present.

- 6. Measure the risk/return profile of the portfolio. Drop and add the single stock that results in the largest return in the risk return profile. This is done by:
 - (a) If there is a stock that will result in a higher return with lower or equivalent risk, drop the stock with the biggest return profile increase first.
 - (b) See if a previously dropped stock can be added back in to decrease risk with no effect on return then add it back in. Else, drop the next stock that increases return without lowering risk.
 - (c) If there are no possible stocks that can increase return without lowering risk then stop and record this as a potential portfolio on the efficient portfolio frontier.

This process can graphically be seen by Figure 3.1.

The modularity of this approach means that different steps in the phase can be replaced by other techniques or approaches as the literature advances or to meet unique investment needs. For example, the GARCH model for variance could be replaced by an IGARCH characterization, the MCDA approach could allow for non-linear representations, and a criteria other than the Sharpe ratio

3.3 Multi-Level Time Series Clustering

Some of the most commonly used tools in machine learning involve clustering. From rather simple tools like K-means to more advanced methods like stochastic maps, being able to separate out data into useful groups has broad applications across disciplines (Jain, 2010). In particular, machine learning based clustering techniques are widely used by the vast majority of machine learning practitioners on countless applied problems. With the advent of cloud computing, parallel processing and advances in GPU architecture many more advanced statistical techniques are now available to practitioners that were computationally unfeasible during the time that CAPM and the Sharpe ratio were developed.

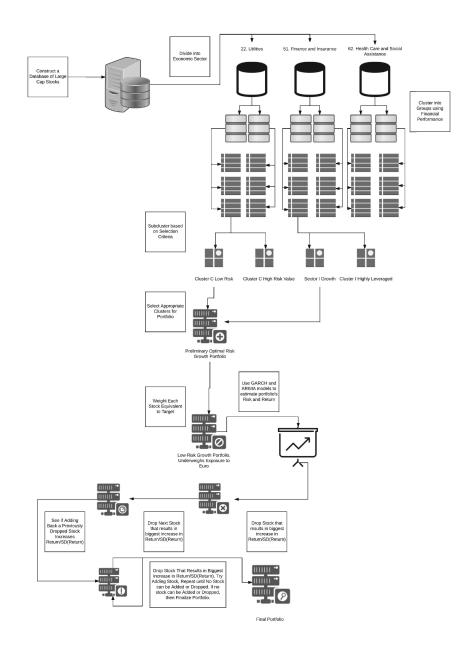


Figure 3.1: Overview of clustering technique

One of the most common features of data in econometrics is a panel based structure². In most data commonly seen in textbook problems, data is independently and identically distributed and each observation's placement in the data set is not inherently meaning-ful³. In time series data, each observation has meaning not only by itself but in relation to other terms. Clive Granger is known for clarifying this relationship by describing a beaded necklace thrown on a table; the distance each bead on the necklace has from the edge of the table is related to the bead before and after it in a necklace, and as such our data has additional dimensionality not described by standard statistical techniques (Granger, 1981). Similar to the way time series econometrics often has its own way of specifying regressions and modeling complex relationships, so too does it need special tools for clustering (Wang et al., 2006).

The particular approach this section focuses on is Multi-Level Time Series Clustering (MLTS clustering). Most time series clustering techniques focus on determining time dependencies and exploiting these to create cluster structures (Xiong, 2004). Rather than focusing on the cumulative effect of each lag dependency, MLTS clustering expands upon the current literature by investigating each lag dependency independently (Doo Young Kim & Tsokos, 2018). While this addresses many computational issues in the matrix representation, it also has some benefits that are more subtle in finance regarding the structural nature of risk (Kotarinos et al., 2019).

MLTS Clustering involves a seven step process occurring over two phases:

Phase 1: Clustering

- 1. Calculate pairwise distance in each lag of interest.
- 2. Construct dissimilarity matrices based on lags.
- 3. Perform clustering using aforementioned dissimilarity matrices.

Phase 2: Portfolio Selection

²referred to in the statistical literature as "time series data." These terms will be used interchangeably. ³This is often empirically tested by checking the significance of auto-correlation terms.

- 1. Select trading interval (target lag k)
- 2. Choose number of stocks within same cluster for each sector.
- 3. Investigate neighborhood lag solutions.
- 4. Make final selection of stocks.

The algorithm by Kim and Tsokos uses a modification of a weighted Mahalanobis distance to construct these dissimilarity matrices then Ward's method to reach an analytic solution⁴ (Ward, 1963).

More formally, the Kim and Tsokos defines a cross lag dissimilarity matrix for a cross lag by defining $_kP_i$ as the daily stock prices of company *i* after removing *k* rows from the front and $P_{j,k}$ as the daily stock prices of company *j* after removing *k* rows from the tail. In other words,

$${}_{k}P_{i} = \begin{bmatrix} h_{i,k+1} & l_{i,k+1} \\ h_{i,k+2} & l_{i,k+2} \\ \vdots & \vdots \\ h_{i,T} & l_{i,T} \end{bmatrix}$$

and

$$P_{j,k} = \begin{bmatrix} h_{j,1} & l_{j,1} \\ \vdots & \vdots \\ h_{j,T-1-k} & l_{j,T-1-k} \\ h_{j,T-k} & l_{j,T-k} \end{bmatrix}$$

where $h_{i,k}$ and $l_{i,k}$ denote the daily maximum and minimum stock prices. Then, the backward and forward distance matrices can be derived by

⁴Other methods can also be used based on the preferences of the analyst.

$${}_{k}D =_{k} P_{i} - P_{j,k} = \begin{bmatrix} {}_{k}d_{h,1} & {}_{k}d_{l,1} \\ {}_{k}d_{h,2} & {}_{k}d_{l,2} \\ \vdots & \vdots \\ {}_{k}d_{h,T-1-k} & {}_{k}d_{l,T-1-k} \\ {}_{k}d_{h,T-k} & {}_{k}d_{l,T-k} \end{bmatrix}$$

and

$$D_{k} = P_{i,k} -_{k} P_{j} = \begin{bmatrix} d_{h,k,1} & d_{l,k,1} \\ d_{h,k,2} & d_{l,k,2} \\ \vdots & \vdots \\ d_{h,k,T-1-k} & d_{l,k,T-1-k} \\ d_{h,k,T-k} & d_{l,k,T-k} \end{bmatrix}.$$

These distance matrices use the Mahalanobis distance given by

$$d_{i,j,k} = \frac{1}{2} \left(\sum_{t=1}^{T-k} \sqrt{k D_t S^{-1} M_t'} \cdot W_t + \sum_{t=1}^{T-k} \sqrt{D_{t,k} S_k^{-1} D_{t,k}' W_{t,k}} \right)$$

In the Mahalanobis distance the weights are defined by

$$_{k}W_{t} = \frac{\frac{1}{2(T-k)}(|H1| + |L1|)}{\sum_{t=1}^{T-k}(|H1| + |L1|)}$$

and

$$W_{t,k} = \frac{\frac{1}{2(T-k)}(|H2| + |L2|)}{\sum_{t=1}^{T-k}(|H2| + |L2|)}$$

with the terms

$$H1 = \sum_{\tau=1}^{t} ({}_{k}d_{h,\tau+T-k-t} - {}_{k}\bar{d}_{h})({}_{k}d_{h,\tau} - {}_{k}\bar{d}_{h}),$$

$$L1 = \sum_{\tau=1}^{t} (_{k}d_{l,\tau+T-k-t} - _{k}\bar{d}_{l})(_{k}d_{l,\tau} - _{k}\bar{d}_{l}),$$

$$H2 = \sum_{\tau=1}^{t} (d_{h,k,\tau+T-k-t} - \bar{d}_{h,k})(d_{h,k,\tau} - \bar{d}_{h,k}),$$

$$L2 = \sum_{\tau=1}^{t} (d_{l,k,\tau+T-k-t} - \bar{d}_{l,k})(d_{l,k,\tau} - \bar{d}_{l,k}).$$

The MLTS technique then advocates choosing some stocks from each basket to compose a portfolio, but does not state how the stocks should be chosen, weighted, or combined to form an appropriate selection of assets.

The approach by Kim and Tsokos is an intriguing addition to the literature on portfolio selection. Consider the alternative approach of using the Sharpe ratio in isolation, given by

$$Sharpe_{portfolio} = \frac{R_{portfolio} - R_{Riskless}}{\sigma_{portfolio}}$$

which expresses the return of a portfolio based on its return, variance, and the performance of a risk-free asset. This is a low dimensional measure of risk that only considers the volatility and return of a portfolio at a single point in time. Used in conjunction with CAPM, this is a rather blunt tool that reduces the dimensionality of the data but also cuts out much of the interesting cross-sectional relationship in the data.

Consider instead the approach by Kim and Tsokos. In this approach, panel data is clustered and grouped by actual differences in the underlying behavior of the series and lagged relations are held and maintained until the end of the process. In this case, note that:

1. The information on significant lagged differences is preserved throughout the process.

- 2. One of the goals of modern portfolio theory is to choose assets that exhibit differing movements over time to gain exposure to diverse markets. Rather than relying on a one-dimensional measure such as correlation, assets from different blocks within a sector will have statistical discrepancies that suggests they face different market exposures.
- 3. Hedge funds will be able to select stocks from similar blocks to create customized portfolio exposures with low levels of correlation to the market by mixing short and long positions on stocks within each block.
- 4. The approach is quite general and can be applied to problems beyond quantitative finance and financial econometrics.

That being said, there are some issues with the current approach as developed by Kim and Tsokos.

- 1. The approach does not have a straightforward implementation for assessing tradeoffs between risk and return, as seen in the Sharpe ratio.
- 2. The approach is more complex than the Sharpe ratio and is not as easy for those outside of quantitative finance to understand.
- 3. The approach does not address whether a given portfolio is appropriate or suitable for a given investor⁵.
- 4. The approach may not address an investor's unique needs or goals.
- 5. The approach for choosing assets does not state how the assets should be chosen from a basket of possible stocks.

While the Kim and Tsokos model offers an excellent way of categorizing and binning assets in a way that has a great deal of potential to financial institutions and investors, its

⁵The Sharpe ratio only investigates suitability in terms of risk and return, and as such also leaves much to be desired.

issues mean that it can face severe difficulties in implementation as a complete framework in its current form without any extensions.

3.4 Example of Multi-Level Time Series Clustering

Price data for securities composing the Standard and Poor's 500 index (S&P 500) was taken over a 10 year interval, with the first 8 years being taken for clustering and the last two years being used for validation. Five day blocks were constructed to mimic the five-day trading structure of most exchanges. The MLTS clustering algorithm was run on 8 sectors⁶ and lead to the creation of blocks of size 4 in each sector. Some manual corrections were made to accommodate some unique features⁷ and discrepancies⁸ in the data set. The 8 sectors were created via the combination of similar sector codes. The results of the clustering can be seen in Table 3.1.

Note that there is a discrepancy between the total number of companies and the number of total companies in the S&P 500. Some tickers were purged or modified as a result of incomplete data (the company dropping out of the index, being acquired, or merging with another company, or a new company joining the index), and in some cases manual corrections of index discrepancies were made. The use of universal CUSIP codes addressed most of the issues regarding ticker discrepancies, but some manual corrections still needed to be made.

- WYNN was moved to the same sector as MAR.
- PCL was moved to the same sector as L.
- AES was moved to the same sector as PCG.

⁶Sectors were based on North American Industry Classification Code. NAISC codes are six digit codes, with the first two fields designating economic activity, the third field designating sub-sector, fourth field designating industry group, fifth the NAICS industry and sixth the national industry (which is 0 if same as NAISC industry). For more information on these codes, see (NAICS, 2017).

⁷The seventh sector was a mix of nineteen companies in unique industries that were poorly described by other blocks.

Sector-Block	Members	Sector-Block	Members
1-1	19	5-1	9
1-2	3	5-2	27
1-3	1	5-3	9
1-4	2	5-4	4
2-1	2	6-1	34
2-2	25	6-2	18
2-3	5	6-3	3
2-4	1	6-4	4
3-1	90	7-1	30
3-2	58	7-2	7
3-3	11	7-3	7
3-4	2	7-4	4
4-1	33	8-1	1
4-2	9	8-2	2
4-3	2	8-3	1
4-4	2	8-4	1

Table 3.1: Clusters and Numbers of Members in Each Cluster

At this point, the Kim and Tsokos approach would select stocks based on a chosen distance metric, and use this in the construction of a portfolio. The MLTS approach is instead extended here to use Multi-Criteria Decision Analysis as an analytic framework for asset selection.

3.5 Multi-Criteria Decision Analysis in an Integrated Framework

Multi-Criteria Decision Analysis (MCDA) refers to a class of techniques often used by consultants to help clients evaluate decisions with multiple decision criteria (Greco, 2016). In the typical MCDA framework, an individual is asked some generic questions to help elicit the criteria he or she believes to be important to the decision making process. After understanding some of the individual's goals and desires, the consultant then proceeds to ask a series of questions that involve the individual choosing between different scenarios and forming a set of preferences. Over time, these questions yield information that can be used for parameter estimation.

One of the most common techniques in MCDA is the Measuring Attractiveness by a Categorical Based Evaluation Technique (MACBETH) approach. Under MACBETH, individuals are given short, relatively simple to answer questions that slowly over aggregation reveal an individual's preferences (Bana e Costa et al., 2005). For example, an individual choosing between two jobs may have criteria regarding location, income, hours worked, and commute time. The MACBETH algorithm would eventually be able to determine an individual's preferences over income and commute time by presenting scenarios where location and hours worked are the same and determining relative preferences between these two variables. Over time, by repeatedly testing trade-offs between these variables the algorithm generates relative weights on the importance of each criteria. One feature to note is that individuals often have irrational preferences, and if MACBETH notices a transitive preference inconsistency⁹ it will ask the user additional questions for clarification.

The MACBETH procedure begins by asking a Decision Maker (DM) pairwise comparisons about which feature is more important in a decision making process. After this, the user is asked to ranked the relative desirability of one condition versus the other based on a six level system: very weak (C_1), weak, (C_2), moderate (C_3), strong (C_4), very strong (C_5), and extreme (C_6). The result is that for a series of pairwise comparisons an upper triangular method is populated with elements stating levels of preferences between states. Not all $\frac{n(n-1)}{2}$ comparisons are needed, and n - 1 comparisons are sometimes sufficient if the user has consistent preferences across states. Additional comparisons are often made to ensure consistency and transitivity of preferences.

For each qualitative judgment, for each x that is an element of the state space X the valuation function v(x) implied by the preferences is checked to see if the following hold:

 $\forall x, y \in X : xIy \Rightarrow v(x) = v(y),$

⁹that is for bundles A, B, and C we have $A \succ B \succ C \succ A$

$$\forall x, y \in X : xPy \Rightarrow v(x) > v(y),$$

$$\forall x, y \in C_1 \cup \ldots \cup C_s$$

and

$$\forall w, z \in C_{i'} \cup \ldots \cup C_{s'} \text{ with } i, s, i', s' \in \{1, 2, 3, 4, 5, 6\}$$

$$i \leq s$$
 and $i' \leq s' : i > s' \Longrightarrow v(x) - v(y) > v(w) - v(z)$.

These conditions form a pre-cardinal set of value information. The conversion of precardinal information to pre-cardinal scaling is done via an algorithm that solves the linear system given by

 $Min[v(x^+ - v(x^-)]$ such that:

Condition 1: $v(x^-) = 0$

Condition 2: $v(x) - v(y) = 0, \forall x, y \in C_0$

Condition 3: $v(x) - v(y) \ge i, \forall x, y \in C_1 \cup \ldots \cup C_s$

with $i, s \in \{1, 2, 3, 4, 5, 6\}$ and $i \leq s$.

Condition 4:
$$v(x) - v(y) \ge v(w) - v(z) + i - s', \forall x, y \in C_1 \cup \ldots \cup C_s$$

and
$$\forall w, z \in C_{i'} \cup \ldots \cup C_{s'}$$
 with $i, s, i', s' \in \{1, 2, 3, 4, 5, 6\}$,

$$i \leq s, i' \leq s', i > s'.$$

The solution process may not be unique, and as such there are a variety of supplementary tools that are sometimes applied to construct and characterize and the nature of the solution space in a decision making environment. Recall that when discussing MLTS clustering, the process described a process for clustering assets for selection but did not state how the assets can be chosen. By combining the clustering structure from a MLTS algorithm with the pre-cardinal scaling from MACBETH, a solution frontier can be developed on an individual level for a customized investment portfolio.

One of the features that is intriguing is that many of these activities are similar to the tasks performed by financial planners and asset managers, although these algorithms are quite broad and as such could very easily be applied to diverse interdisciplinary problems. Asset managers operating under a fiduciary duty must put the needs of the client ahead of his or her own needs. In order to meet this responsibility, asset managers must discuss with clients their financial goals, objectives, any financial obligations, and develop a plan to help achieve these goals. In this process, the consultant must make a plan that takes multiple objectives, criteria, and performance goals into account.

Only the very wealthy typically have access to comprehensive financial planning. A shortage of financial planners combined with high opportunity costs for advisers leads to most individuals having little to no access to affordable financial planning options. Further, most financial planners servicing poorer clients will typically offer limited services and offer more generic plans such as "60-40" bond-stock allocations. The lack of readily

available financial planning combined with the fact that unsophisticated investors often make numerous investing errors means that many individuals outside of the top 10% of the United States income distribution are left out of the gains from markets, investments, and passive income.

Even with the widespread increase in "robo-investing" style options these tend to be generic and lack the specificity offered by a comprehensive financial planner. While MACBETH provides many of the tools needed to generate financial plans and goals, it does not offer a complete framework for portfolio selection which has limited its applications to finance. While it can state whether a security is locally appropriate, it is not ideal for stating whether a security is globally appropriate. With the move to modern portfolio theory this proves to be a major challenge to MACBETH.

Recall from section 3.3 that after performing a time series cluster, the authors took assets and grouped them within clusters based on their actual variation in time series signals. The result is that these signals have an additional level of diversification beyond that given by the Sharpe ratio. As an example, consider the stocks of Facebook, Alphabet (Google), Visa and Nvidia. In a traditional asset evaluation framework, all four of these would be considered "tech stocks" and one would note that Alphabet and Facebook appear highly correlated. With the rise of Exchange-traded funds, this co-integrated structure between Alphabet and Facebook has increased in recent years and often extends beyond company fundamentals due to the new ETF based market structure. One major issue is that the high correlation between Alphabet and Facebook could be drowned out by larger market signals in the tech market or even the economy as a whole. However, in the MLTS clustering framework these signals would be blocked within a sub-cluster within the technology sector. The principle of diversification would lead us to most likely choose one of these but not the other in constructing a portfolio based on the unique needs of a client.

25

This process leads to a procedure for the selection of assets by combining MLTS with a MACBETH based MCDA implementation: use MACBETH to determine appropriate securities, use MLTS to cluster the securities into blocks as a whole, and skim the top percentage off each block to generate potential assets to be considered in a screened portfolio.

One interesting theoretical implication of the new proposed technique is on the traditional risk-return paradigm that underlies financial theory. This approach extends the risk-return framework into a risk-return-suitability framework that incorporates the appropriateness of an asset in the selection process. Modern applied finance and regulations often focus not just on risk and return but whether or not an asset is suitable for a given investor. As such, even though theoretical results focus on risk and return financial practitioners often focus on suitability, appropriateness, and financial goals as a major concern in the selection of assets and as such this extension of the framework in many ways updates the theory of finance to match the actual practice and application of the discipline.

3.6 A Note on Suitability

The assets chosen through the MACBETH process extend the traditional mean-variance framework of returns into a broader context of appropriateness. Consider for example an individual working for a pharmaceutical company. This individual would already have too much exposure from his or her employment to the pharmaceutical and healthcare industry, and as such would not be best served by investing in this sector. If the individual is a younger investor he or she may also wish to have more growth exposure.

Through an MCDA framework, an individual wanting broad market exposure could get a basket of assets similar to the market but with strong growth prospects. The use of an MCDA framework creates an "intelligent" investing system that mimics the same processes and procedures a financial planner would use to engage with a client but in a cost effective, efficient system. The end result is a portfolio that can be characterized not just in terms of a two parameter specification (mean and variance), but in a high dimensional characterization (mean, variance, growth, dividend yield, strength of book, liquidity, among other factors as needed).

Many of these factors are implicitly understood to be used in traditional asset management, even if not explicitly stated. For example, it is understood that derivative contracts are often not discussed just in terms of risk and reward but also in terms of complexity, leverage, and asset insurance. This is similar to the reasoning why the Securities and Exchange Commission does not allow young investors in the United States with limited assets to invest in hedge funds, which would be the case if one only considered the desires over risk and return.

It is important to note that although portfolios will be discussed in terms of risk and return that these measures should be considered crude approximations rather than absolutes, as the portfolios generated by this combined framework are designed to address concerns extending beyond the two parameter characterization of investments and focus instead on a multi-parameter Pareto frontier.

3.7 Example of Portfolio Selection for a Value Based Investor

3.7.1 Selection Criteria

Suppose an investor had a long term investment horizon and was focused on having a stock portfolio weighted towards undervalued companies. For example, a venture capitalist may already have significant exposure to growth based start-ups and as such would want to weight his or her stock holdings more heavily in established companies with high dividend yields, attractive pricing and favorable book ratios. Suppose that through consultation, the analyst narrows down the client's MACBETH screening criteria to the dividend payout ratio, book to market, Shillers cyclically adjusted P/E ratio, dividend yield, enterprise value multiplier, price to cash flow, diluted price to operations earnings, price to sales and price to book. The MACBETH weights on these criteria is shown in Figure 3.2.

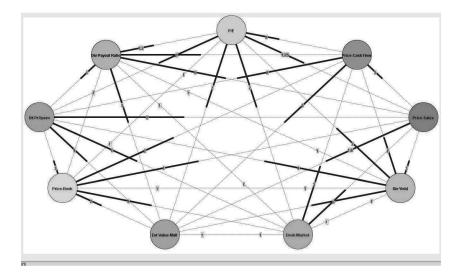


Figure 3.2: MACBETH criteria weights

The dividend payout ratio is the ratio of dividends paid out to shareholders relative to net income. From a valuation perspective, dividends paid out represent a real cash flow. As cash flows are discounted and as such present cash flows are more valuable than those in the future, firms that pay out high amounts of dividends could represent a relatively strong value pick. The main concern when discussing the payout ratio is whether the firm is stifling research and development and other growth options to finance present dividends, such as an oil company paying out higher dividends rather than investing in new drilling ventures. As such, dividend payouts and the health of a firm is of principal interest to investors and often a topic of research. It has been found that increased dividend payouts is correlated to higher real earnings growth in the future, but not necessarily higher real dividend growth (Gwilym, 2006). More recent studies provide consistent if different results, suggesting that firms with lower dividend payouts tend to be more volatile and under-perform, with this under-performance most noticeable among growth oriented firms (Connover, 2016). Similarly, the dividend yield is the is simply the ratio of dividend payouts to share price. Many of the arguments relating to dividend payout ratio are quite similar to dividend yield, although many investors seem to currently have a preference for the latter over the former. In recent years high dividend yield stocks

are increasingly of interest to institutional investors, increasing demand for them¹⁰ and raising the value of activity among insiders (You, 2017).

The Shillers cyclically adjusted P/E ratio is the ratio of the price to a 10 year moving average of real earnings. P/E ratios in general are considered one of the most important ratios in financial valuation and a standard component of a valuation framework (so much so that firms often attempt to manipulate or "cook the books" prior to raising capital in equity markets) (Chu, 2016). The Shillers P/E ratio, in particular, has been the workhorse of financial firms for many years. In particular, when used with consistent earnings data¹¹ the ratio produces consistent, reliable results as to the relative cost of companies and future real earnings returns (Siegel, 2016). In particular, the federal reserve is believed to track this ratio (or possibly other ratios that are highly correlated to it, based on empirical results) on a macroeconomic level to determine if markets appear to be "overheating" or forming asset bubbles (Hafner, 2017). Given the wealth of research on this topic and its importance, almost all portfolio managers use some sort of earnings multiplier in some part of a selection procedure¹². The enterprise value multiplier is a multiplier that is considered the merger and acquisition alternative to the P/E ratio, as it measures enterprise value relative to EBITDA. This value, while not often used in the valuation process for capital investors of particular note to merger and acquisitions specialists. In particular, it has had some success empirically in clustering firms within M&A circles and as such the authors feel its omission would be amiss (Asche & Misund, 2016). The authors of this paper often hear this ratio mentioned among the many investment bankers in the Tampa area who regularly use metrics based on enterprise value in their selection process.

¹⁰Increased demand for a high dividend yield stock increases its price and in turn reduces the relative yield on the stock over time. This paradigm is one of the reasons why maintaining a high relative dividend yield is far more difficult than dividend growth.

¹¹Earnings data based on GAAP is sometimes inconsistent, as GAAP are constantly being modified by FASB. In particular there has been a push recently to create additional similarities between GAAP and IFRS as more countries allow companies to issue annual reports using either standard. Inconsistent earnings data due to changes in GAAP have been shown to lead to some issues in estimation.

¹²ignoring purely momentum driven strategies

The price to cash flow ratio is, as it sounds, the ratio of a stock's price to the company's cash flows. Unlike earnings, some may argue that cash flows represent a more "honest" accounting of a firm's profitability as non-cash expenses are removed from the computation, with the figure in question coming from the company's statement of cash flows. Disruptions and unstable cash flows have recently been considered a red flag to investors, as firms who cannot provide stable cash flows are considered extremely risky to run into financial instability (Erickson et al., 2016). The price to sales ratio is similar, expressing a firm's stock price relative to earnings. When combined with the price relative to earnings, book, and cash flow this metric has been shown to be a strong predictor of equity returns (Shittu et al., 2016). As such, even though sales are often in and of themselves not as of interest as cash flow it does capture an important aspect of a firm's financial performance and is useful for clustering purposes. Put another way, a firm with declining or low sales may need to buffer their performance by cutting other areas such as research and development and as such a firm that performs poorly in this area but well in other areas may be "cutting corners" on long-term growth to bolster other ratios and as such may be perceived as being incongruous to the portfolio.

The final two metrics are the diluted price to operations earnings and price to book. This ratio is sometimes of interest in the case of large stock firms who issue exotic instruments with debt and equity like characteristics. This ratio simply assumes all possible convertible shares would be converted to stocks and is often similar to the earnings per share. Price to book is the price of a firm relative to its book value. The book value of a firm recognizes the historical cost principle from GAAP that states that firms acquired by a company must be recognized in accounting books at their historical price less depreciation rather than the market price. Because of this, firms trading under book are often thought to represent excellent value for an investor and as firms dip closer towards book are often targeted as strong buys. Because of this, firms trading close to book often have some issue that causes their valuation to be so low. In particular, this is often considered

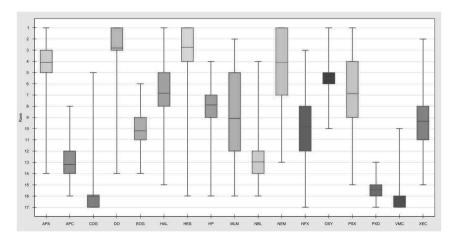


Figure 3.3: Box-and-Whisker plots of company utility

and has been in many cases to be a more conservative estimate of value than other valuation metrics (Nezlobin et al., 2016). Because of its conservatism this metric is often used in cases when firm's books may be considered suspect, prepared in accordance with less strict accounting standards, and or in markets with very high levels of uncertainty (Chen et al., 2014).

3.7.2 MCDA Process Example

Consider the previous 8 sectors, each with 4 sub-sector blocks. in the first block of sector 1 there are 19 members. By trimming each block to around the top 20% of companies, the space of potential companies can be reduced to just under 100, which can then be further optimized in terms of variance and return.

Figure 3.3 shows relative utilities for each company in the MCDA framework. Since the MACBETH process results in an additive model, it is typically the case that the mathematical calculation of robustness tends not to be a very computationally intensive process in most applications. For this particular example, DO, HES, APA, and OXY are the chosen tickers to represent this group as suitable candidates.

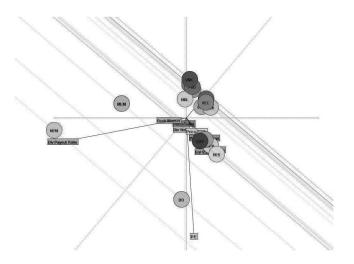


Figure 3.4: Two-Dimensional visualization

It is also possible to visualize the additive effects of each company and understand the relationship between the various criteria in the decision making process. Figure 3.4 shows that the selection of DO is heavily influenced by the Price to Earnings of the company.

Note that Figure 3.4 is not always a useful visualization for high degree problems, and in general caution should be taken when attempting to draw conclusions.

This process is continued for the remaining 31 blocks.

This leads to the selection of tickers given below.

3.7.3 Time Series Methods for Chosen Tickers

While each of these individual stocks may be appropriate, it is possible that every stock taken in aggregate could be inappropriate (locally optimal but not globally optimal). Because 95 tickers were chosen, it is reasonable to conclude that not all of these companies are necessary to construct a balanced portfolio.

Recall that the Sharpe ratio is a commonly used statistic that penalizes a portfolio's return in terms of its risk (Sharpe, 1966). The Sharpe ratio is given by

$$Sharpe_{Portfolio} = rac{r_{portfolio} - r_{free}}{\sigma_{portfolio}}$$

Sector-Block	Members	Number Selected	Sector-Block	Members	Number Selected
1-1	19	4	5-1	9	2
1-2	3	1	5-2	27	5
1-3	1	1	5-3	9	2
1-4	2	1	5-4	4	1
2-1	2	1	6-1	34	7
2-2	25	5	6-2	18	4
2-3	5	1	6-3	3	1
2-4	1	1	6-4	4	1
3-1	90	18	7-1	30	6
3-2	58	12	7-2	7	1
3-3	11	2	7-3	7	1
3-4	2	1	7-4	4	1
4-1	33	7	8-1	1	1
4-2	9	2	8-2	2	1
4-3	2	1	8-3	1	1
4-4	2	1	8-4	1	1

Table 3.2: Number of Companies Chosen from Each Block

Table 3.3: Chosen Tickers

DO	AEE	EXC	XOM	RAI	MUR	APD	NSC	NDAQ
HES	CNP	LLL	NOC	PFE	CVX	WHR	CMCSA	UNM
APA	D	MRK	MPC	SJM	INTC	PBI	NWSA	WM
OXY	FE	ADM	GILD	JNJ	NUE	HPQ	CSX	DRI
COP	PCG	GE	RTN	TSN	WDC	PH	FIS	RSG
DVN	DUK	TAP	HOG	MDLZ	DOW	MRO	VZ	UHS
RRC	ETR	GT	VLO	AMAT	CF	SPLS	СВ	DGX
KSS	CVS	TGT	PG	CAH	KR	BBY	TRV	LUK
GME	М	HD	LUV	TWX	Т	CA	AET	IBM
JPM	AFL	TNK	MTB	NTRS	PNC	PRU	AIZ	R
BHI	AA	GOOG	СХО	UDR				

where $r_{portfolio}$ is the return of a given portfolio, r_{free} is the risk free rate, and $\sigma_{portfolio}$ is the standard deviation of the portfolio. Typically the risk free rate is treated as known, but the authors for the purposes of this paper decided to treat all three parameters as unknown.

Three portfolios are generated: one using the Sharpe ratio as given ("Sharpe"), one dividing by ∂^2 instead of the normal estimator for sigma ("Sharpe Sensitive"), and one using the traditional Sharpe ratio but allowing for shorting ("Hedge Fund variant"). Shorting involves a trader borrowing a stock, selling it, and then returning the shares at a later date plus any dividends that were acquired. The result of a short is that an investor holds a negative equity position, which may not be suitable for all investors. In order to generate these portfolios, the authors first estimated the risk and return of each portfolio using ARIMA and GARCH methods including every possible stock in the candidate list. The authors considered every possible portfolio generated by dropping one stock, and chose the resulting portfolio that lead to the biggest increase in returns relative to risk. For the case of the hedge fund variant, the authors then saw if the risk-return profile would be improved by taking the stock in a short variation. Once the portfolio's risk-return structure could no longer be improved by dropping stocks, the authors attempted iteratively adding stocks back in that were previously dropped, and then resumed the process of attempting to drop stocks until a final allocation was reached.

3.7.4 Note on the Risk Free Rate and Riskless Assets in the Sharpe Computation

There are a variety of issues with the risk-free rate. Investors would first have to determine the appropriate investment time frame. Once this is determined, the analyst would have to determine whether to use the LIBOR rate, the federal discount rate, or the treasury rate for the corresponding asset. The LIBOR rate was originally the preferred rate, but has fallen out of favor after it was revealed that the LIBOR rate was being artificially set to maximize returns on derivative positions (mostly swaps) by British banks (FSA, 2013). It should be clear that one of the ideal properties of an estimator is that it should not be set artificially to maximize bank profits. The federal discount rate has fallen out of favor and is seen as powerful but in many ways symbolic (Kocherlakota, 2017). Central banks (including the Federal Reserve) have been increasingly reliant on quantitative easing and the unfreezing and freezing of bank reserves rather than on traditional rate-setting in recent years. The treasury rate also has issues. For one, it is moderately susceptible to the aforementioned quantitative easing policy of the Federal Reserve. Further, the Sharpe ratio was not designed for negative interest rates, which have been nominally positive but negative in real terms. Perhaps what's most striking is that Treasury bonds aren't technically considered risk free, as Standard and Poor's downgraded their rating to AA+ from AAA. One commonly used approach is to select the appropriate risk free rate arbitrarily after looking at different possible rates, which may be appropriate for some practitioners but is not a practice that would be difficult to justify statistically.

Bearing these ideas in mind and the fact that over the last several years the short term risk free rate has been essentially zero, a value of zero was chosen for this particular example. It could reasonably be argued that this is an unrealistic assumption of the riskfree rate, but one could also argue that almost any risk-free rate assumption is highly unreasonable in some way for some class of investor.

3.7.5 Comparison to Benchmark

The three different portfolios for a value driven investor are compared to the benchmark equity portfolio given by the SPDR S&P500 electronically traded fund (S&P500 ETF). The S&P 500 consists of 500 of the largest domestic large cap securities across multiple sectors. It is the most widely known ETF, and investments in the SPDR S&P500 ETF in many ways began the "ETF revolution" in finance. Its popularity is based on its broad exposure, low fee structure, and high liquidity, making it an excellent fit for most investors. The return on the S&P 500 is sometimes broadly referred to as the return on

Portfolio	Annual Comp Ret	Annualized Volatility	Alpha	Analytic VAR (5%)	Sharpe
Hedge Fund	11.68%	11.31%	2.96%	-4.28%	1.02
Sharpe	11.58%	12.76%	1.93%	-4.88%	.91
Sharpe Sensitive	11.56%	12.97%	1.77%	-5%	.89
Benchmark	9.86%	12.42%	0%	-5.05%	.8

Table 3.4: Comparison of Generated Portfolios to Benchmark

the "market." If no market is specified as the "market return," it is typically assumed to be either the S&P 500 or the Dow Jones Industrial Average. Because the stocks chosen are all large cap domestic U.S. equities, the S&P 500 is a natural comparison (and, as such, the authors chose the SPDR S&P 500 ETF). Table 3.4 shows the portfolios compared to the benchmark ETF on some traditional metrics. The data used to select the portfolios is from January 2012 to December 2014, with the performance results generated from the portfolio's performance from January 2015 to December 2017. The S&P 500 consists of 500 of the largest domestic large cap securities across multiple sectors. It is the most widely known ETF, and investments in the SPDR S&P500 ETF in many ways began the "ETF revolution" in finance. Its popularity is based on its broad exposure, low fee structure, and high liquidity, making it an excellent fit for most investors. The return on the S&P 500 is sometimes broadly referred to as the return on the "market." If no market is specified as the "market return," it is typically assumed to be either the S&P 500 or the Dow Jones Industrial Average. Because the stocks chosen are all large cap domestic U.S. equities, the S&P 500 is a natural comparison (and, as such, the authors chose the SPDR S&P 500 ETF). Table 3.4 shows the portfolios compared to the benchmark ETF on some traditional metrics. The data used to select the portfolios is from January 2007 to December 2014, with the performance results generated from the portfolio's performance from January 2015 to December 2017.

All three portfolios generate a significant annual return over the benchmark of around 2%. This is not surprising, as the stocks chosen for a value driven investor tend to have solid fundamentals, low Price to Earnings ratios, and strong dividend yields. The old mantra "but low P/E stocks" unsurprisingly tends to return strong performance when the companies have strong underlying fundamentals. It is also appealing that no re-balancing

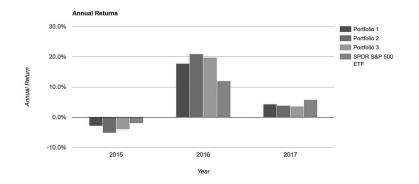


Figure 3.5: Performance over time

of the portfolio is required and that the returns were generated without the use of financial trickery such as reliance on small cap, low volatility stocks. These returns do not come at the expense of volatility, as the hedge fund variant has lower annualized volatility while the Share and Sharpe sensitive variants have volatility in the same neighborhood as the benchmark.

Another way of assessing risk other than estimated volatility is the variation at risk. This is the probability of the portfolio generating an annualized return of -5% or less. In other words, it's the likelihood an investor loses more than 5% of his or her investment given past returns. This takes the entire curve into consideration rather than just volatility. One of the more recent ideas in portfolio management is that selective management of risk can reduce VAR and thus reduce the investor's exposure to adverse events. The fact that all three portfolios beat the benchmark suggests that some of the higher volatility generated by the Sharpe and Sharpe sensitive portfolios are "noise" rather than uncertainty in the underlying fundamentals. This may offer some additional insurance against "jump" type events, but is not a guarantee of risk diversification.

All three portfolios perform very well on the estimated Sharpe ratio, which is not a surprise since this was a criteria used to generate them. The market is considered to be the S&P 500, which is why the benchmark returns an alpha of zero. It is not surprising that the hedge fund variant outperforms all the other portfolios. The ability to short means

that this portfolio is able to take market actions and positions that the other portfolios cannot adopt. Only a small percentage of the hedge fund portfolio is short, which means that the algorithm used shorts in a fairly conservative manner.

Shorting is not appropriate for all investors. Short positions require a margin account, which could lead to more expensive underwriting for investment firms. Individual traders engaging in short positions need to meet more stringent requirements set by FINRA and implemented by FINRA registered broker dealers. These tools are not meant for the average investor, and as such the authors of this paper suggest caution. As always, an individual should discuss the appropriateness of margin and other positions with a financial planner before engaging in such activities.

While much of the focus in asset management has been on return relative to risk and the Capital Asset Pricing Model, this chapter attempts to broaden the discussion. The generated portfolios not only generated strong financial results but are appropriate investment vehicles for the theoretical investor. They are composed of companies that meet his or her unique profile and should meet the individual's unique financial needs.

3.8 Contributions

In this section a new modular framework for constructing portfolios was developed. This framework extends the MLTS clustering framework to allow for a dynamic, high dimensional selection criteria and a high level of customization, addresses issues found in other asset allocation frameworks and embodies many of the traditional techniques currently being used in data analytics and modern portfolio theory.

This section shows some of the complexities and subtleties that often occur in the implementation of frameworks, in particular noting some of the decisions and choices that need to be made in data scrubbing and cleaning to allow for decisions to be drawn. It is also shown that MLTS clustering addresses many of the structural issues that limited MCDA implementations, while the MCDA framework (in particular MACBETH) allows

for a robust process for asset filtering extensions to MLTS clustering. It can be argued that together, these two methods generate results that are greater than the sum of their parts.

This chapter also notes the importance of reliable, customizable automated asset selection tools and the importance of expanding access to these tools to individuals from diverse financial backgrounds. Some of the issues with traditional asset management are discussed, with a discussion on the nature of the fiduciary responsibility in a high dimensional environment. The chapter notes that many of these extensions lead to non-trivial decisions that appeared trivial in the low-dimensional environment, such as the treatment of the risk-free rate.

It is shown that the new portfolios generated require little re-balancing, can contain many different kinds of stocks, and are still competitive with generic market portfolios. It is shown that there does not need to be an inherent trade-off between suitable investment strategies and returns for some classes of investors and that there is an argument for elevating the status of suitability of investments in the quantitative literature to the same level as risk and return.

4 Forecasting Single Asset Return in ETF Driven Markets: Revisiting the K-th Weighted ARIMA Model

4.1 Introduction

The previous chapter focused on modeling and developing a customized framework for a portfolio. While modern portfolio theory emphasizes the performance of a complete portfolio over individual securities, there is still a desire to understand the volatility of individual securities. Understanding individual security performance is often important in screening possible assets to include in a portfolio as well as understanding the asset specific risks and opportunities of an investment. Over the last several years there has been a structural shift in the markets towards Exchange-Traded Funds (ETFs). This shift in investing has opened up markets and allowed investors access to low cost, diversified investment options that previously were not available, while also leading to less oversight of individual companies.

The result of this structural shift towards ETFs is that it is now possible to gain broad exposure to sectors and areas of the market at a low cost, lowering the barriers to sector diversification. This has made strategies such as sector rotating more accessible to new classes of investors and significantly changed the investing framework. Since assets in major ETFs tend to move together, it also has created opportunities for exploiting investing opportunities in correlated movements of companies whose stock prices are correlated owing to their common inclusion in a major ETF but whose performance and underlying fundamentals are not. In order to do so, it is important to understand the performance and volatility of individual stocks over time. This has historically been done using an ARIMA based approach, sometimes extended into an ARIMA based econometric model. Shou Hsing Shih proposed an extension of this approach based on K-th weighted moving average filtering, which was designed to address structural issues in the securities market. It was shown this approach outperformed the standard ARIMA approach for most securities that were studied.

Since the introduction of the K-th weighted moving average ARIMA model, there has been a structural shift in As a result, it is important to re-evaluate the predictive nature of the approach and also investigate its usefulness at predicting the performance of asset classes besides individual stocks. This chapter investigates the applicability of the K-th weighted moving average approach in a changing financial market by applying it to six different highly liquid index funds. By doing so, it is possible to determine how the K-th weighted approach performs in current market environments. This chapter also sets the foundation for the empirical basis for the next chapter, which aims to extend the ideas from the K-th Weighted framework to GARCH volatility forecasts.

4.2 Market Sectors and Exchange Traded Funds

Today's economy is commonly classified into sectors in order to describe its various industries and stakes in a global perspective. The eleven sectors officially named by the Global Industry Classification Standard Framework (GICS) are Consumer Discretionary, Consumer Staples, Energy, Financials, Health Care, Industrials, Information Technology, Materials, Real Estate, Telecommunication Services, and Utilities (Kile & Phillips, 2009). These eleven sectors describe the current market and distinguishes one company from another. Investments are often made in various sectors in order to diversify a portfolio and take into account and balance the fluctuations throughout the stock market. Exchange-Traded Funds (ETFs) have become popular as they are able to represent several sectors of the economy as one asset in today's stock market.

This chapter will be analyzing six different ETFs: iShares MSCI Emerging Markets (EEM), iShares China Large-Cap (FXI), iShares Russell 2000 (IWM), SPDR Barclays Capital High Yield Bond (JNK), SPDR S&P 500 (SPY), and Financial Select Sector SPDR Fund (XLF). These six ETFs were selected because they are six of the highest daily trading volume ETFs and all represent different markets and segments.

The SPDR S&P 500 (SPY) was the first listed ETF in the United States, and directly corresponds to the Standard & Poor's 500 (S&P 500), based on 500 large corporations from various markets in the economy. In one asset, the SPDR S&P 500 is able to encapsulate the diversity of today's economy with stakes in all eleven GICS sectors, therefore serving as a desirable ETF (Fuhr, 2001). The iShares Russell 2000 ETF (IWM) contrasts from that of the SPDR S&P because of the importance it gives to smaller public United States companies, rather than larger corporations. It provides the diversification of all GICS sectors but focuses primarily on smaller companies rather than large caps (Hameed et al., 2017).

The iShares China Large-Cap ETF (FXI) gives investors a share in 50 large companies in China, having a similar purpose to SPY and IWM in terms of diversification, but specialized to give exposure to the Chinese market. Emerging markets have been a popular area of investment in recent years. Because of the tendency of China to outperform other emerging markets and the development of large multi-national conglomerates in China, the Chinese market has been characterized by highly active investments in recent years. There is an argument that ETFs do not have a good track record of reflecting the value of the underlying fundamentals, suggesting that traditional views about market efficiency and pricing them may not be correct. This happens more often and at a greater magnitude with funds that have foreign and illiquid assets, so the international ETFs, such as that of the IWM index, could be significantly different from the SPY ETF in prediction results (Petajisto, 2017).

As a combination of SPY, IWM, and FXI, the iShares MSCI Emerging Markets ETF (EEM) serves to provide diversification in the economy as well as internationally, provid-

ing investors shares in over 800 stocks in the emerging market at a global scale (Fapetu and Aluko, 2017). As opposed to SPY, FXI, EEM, and IWM, the Financial Select Sector SPDR Fund (XLF) is restricted primarily to the Financial Select Sector index, accounting for diversification within the financial sector.

ETFs specific to a certain sector are very responsive to the news of that sector. However, on the other hand, ETFs that cover multiple sectors, such as that of the SPDR Bloomberg Barclays High Yield Bond ETF (JNK), may not react to new changes in the credit spread as fast as expected, and could often be mispriced. This is, therefore, an ETF which is difficult to forecast, and represents a high-risk ETF and provides diversification specific to high-yield corporate bonds and, therefore, entails more risk for the investor with greater volatility (Bhorjaj et al., 2017). The JNK index is an intriguing addition that extends the K-th weighted literature to the fixed income market and a high risk asset class. Since the K-th weighted literature has primarily focused on vanilla assets, this makes for an intriguing application.

4.3 K-th Moving Average ARIMA Model

Many times, a time series requires a transformation in order to make it less volatile and therefore obtain better forecasting predictions. An attempt to do so is the k-th Moving Average ARIMA model, where the average of the "k" recent data points is taken as a moving average, and this data is then used to suit a traditional ARIMA model. Through this process, the original data becomes smoother, but at the same time, the original qualities of the data are retained, therefore serving two purposes: allowing the user to understand the trend of the original data and providing more precise forecasts for a set of data.

Various parameters were investigated for the moving average smoothing parameter, with a final smoothing parameter of five chosen. In order to choose a smoothing parameter the full modeling process is run with various choices of k, with the value five performing best in terms of AIC and fit. Because the ETFs chosen in this section and the next section are traded on U.S. exchanges, the choice of five as a parameter is a natural choice that mimics the structure of these exchanges which are open Monday to Friday and closed on weekends. These methods and the GARCH methods introduced in the next chapter are broadly applicable, and the user is in no way restricted to the use of 5-day smoothing. As such, other parameterizations such as 3 or 4 are acceptable options for some problems.

For data points $a_1, a_2, ..., a_5$ the 5th-average is defined as (Shih & Tsokos, 2007):

$$b_k = \frac{a_{k-4} + a_{k-3} + a_{k-2} + a_{k-1} + a_k}{5}$$

After taking the 5th-average for n - 4 data points an ARIMA model is fitted to the transformed data.

4.4 K-th Weighted Moving Average ARIMA Model

In modern markets it is more appropriate to give greater importance to the most recent data points rather than older ones, especially when making predictions. The k-th weighted moving average model accounts for this importance by weighing the most recent data point out of a set of "k" data points the greatest, the second most recent data point the next greatest, and so on with the k-th data point worth $\frac{1}{k}$ the value of the most recent observation.

For data points $a_1, a_2, ..., a_5$ the 5th-weighted average c_k is defined as:

$$c_k = \frac{a_{k-4} + 2a_{k-3} + 3a_{k-2} + 4a_{k-1} + 5a_k}{15}$$

Similar to the moving average, an ARIMA is then fit to the 5-th weighted moving average series to generate a prediction.

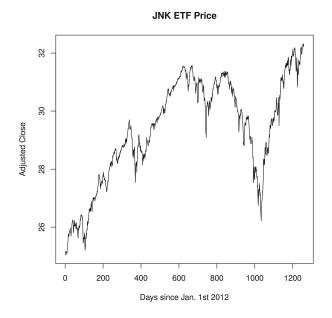


Figure 4.1: Daily adjusted JNK price

In order to make comparisons, models were fit on total return data¹ from January 2012 to December 2016, and used to make predictions from January to March 2017. Predictions are made one-step ahead at a time, using a daily updating process.

4.5 Example: Model for JNK

Daily Prices were recorded for the JNK index from January 2012 to December 2016. These returns can be seen in Figure 4.1.

The Augmented Dickey-Fuller test on the data produced a p-value of .55, which along with a graphical inspection suggests the series is not stationary in nature. Differencing the series once lead to a p-value below .01, which suggests the series is weakly stationary. The differenced series can be seen in Figure 4.2.

The new series appears stationary, and as such an ARIMA model with first order differencing is fit. The series has auto-correlation shown by Figure 4.3 and Partial Autocorrelation shown by Figure 4.4.

¹The total return is the capital gain adjusted for dividends, cash payments from acquisitions, and similar transactions.

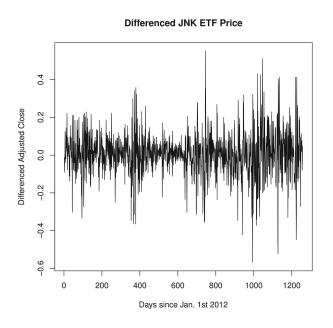


Figure 4.2: First order differenced JNK price

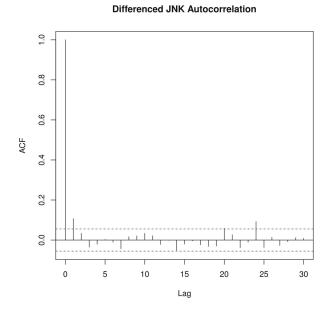


Figure 4.3: Auto-correlation of differenced JNK prices

Differenced JNK Partial Autocorrelation

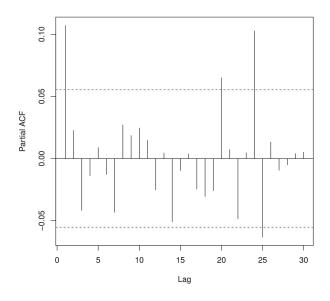


Figure 4.4: Partial Auto-correlation of differenced JNK prices

The auto-correlation is often used as a measure of the moving average component of the model and the auto-correlation as a measure of the auto-regression part of the model. After looking at the auto-correlation and partial auto-correlation the AIC and residuals are analyzed for different possible candidate models. After checking various parameterizations in a trial and error process a suitable model can be chosen that is parsimonious while also modeling the underlying behavior of the process. This trial and error process often requires some amount of familiarity with time series techniques and model fitting.

After trial and error, a first order autoregressive, second order moving average once differenced model was chosen.

This process leads to the model

 $\hat{r_{ndifference}} = -.1262r_{n-1_{difference}} - 9078r_{n-2_{difference}} + .0778r_{n-3_{difference}}$

$$+.2352\epsilon_{n-1_{difference}}+.9594\epsilon_{n-2_{difference}}$$

JNK 5-day MA ETF Price

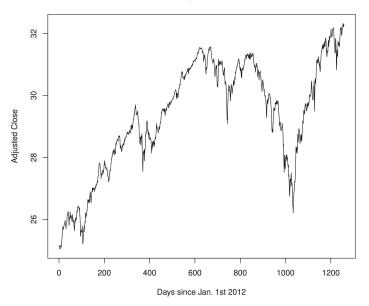


Figure 4.5: Daily adjusted fifth MA JNK price

with

$$\hat{r}_n = \hat{r}_{n_{difference}} - \hat{r}_{n-1_{difference}}.$$

This model produced an AIC of -1842.

This process is then repeated for the Fifth Order MA filtered data.

After differencing once, stationarity is achieved.

Using a similar process as with the unfiltered series, the final model of ARIMA(3,1,0) was chosen, with form given by

$$\hat{r}_{n_{differenceMA}} = .109 r_{n-1_{differenceMA}}$$

with an AIC of -1837.34. This corresponds to the final model given by

$$\hat{r}_{n_{differenceMA}} = .109 r_{n-1_{differenceMA}}$$

Differenced 5-day MA JNK ETF Price

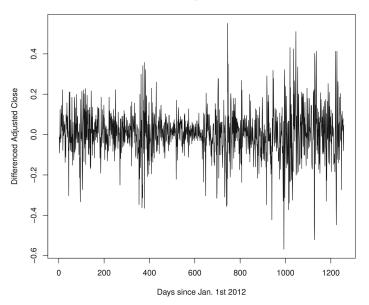


Figure 4.6: First order differenced fifth MA JNK price

$$\hat{r}_{n_{MA}} = \hat{r}_{n_{differenceMA}} - \hat{r}_{n-1_{differenceMA}}$$

$$\hat{r}_n = \frac{\hat{r}_{n_{MA}} + \hat{r}_{n-1_{MA}} + \hat{r}_{n-2_{MA}} + \hat{r}_{n-3_{MA}} + \hat{r}_{n-4_{MA}}}{5}.$$

This process is repeated for the Weighted MA filter.

Differencing once produces a stationary process.

An ARIMA(2,1,4) model was chosen, with form given by

$$\hat{r}_{n_{differenceWMA}} = -.1847 r_{n-1_{differenceWMA}} + .1661_{r_{n-2_{differenceWMA}}} + 1.0919_{\epsilon_{n-1_{differenceWMA}}} + .10919_{\epsilon_{n-1_{differenceWMA}}} + .10919_{\epsilon_{n-1_{differenc$$

$$+.7274_{\epsilon_{n-2_{differenceWMA}}} + .4512_{\epsilon_{n-3_{differenceWMA}}} + .2009_{\epsilon_{n-4_{differenceWMA}}}$$

with an AIC of -4571. This corresponds to the final model given by

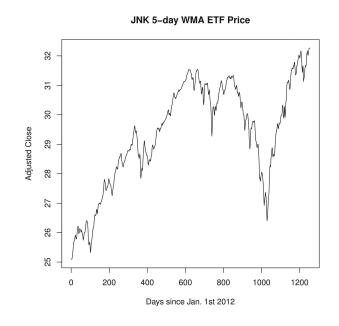


Figure 4.7: Daily weighted MA adjusted JNK price

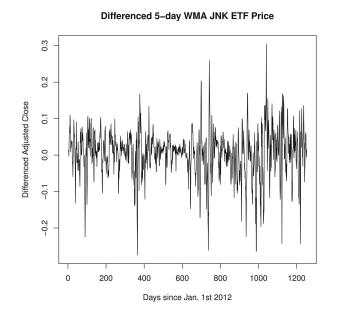


Figure 4.8: Daily differenced weighted MA adjusted JNK price

Table 4.1: EEM residuals

	ARIMA	kth MA	kth Weighted MA
\overline{r}_{EEM}	0.599	0.494	0.309
$V(r_{EEM})$	0.597	0.376	0.123

$$\hat{r}_{n_{differenceWMA}} = -.1847 r_{n-1_{differenceWMA}} + .1661 r_{n-2_{differenceWMA}} + 1.0919 \epsilon_{n-1_{differenceWMA}}$$

$$+.7274_{\epsilon_{n-2_{differenceWMA}}}+.4512_{\epsilon_{n-3_{differenceWMA}}}+.2009_{\epsilon_{n-4_{differenceWMA}}}$$

 $\hat{r}_{n_{MA}} = \hat{r}_{n_{differenceWMA}} - \hat{r}_{n-1_{differenceWMA}}$

$$\hat{r}_n = \frac{\hat{5r}_{n_{MA}} + 4\hat{r}_{n-1_{MA}} + 3\hat{r}_{n-2_{MA}} + 2\hat{r}_{n-3_{MA}} + \hat{r}_{n-4_{MA}}}{15}$$

This modeling process of determining the level of stationarity, choosing an appropriate mix of AR and MA parameters, converting the differenced predictions back to normal predictions, and inverting the filter was performed for all six ETFs. Every day, a one-stepahead forecast was produced of the next day's return. These returns were compared with the actual returns to form a series of residuals which are used to assess the accuracy of the model.

4.6 Model Performance

4.6.1 iShares MSCI Emerging Markets ETF (EEM)

Residuals are shown graphically and tabularly (mean and variance) below.

The k-th weighted moving average significantly outperformed the other two models. It appears that much of the gains from the model result from not just more accurate

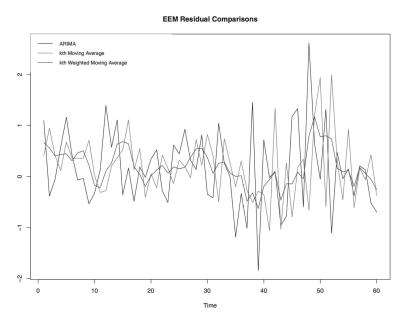


Figure 4.9: EEM residuals over time

Table 4.2: FXI residuals

	ARIMA	kth MA	kth Weighted MA
\bar{r}_{FXI}	0.606	0.510	0.302
$V(r_{FXI})$	0.610	0.394	0.130

predictions but also far less volatile predictions. The weighted average residuals appear much smoother and tend not to over-correct as much as the ARIMA and kth MA ARIMA alternatives, and is a significantly better predictor of performance in emerging markets.

4.6.2 iShares China Large-Cap ETF (FXI)

Residuals for FXI are shown below.

The FXI ETF exhibited similar trends to the EEM ETF. This is not surprising, as China's large role in EEM leads to a high level of correlation between these two indexes. In general, emerging markets tend to have similar exposures (such as debt linked to the dollar) and are highly sensitive to factors such as interest rates and slowdowns in the United States economy. Given the similar exposure of both of these ETFs have it is not surprising that the predictors showed similar results.

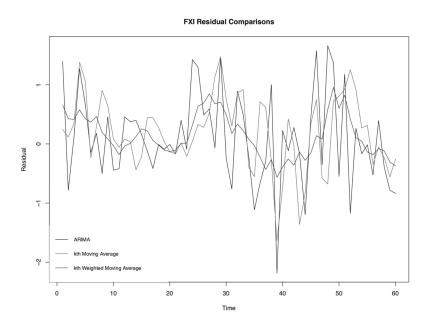


Figure 4.10: FXI residuals over time

Table 4.3: IWM residuals

	ARIMA	kth MA	kth Weighted MA
\bar{r}_{IWM}	0.689	0.807	0.284
$V(r_{IWM})$	0.772	0.952	0.117

4.6.3 iShares Russell 2000 ETF (IWM)

Residuals for IWM are shown below.

The Russell 2000 ETF consists of many small and mid-cap companies. This market segment is often believed to have higher levels of volatility and liquidity risk, and as such many of these companies on an individual level are considered inappropriate for unsophisticated investors. While IWM has less liquidity risk, much of the volatility of its members transfers over to the index.

the K-th moving average approach appears to have difficulty making predictions on these more volatile assets. The weighted MA approach delivers highly significant improvements in residuals against both approaches, suggesting that perhaps volatility re-



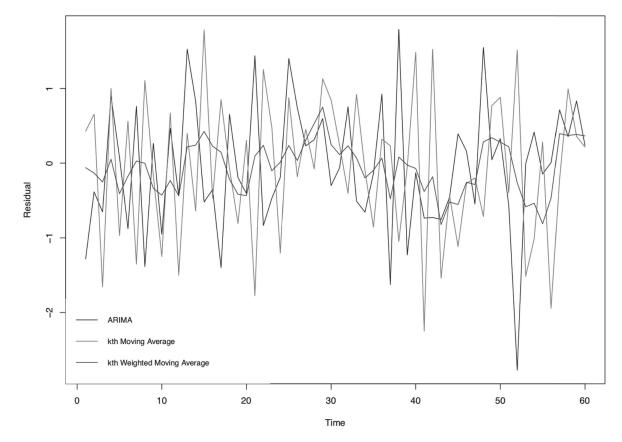


Figure 4.11: IWM residuals over time

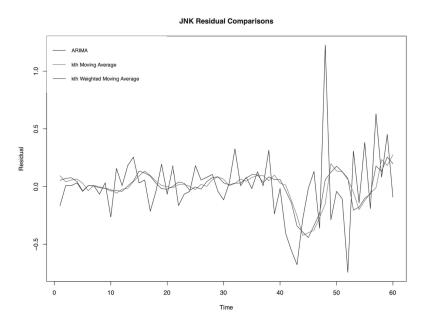


Figure 4.12: JNK residuals over time

Table 4.4: JNK residuals

		ARIMA	kth MA	kth Weighted MA
	\bar{r}_{JNK}	0.189	0.090	0.094
V	$r(r_{JNK})$	0.085	0.018	0.018

sulting from lack of liquidity could be creating issues with the 5-day moving average selection.

4.6.4 SPDR Bloomberg Barclays High Yield Bond ETF (JNK)

Results for JNK are shown below.

The aptly named JNK ETF consists of high yield dollar based corporate bonds. These bonds are typically not investment grade, and are often informally called "junk bonds." Fixed income investments are considered to be generally more resilient than equities. High levels of volatility in the bond market almost always occurs after high levels of volatility in stocks, and are usually thought of as lagging indicators during a downturn and leading indicators during a recovery.

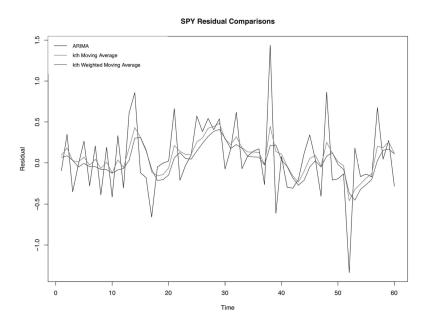


Figure 4.13: SPY residuals over time

Table 4.5: SPY residuals

	ARIMA	kth MA	kth Weighted MA
\bar{r}_{SPY}	0.325	0.169	0.144
$V(r_{SPY})$	0.188	0.038	0.033

The K-th moving average outperforms its weighted alternative by a slight margin. The relative performance of the two measures on this ETF appear very closely linked and similar across time. Both of these techniques vastly outperform the standard ARIMA alternative. Perhaps the relatively more steady behavior of the index during this time period explains the lack of performance improvement in the K-th weighted approach, as new information during this period was typically not severe enough to change the underlying likelihood of default of most companies in the high yield basket. Further research on JNK co-integrated with changing interest rate spreads may yield additional insights into this relationship.

4.6.5 SPDR S&P 500 ETF (SPY)

Residuals for SPY are shown below.

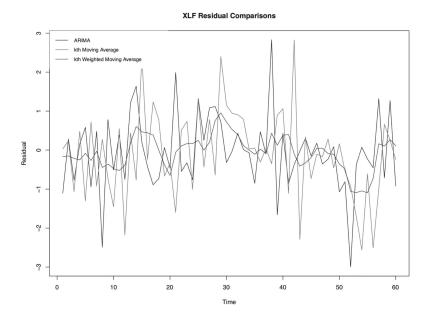


Figure 4.14: XLF residuals over time

Table 4.6: XLF residuals

	ARIMA	kth MA	kth Weighted MA
\bar{r}_{XLF}	0.717	0.863	0.333
$V(r_{XLF})$	0.942	1.24	0.195

The SPY ETF was the first ETF and is one of the most well known assets on any exchange. Offering exposure to every member of the S&P 500, the SPY ETF is often considered one of the most vanilla investments and is often used as an entry level investment, as a benchmark, as a shorting tool, and in countless ways. Being able to perform well on the SPY index is essential for any good predictor. As such, the fact that there is such a marked improvement in the new proposed approaches is a significant statement about the improvement in the quality of prediction.

4.6.6 Financial Select Sector SPDR ETF (XLF)

Residuals for XLF are shown below.

The weighting procedure produced significant improvements over both alternatives in the financial sector select ETF. One possible explanation for the poor performance of the unweighted MA is that for the last several years the financial sector has been a very information dependent market, and changes in information can quickly leave old information outdated. When new information is released such as an announcement of an interest rate change or new capital requirements, the weighting feature can greatly enhance the reliability of the prediction signal.

The strong performance of the K-th weighted approach suggests that the procedure can have applications to specific markets. The poor performance of the unweighted variant would be consistent with a theory that while the moving average approach is generally superior it may have issues in specialized markets or in individual sectors where time dependency is critical.

4.7 Conclusion

The K-th moving average outperformed the standard ARIMA method and produced superior or similar results to the unweighted forecast across all six ETFs. This suggests that during the time period for the study the K-th weighted approach produced reliable, consistent estimates. Visual inspection shows that this tended to hold across various periods in the study, including times of high and low volatility.

The proposed forecasting models utilize two unique factors that the traditional Autoregressive Integrated Moving Average (ARIMA) Model does not use. One, the model uses the idea that taking the average of a set of data points at a time smooths a time series model, thus allowing the ARIMA model performed on a series of averages to outperform the ARIMA model performed on a series of individual data points. This is the logic used by the k-th Moving Average ARIMA model. By, in other words, "smoothing out" a time series, the k-th Moving Average Model allows predictions to be made more accurately, thereby giving smaller and less variable residuals, on average. Second, the proposed model of the k-th Weighted Moving Average uses the idea that the most recent data point in the set of data points has a greater weight than previous data points when making short term predictions for a time series. The k-th Weighted Moving Average model then uses two unique factors in order to achieve more reliable results with greater precision. In using a smoother time series and weighing recent points greater, the k-th Weighted Moving Average drastically outperforms the traditional ARIMA model.

4.8 Contributions

ETFs are worth over \$4 trillion in assets in the current market. While the works of Hsing-Shih in the K-th moving average and K-th weighted moving average proved interesting, they had not been tested across asset classes and it was unknown if they were still relevant after the financial crisis.

By testing these new approaches across markets and asset classes, this chapter shows how the earlier work on filtered ARIMA processes has gained new relevance in recent years. In particular, it clarified the nature of the difference in performance by showing empirical results consistent with a theory that information sensitive sectors or models will experience a performance drop in the K-th MA approach but will have a significant expected performance improvement in the K-th weighted scenario. This chapter extends the understanding of time series filtering by showing that the results are consistent with an empirical theory that weighted MA processes realize significant gains in accuracy in environments where sudden changes that render old information unreliable. This added characterization on the nature of the improvement in this process adds to the understanding of possible mechanisms through which the improvement in the k-th weighted approach would theoretically generate the observed improvements which is consistent across all six ETFs. This theory on the nature of information reliability is a consistent extension of the K-th MA framework that offers a compelling explanation on the relationship between the three proposed models.

This chapter also marks the first empirical extension of the K-th filter approaches into the fixed income asset space. These results strongly support the notion that the K-th filter approaches have extensions to sectors beyond just stocks and may have important applications in other asset classes and in a time series econometrics framework.

The results from the six ETFs are highly suggestive that the K-th weighted approach is a reliable and consistent estimator across asset classes and should be strongly considered by those looking to characterize the nature of assets. This work further advanced the notion of the potential gains from smoothing bumpy or non-stationary time series data. Finally, this chapter sets the foundation for the next chapter on a new proposed GARCH method.

5 New Moving Average Approaches for Forecasting and Characterizing Volatility

5.1 Introduction

Volatility is the numerical measure of the range of values of the returns for a specific stock market index. Traditionally, volatility is used as a measure of risk in determining the relative response of a market index to changes in the economy. Thus, it often helps investors determine whether and at what time during the economic cycle to purchase or sell a specific market asset. Historically, volatility was calculated by looking at each quarter, taking the daily returns and computing the sample standard deviation of the return.

The ARCH process described in chapter two lead to a new method of estimating volatility that took spatio-temporal relationships into consideration. This was later generalized into GARCH, which has been the standard model for characterizing volatility in quantitative finance. This chapter aims to extend the work on modeling returns of ETFs in chapter 4 to modeling volatility, and proposes two new GARCH extensions: the K-th moving average GARCH model and the K-th weighted moving average GARCH model.

Nine different ETFs were used across various market segments to assess the reliability and effectiveness of the newly proposed models. After showing the new GARCH methods to be effective tools, this chapter explores an application of this filtering extension to a supervised volatility modeling environment on periodicity. In this section, empirical evidence suggests that the filtering process on the GARCH estimates generates the biggest improvements during times of economic expansion. Filtered and unfiltered estimators tend to perform poorly during times of economic contraction, and it is shown that the transformed data performs worse during this time period. This suggests that the modeling approach is particularly powerful during periods of growth and weak during periods of contraction. In particular, this suggests that the accuracy of the parameter estimates exhibits cycle dependency and that increasingly inaccurate volatility estimates are a sign of a financial correction.

The results from the newly proposed method of signal decomposition and from the model fits show that the GARCH approach is a useful technique that should be considered as an alternative or in conjunction with the spectral decomposition of a series. These approaches and GARCH in general has applications beyond finance and should be considered as a useful tool along with spectral decomposition across fields that use panel data.

sectionK-th Moving Average GARCH Model

As discussed in earlier chapters, the GARCH (Generalized Autoregressive Conditional Heteroskedasticity) Procedure analyzes a time series and measures for their variance by taking into account the variance of past data points (following an autoregressive format). GARCH differs from ARCH, becoming generalized as it takes into account the Autoregressive Moving Average Model (ARMA) to model the deviation of error. The GARCH(s, l) model denotes a model with s lags and l autoregressive terms.

The premise of the k-th Moving Average model is to reduce the fluctuations of the original plot of returns, therefore making it more suitable to be modeled. The problem of drastic fluctuations in the returns of a time series is solved by averaging groups of in a moving average. For a model of a data set $a_{1,}a_{2,...}, a_T$ A moving average model of length k uses the filter

$$j_t = \frac{a_{t-1} + a_{t-2} + \dots + a_{t-k}}{k}$$

to construct a new series $k_1, k_2, ..., k_{T-n-1}$. After fitting a GARCH model and obtaining an estimate $\hat{\sigma}_t^2$ this is inverted to form

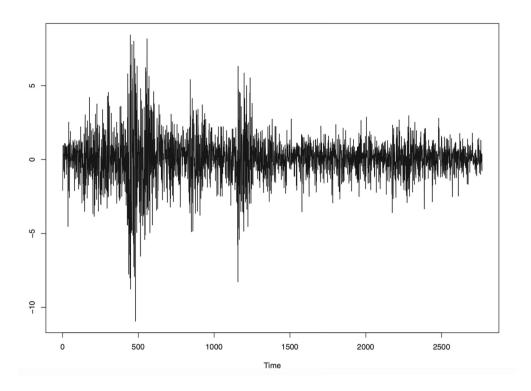


Figure 5.1: Time series of IWM returns; unmodified

$$\hat{\sigma_{a_t}^2} = (\hat{\sigma_{j_t}^2} + \sigma_{j_{t+1}}^2 + \sigma_{j_{t+2}}^2 + \dots + \sigma_{j_{t+k-1}}^2)/k.$$

Similar to the way that GARCH is considered the ARMA equivalent for variance, the K-th Moving Average GARCH model is the K-th Moving Average equivalency for variance.

For the IWM ticker that was discussed in the previous chapter, Figures 5.1 and 5.2 show the unfiltered and filtered series process for a 5-th order moving average. Similar to the previous chapter a fifth order moving average was chosen to model the weekly exchange structure.

As can be seen, not only does the amount of fluctuation in the data set significantly reduce, but the range of returns also reduces by about 25%. It is believed this is one of the mechanisms through which the improvement in estimation occurs.

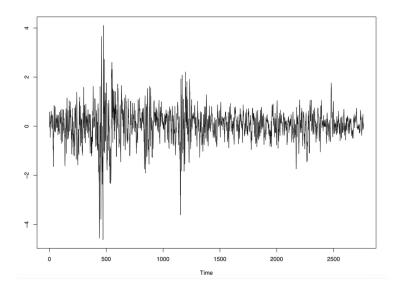


Figure 5.2: Time series of IWM returns; modified with K-th moving average

5.2 K-th Weighted Moving Average G ARCH Model

The k-th Weighted Moving Average takes into consideration what the traditional GARCH and ARMA procedures do not: added importance to more recent data points to form better overall models. In order to increase the importance given to more recent data points, for a given k the transformed weighted data point w_t the un-transformed data point a_t is assigned the value k, a_{t-1} is assigned the value k - 1 and so on with the final terminal value a_{t-k-1} assigned a value of 1. The general form of the transformed series for some n is

$$w_t = \frac{ka_t + (k-1)a_{t-1} + (k-2)a_{t-2} + \dots + a_{t-k-1}}{\sum_{i=1}^k i}$$

$$=\frac{ka_t + (k-1)a_{t-1} + (k-2)a_{t-2} + \dots + a_{t-k-1}}{(k(k+1))/2}$$

After obtaining a variance estimate $\hat{\sigma}_t^2$ for the filtered process, the data can be converted using

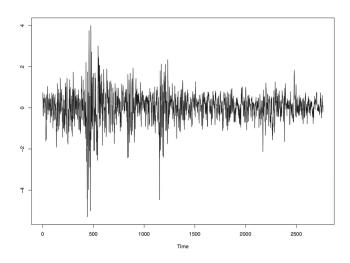


Figure 5.3: Time series of IWM returns; modified iwth K-th weighted moving average

$$\hat{\sigma_{a_t}^2} = 2(\hat{k}\sigma_{w_t}^2 + (k-1)\sigma_{w_{t+1}}^2 + (k-2)\sigma_{w_{t+2}}^2 + \dots + \sigma_{w_{t+k-1}}^2)/k(k+1).$$

This chapter will use k = 5 given by

$$w_t = \frac{5a_t + 4a_{t-1} + 3a_{t-2} + 2a_{t-3} + a_{t-4}}{15}.$$

Applying our 5-th Weighted MA to IWM leads to Figure 5.3.

The k-th weighted average not only reduces the range of data by about 25%, thus making it more suitable to form models upon, but places a heaver weight on more recent data points.

5.3 Exchange Traded Fund Selection

Nine different ETFs are used in the data collection process: SPDR S&P 500 (SPY), Financial Select Sector SPDR (XLF), iShares Russell 2000 (IWM), iShares Investment Grade Corporate Bond (LQD), iShares 1-3 Year Treasury Bond (SHY), iShares Emerging Markets (EEM), iShares China (FXI), Vanguard Total Bond Market Index Fund (VBMFX), and SPDR S&P Transportation (XTN). In addition to the five ETFs used in the previous chapter, four new ETFs are included (High yield bonds was dropped). These included ETFs representing investment grade bonds, treasuries, and a transportation and logistics ETF.

The new ETFs were selected to represent various market sectors as well as better understand the role of volatility in bond markets. As discussed earlier, during times of economic uncertainty before a correction stocks historically experience higher levels of volatility than bonds. After the correction, bonds typically show higher volatility before a recovery begins. As such, understanding both fixed income and equity markets is an important area of study in volatility estimation.

5.4 Data Collection

The closing price of SPDR S&P 500 (SPY), Financial Select Sector SPDR (XLF), iShares Russell 2000 (IWM), iShares Investment Grade Corporate Bond (LQD), iShares 1-3 Year Treasury Bond (SHY), iShares Emerging Markets (EEM), iShares China (FXI), Vanguard Total Bond Market Index Fund (VBMFX), and SPDR S&P Transportation (XTN) from January of 2007 to December of 2017 was collected in order to use for the GARCH procedures and respective volatility modeling. The total return on equity was computed (the stock price plus any cash distributions) for a given day n by

$$r_n=\frac{p_n-p_{n-1}}{p_{n-1}}.$$

These results were scaled by 100, which will not change the final model chosen by AIC. The set of returns was then modified with a 5-th moving average and a 5-th weighted moving average to produce the three models for each data set.

5.5 Example: SPY Ticker Model Fitting

This section will focus on the process of fitting the GARCH models, using the SPY data as an example. Once the data is converted into returns, the data is scaled by 100 (converted into a percent) then tested for stationarity.

For the unfiltered data, with a test statistic of -14.825 in the augmented Dickey-Fuller test and a p-value below .01 it is reasonable to conclude that an ARMA model is appropriate for mean return.

Using the Yule-Walker estimation method, an AR(2) model was found to be the most appropriate for the unfiltered data. This formed the model¹

$$\hat{r}_{n_{SPY}} = -.0907r_{n-1_{SPY}} - .0785r_{n-2_{SPY}} + .0391 + \epsilon_{n_{SPY}}$$

with $\epsilon_{n_{SPY}} \sim N(0, \sigma^2)$

This model had an AIC of 9096. A standard GARCH(1,1) model was then fit using the AR(2) with non-zero mean model. Using maximum likelihood estimation, this lead to a model parameterization of

$$\hat{\sigma}_{n_{SPY}}^2 = .024 + .124\epsilon_{n-1_{SPY}} + .857\sigma_{n-1_{SPY}}^2 + \epsilon_{n_{SPY}}$$

where $\epsilon_{n-1_{SPY}}$ corresponds to the aforementioned AR(2) model and $\sigma_{n-1_{SPY}}^2$ is computed via an exponentially weighted moving average. This lead to a model with an AIC of 2.53. Details of the model are in table 5.1.

For the 5-th MA filter model, the data was filtered with the fifth order moving average. Analyzing trends in the AIC across candidate models lead to the choice of a second order autoregressive and second order moving average model. This corresponds to the estimate

¹Note that the final error term is often not explicitly stated due to its zero mean.

Parameter	Estimate	SE
μ	.075	.013
MA(1)	0907	.021
MA(2)	0785	.021
ω	.024	.004
α	.124	.014
β	.857	.015

Table 5.1: Unfiltered GARCH model parameters

Table 5.2: Fifth Moving Average GARCH model fit

Parameter	Estimate	SE
μ	.053	.009
AR(1)	550	.021
AR(2)	.0426	.021
MA(1)	1.62	.001
MA(2)	1	.000
ω	.024	.000
α	.124	.017
β	.857	.017

$$\hat{r}_{n_{SPYMA}} = -.550r_{n-1_{SPYMA}} + .0426r_{n-2_{SPYMA}} + 1.62\epsilon_{n-1_{SPYMA}} + \epsilon_{n-2_{SPYMA}} + .0525 + \epsilon_{n_{SPYMA}} + .0525 + \epsilon_{n_{SPYMA$$

$$\hat{r}_{n_{SPY}} = \frac{\hat{r}_{n_{SPYMA}} + \hat{r}_{n-1_{SPYMA}} + \hat{r}_{n-2_{SPYMA}} + \hat{r}_{n-3_{SPYMA}} + \hat{r}_{n-4_{SPYMA}} + \hat{r}_{n-5_{SPYMA}}}{5}$$

and would involve inversion of the function as appropriate. For the GARCH(1,1) fit, this leads to

$$\hat{\sigma}^2_{n_{SPYMA}} = .001 + .158\epsilon_{n-1_{SPYMA}} + .828\sigma^2_{n-1_{SPYMA}} + \epsilon_{n_{SPYMA}}$$

with an AIC of -.255.

This can be extrapolated to form the final model given by the hierarchical model²

²Note that all three hierarchical stages are needed, as the daily volatility is a function of the previous return volatilities, which are themselves a function of the mean estimation errors.

$$\sigma_{n_{SPY}}^{2} = \frac{\hat{\sigma}_{n-1_{SPYMA}}^{2} + \hat{\sigma}_{n-2_{SPYMA}}^{2} + \hat{\sigma}_{n-3_{SPYMA}}^{2} + \hat{\sigma}_{n-4_{SPYMA}}^{2} + \hat{\sigma}_{n-5_{SPYMA}}^{2}}{5},$$

$$\hat{\sigma}_{n_{SPYMA}}^2 = .001 + .158\epsilon_{n-1_{SPYMA}} + .828\sigma_{n-1_{SPYMA}}^2 + \epsilon_{n_{SPYMA}},$$

$$\hat{r}_{n_{SPYMA}} = -.550r_{n-1_{SPYMA}} + .0426r_{n-2_{SPYMA}} + 1.62\epsilon_{n-1_{SPYMA}} + \epsilon_{n-2_{SPYMA}} + .0525 + \epsilon_{n_{SPYMA}}.$$

For the Fifth Order Weighted Moving Average, the data is filtered using a fifth order weighted moving average. A second order autoregressor and first order moving average model was fit to the weighted filtered data, leading to the model

$$\hat{r}_{n_{SPYWMA}} = 1.3r_{n-1_{SPYWMA}} - .488r_{n-2_{SPYWMA}} - .6\epsilon_{n-1_{SPYWMA}} + .0388$$

with an AIC of 3079. This corresponds to the model

$$\hat{r}_{n_{SPY}=\frac{5\hat{r}_{n_{SPYWMA}+4\hat{r}_{n-1_{SPYWMA}+3\hat{r}_{n-2_{SPYWMA}+2\hat{r}_{n-3_{SPYWMA}+\hat{r}_{n-5_{SPYWMA}}}{15}}{15}}$$

The GARCH(1,1) model fit to the corresponding weighted filtered series lead to the model

$$\hat{\sigma}_{n_{SPYMA}}^2 = .003 + .12\epsilon_{n-1_{SPYWMA}} + .862\sigma_{n-1_{SPYWMA}}^2$$

with an AIC of .349.

This corresponds to the final hierarchical model of

$$\sigma_{\hat{n}_{SPY}}^{2} = \frac{\hat{5\sigma}_{n-1_{SPYWMA}}^{2} + 4\hat{\sigma}_{n-2_{SPYWMA}}^{2} + \hat{3\sigma}_{n-3_{SPYWMA}}^{2} + \hat{2\sigma}_{n-4_{SPYWMA}}^{2} + \hat{\sigma}_{n-5_{SPYWMA}}^{2}}{15},$$

Parameter	Estimate	SE
μ	.003	.001
AR(1)	1.30	.006
AR(2)	488	.004
MA(1)	600	.006
ω	.003	.000
α	.120	.014
β	.862	.015

Table 5.3: Fifth Weighted Moving Average GARCH model fit

Table 5.4: AIC values for modeling approaches

Ticker	Traditional	Moving Average	Weighted MA
SPY	2.53	255	.349
XLF	3.22	.316	1.03
IWM	3.20	.060	1.01
LQD	.500	-2.18	-1.69
SHY	-2.83	-6.00	-4.99
EEM	3.41	.714	1.22
FXI	3.89	1.16	1.63
VBMFX	404	-2.95	-2.60
XTN	.710	-2.20	-1.49

$$\hat{\sigma}_{n_{SPYMA}}^{2} = .003 + .12\epsilon_{n-1_{SPYWMA}} + .862\sigma_{n-1_{SPYWMA}}^{2},$$

$$\hat{r}_{n_{SPYWMA}} = 1.3r_{n-1_{SPYWMA}} - .488r_{n-2_{SPYWMA}} - .6\epsilon_{n-1_{SPYWMA}} + .0388$$

This process was repeated for the remaining eight indexes, with the results discussed in the next section.

5.6 Results

The results of the modeling process are shown in the table below.

The standard Moving Average outperformed both the traditional and the Weighted MA approach in every model, with the weighted MA being the second best model in every case.

The MA approach's performance advantage is highly convincing, outperforming across market segments and across time and by a significant margin relative to the next closest models. This suggests the filtering process is a useful technique worthy of further study. In order to understand the nature of the improvement, this chapter proceeds with a discussion of periodic effects and model fit across time.

5.7 Periodicity Modeling with Supervised Models

The typical approach for analyzing periodicity in time series methods is via a spectral analysis. In this approach, a signal is decomposed using a Fast Fourier Transform into a combination of sine and cosine functions, which can help identify patterns and trends. This would be considered a kind of unsupervised approach, with no labeling or structure given before the decomposition.

This section of the chapter focuses on analyzing structural trends in a supervised environment to better understand the performance of volatility estimation in different business cycles. Data was collected on the difference between Moody's seasoned corporate bond yield on 10 year issues less the federal funds rate to form a measure of credit spread, known as the AAAFF. This is considered a measure of the spread between safe and low risk fixed income investments. The instrument tends to be relatively stable during times of economic certainty and tends to vary as uncertainty in future corporate performance develops.

The data over time was labeled before the analysis based on the period of the economic cycle, categorized as peaks and troughs based on GDP growth. The peaks analyzed were those during that of March 2001 and December 2007, while the troughs analyzed were those during that of November 2001 and June 2009. The returns of the AAAFF credit

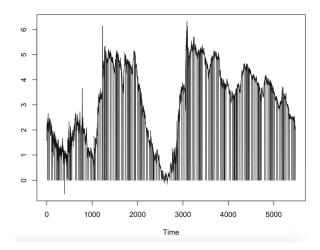


Figure 5.4: Time series of AAAFF prices

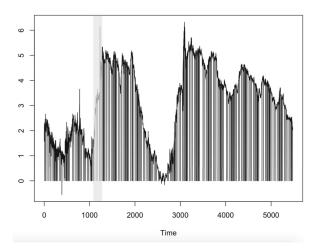


Figure 5.5: Region of AAAFF corresponding to March 2001 peak

spread were truncated to form data sets that match the following periods: January 2001-May 2001 (March 2001 Peak), September 2001-January 2002 (November 2001 Trough), October 2007-February 2008 (December 2007 Peak), and April 2009-August 2009 (June 2009 Trough).

The prices of AAAFF in the large data set are truncated to the periods listed above. Then, the prices in this region are plotted on graphs, and their data is modeled using the GARCH procedure. The AIC values are then compared amidst the regions to determine

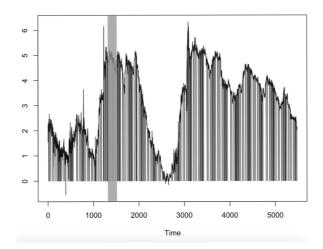


Figure 5.6: Region of AAAFF corresponding to November 2001 trough

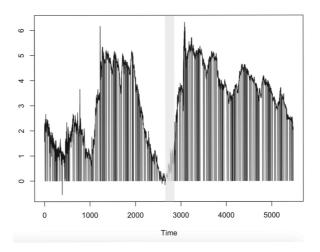


Figure 5.7: Region of AAAFF corresponding to December 2007 peak

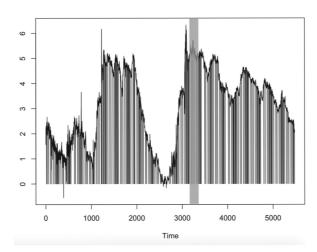


Figure 5.8: Region of AAAFF corresponding to June 2009 trough

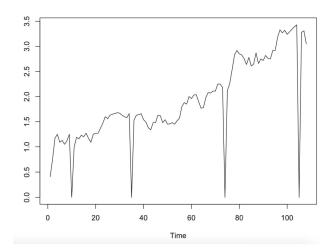


Figure 5.9: Price of AAAFF during March 2001 peak

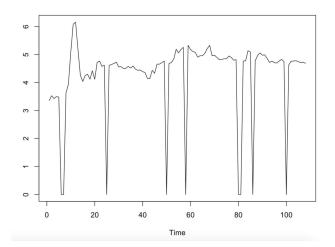


Figure 5.10: Price of AAAFF during November 2001 trough

any significant difference in AIC during economic periods of certainty versus those of uncertainty.

Under the standard GARCH model the AIC values during peaks are significantly lower than AIC values during troughs. This implies that during periods of economic uncertainty, the traditional GARCH estimator does not perform as well as during periods of economic certainty.

The analysis on volatility prediction was repeated for the 5-day moving average approach.

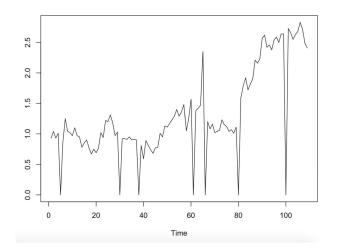


Figure 5.11: Price of AAAFF during December 2007 peak

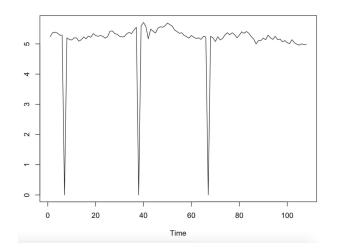


Figure 5.12: Price of AAAFF during June 2009 trough

Table 5.5: AIC values of AAAFF GARCH volatility model during different business cycle periods

	March 2001 peak	Trough Nov. 2001	Peak Dec. 2007	Trough June 2009
AIC	1.71	3.51	1.49	2.60

Table 5.6: AIC values of AAAFF 5-day moving average GARCH volatility model during different business cycle periods

	March 2001 peak	Trough Nov. 2001	Peak Dec. 2007	Trough June 2009
AIC	-1.33	3.59	1.36	2.63

While the 5-day approach tends to outperform the traditional approach, these gains occur during periods of financial stability and growth. The smoothing of the moving average appears to work best during relative stability, but the model performs worse during periods of economic uncertainty and during contractions. The traditional focus on spectral decomposition and variance often focuses on the nature of variance over time, but not on the nature of the prediction of variance over time. The GARCH process of decomposing volatility appears to have issues during troughs, and as such it appears that the likelihood representation of models and relative performance exhibits non-spectral state dependency.

This has broad implications to the understanding of markets and cycle rotation. This suggests that the information contained in the market regarding volatility is highest during periods of economic growth. Uncertainty peaks at the trough in the economic cycle. This suggests that contrary to traditional theories on uncertainty and risk that peak uncertainty on volatility is actually a sign that the market is due for possible expansion and that low volatility estimates are a suggesting of market complacency.

5.8 Theoretical Implications of Volatility Periodicity Estimation

Consider the case of an individual choosing to invest and makes a decision to allocate a certain amount of money towards bonds. In this simplified model, assume that the investor is not an accredited investor and is discouraged from investing in bonds with moderate to high risk of default. Then, this investor would have the choice of investing in either investment grade corporate bonds, municipal bonds, or corporate bonds.

Since the individual is not an accredited investor, it is reasonable to assume that municipal bonds would not be an appropriate investment choice and thus the investor makes a choice to invest in low risk corporate bonds or United States treasury bonds. Let us denote corporate bonds by C and treasury bonds by T.

Consider the case if a treasury bond and a corporate bond of the same maturity had equivalent yields. The treasury bond is backed by the full faith of the United States federal government and would be considered less risky. If the market for C and T is made of many investors identical to the given investor, then in order to have

 $C \sim T$

then it must be the case that

Yield(C) > Yield(T)

based on the level of risk, and that

$$1 > R = \frac{Yield(T)}{Yield(C)} > 0$$

with

R'(D) < 0

where D denotes the probability of a corporate default on debt for low yield corporate paper. In other words, when the probability of default is quite high then the yield of corporate bonds must also be quite high relative to treasuries, so R must be closer to 0.

Consider the level of risk of default of a corporate bond in an economy that is growing and continues to grow. The company choose to invest in projects, and since the success of the projects is highly correlated to performance in the economy it is likely that the probability of financial difficulty is low and the probability of default is also relatively low. Using S to denote the state of the economy, then with perfect information then this would suggest that

$$D|Growth < D|Correction \Longrightarrow (S_i(Growth) \rightarrow S_{i+1}(Correction) \Rightarrow R \uparrow)$$

which is to say all else being equal, for some state i transitions from growth to correction should lead to an increase in R, with similar implications for movements from correction to growth states.

If it is the case that:

- 1. Market participants have unbiased estimators of D and have full information on transition probabilities.
- 2. Markets are fully rational in the pricing of R.
- 3. As the transition happens more information becomes available investors have more information regarding the probability of being in a transition.
- 4. R is more sensitive to changes from growth to correction than correction to growth.

Then we would expect that:

- 1. R would exhibit more volatility in $S_i(Growth) \rightarrow S_{i+1}(Correction)$ than in $S_i(Correction) \rightarrow S_{i+1}(Growth)$
- 2. The moving average approach of smoothing noise should perform well in $S(Growth) \rightarrow S(Growth)$ and $S_i(Correction) \rightarrow S_{i+1}(Correction)$ relative to the standard approach.
- 3. Because the market participants have full information on the likelihood of transition then the GARCH estimation process of volatility in the standard case should not perform better in one state or another.
- 4. The moving average process will do worse in the transition as it smooths vital information that reveals a market change is occurring.

Empirically it was in this chapter that that (1), (2), and (4) held true. However, (3) did not. That is to say, the estimated models actually perform worse during all transitions in both cases. This is not consistent with a theory that there is accurate information on transition probabilities between states. That is to say, the empirical results observed suggests that the 3 by 3 transition matrix between Market states M of growth, constant, and correction is such that $\hat{S}_i(M) - S_i(M) \neq 0_{3,3}$.

This suggests that either investors are not rationally pricing the trade-off between bonds and treasuries, that market participants are not learning more over time about the likelihood of being in a given state, or that the estimation of D as a function of states is not unbiased to produce such bias in the estimation.

The most likely suggestion from this is that market participants are not properly updating default likelihoods as fast as they should in regards to state dependency, which is consistent with prospect theory predictions about investors not updating conditional beliefs fast enough (Tversky & Kahneman, 2013). The notion that GARCH estimation errors peak in both estimation procedures in peaks suggest that one possible leading indicator of a market correction is investors underestimating volatility and a sign of a market improvement is investors overestimating volatility.

This mismatch in volatility estimation suggests that market participants are too bullish during growth periods and too bearish during corrections. This empirical work is a consistent time dependent characterization of Tversky and Kahneman's work on prospect theory and is not consistent with a theory of rational market participation.

5.9 Contributions

This chapter proposed two new methods for volatility estimation that extends the GARCH framework with new filtering techniques. These new techniques were shown to outperform traditional GARCH measures across nine different ETFs. This chapter also proposed a new alternative technique to decomposing a time series that does not use a Fourier transform that is more appropriate for data structured for use in a supervised learning environment. This consistent improvement in estimation across markets

suggests that the earlier filtering process designed for estimating returns has important applications in structuring and estimating volatility as well and extends the work on estimation of volatility beyond the spectral approach into a GARCH framework.

Using this new approach on data labeled by market conditions, this chapter was able to demonstrate the structural nature of the improvement in GARCH estimation, which suggests that the structure of the estimation process under smoothing performs better in processes that are relatively stable and poorly in processes where more recent information dependency is vitally important. These results were contextualized to show that the empirical results are not consistent with rational market participants with full information and were more in line with predictions from prospect theory.

This chapter's results suggest that the new moving average framework is a powerful innovation for prediction on processes with relative stability over time, suggesting that a series with a relatively flat spectral representation would benefit in estimation from a smoothing process, while processes with bumpy spectral forms may have more accurate representation with standard GARCH techniques.

This chapter extends the work on estimation from the previous chapter to the second moment of a function, which is a vital step in establishing a risk-reward paradigm for an individual series. While this technique has classical applications to finance, it is a broad algorithm that forms a generic framework with applicability to problems beyond financial systems.

6 Future Research

6.1 Applications of the Combined MLTS-MACBETH Approach Across Disciplines

While earlier chapters focused on the financial applications of the MLTS-MACBETH combined approach, it should be noted that the algorithms are quite broad in construction and can be used across disciplines. In particular, one area of note is in design of roll-outs in differences-in-differences experiments. For example, if a group was considering doing a roll-out in multiple phases and coordinating the roll-out with a developmental study combined with a differences-in-differences OLS model the researchers could group cities using the MLTS approach. Once cities are grouped and blocked, the investors could use a MACBETH algorithm on potential confounding variables to sort and assess the factors that are believed to be the biggest econometric issues. Once this process is completed, this would lead to an approach to determine the optimal roll-out patterns to assess econometric differences.

The approach could also have similar applications in clinical drug trials, where individuals could be clustered and then assigned to treatment and control groups. If data on subjects do not have a panel data structure than alternative clustering algorithms could be combined with the MACBETH algorithm for inferences.

6.2 Nonlinear Multi-Criteria Decision Analysis Topological Characterizations of Preferences

One area of interest is an expansion from the MACBETH based MCDA algorithm to a nonlinear topological characterization of preferences. Since the approach developed was modular in nature, the MACBETH approach can be replaced with another decision making tool if so desired. Recall that MACBETH focuses on linear preferences and does not assume a preferential structure that could be additive, non-linear or exhibit conditionality. Rather than use a MACBETH approach, a non-linear method would characterize and assess preferences. This would involve an extension of the MCDA framework into a topological Pareto frontier.

6.3 Theoretical Exploration of Non-Transitive Preferences and its Implications on Selections

One area of future research is the exploration of preference systems that do not exhibit transitivity. Non-transitive preferences is a complex area of price theory that does not assume a transitive preference structure. Rather, preferences exhibit a generic topological structure.

Under a non-transitive preference relation, the typical solution would be a topological characterization of the solution space. Further research could focus on stability and convergence conditions, Classes of Nash Equilibrium conditions with multiple players all with non-transitive beliefs, and if there are mathematical conditions or criteria that would allow for convergence.

A typical example of non-transitive preferences is decision paralysis, where an individual faces so many decision options that he or she is unable to make a decision, possibly due to non-transitive rankings. This could further extend the work on preferences, beliefs, and decisions with ideas from behavioral economics such as irrational decision structures, poorly defined expectations and additional insights from prospect theory.

6.4 Applications of Moving Average Procedures to Various GARCH Class Models

The new GARCH weighting procedure is a broad approach that is compatible with approaches beyond the standard GARCH technique. Further research could investigate how well this method performs with various other GARCH formations beyond the vanilla GARCH technique.

In addition to extending this work to additional GARCH techniques, additional work could focus on ways to combine the spectral decomposition with data from a labeled environment to create a semi-supervised approach to data modeling in a time series environment. This work would continue the philosophy purported across this dissertation of combining time series methods with machine learning philosophies and ideas to create new intelligent systems for panel data problems.

6.5 Combination of a Decision Making Environment for Panel Data with a Bayesian Hierarchical Framework

The techniques developed in the construction of a portfolio are quite broad and modular in nature. In particular, if an individual has beliefs about returns or possible states this information could be expressed as a prior distribution. Future work could investigate methods to create computationally feasible algorithms, approaches, and implementations that structure a Bayesian parameter estimation framework over time.

These approaches could develop customized Markov Chain Monte Carlo algorithms or other approaches for application in a high dimensional environment and utilize cloud computing for parameter estimation. As GPU computing resources continue to grow and develop, higher dimensional computing problems such as a Bayesian addition to the asset allocation framework will increasingly become more cost effective options.

6.6 Determination of Optimal Rebalancing Periods

Every so often as preferences change or asset weights naturally drift the target allocation will become unbalanced and need re-balancing. One of the issues with re-balancing is that it necessitates exchange transactions that come at a cost. As such, there is a natural trade-off between re-balancing and maintaining a given allocation. Future work on re-balancing over time could use a time dependent frontier that combines preferences over allocation with future beliefs and predictions about how this imbalance between desired and actual allocation may change over time. This work could incorporate and investigate marginal preferences with a stochastic differential representation to extend the framework developed earlier in this dissertation and help determine when re-balancing should be used and to what magnitude assets need to be reconfigured. In addition, future work could explore in low liquidity environments optimal approaches for allocating and performing allocations with liquidity concerns over time.

7 References

Frank Asche and Bard Misund. Who's a major? a novel approach to peer group selection: Empirical evidence from oil and gas companies. Cogent Economics and Finance, 2016.

Carlos Bana e Costa, Jean-Marie Corte, and Jean-Claude Vansnick. On the mathematical foundation of macbeth. Multiple Criteria Decision Analysis: state of the art surveys, 2005.

Tim Bollerslev. Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics, 1986.

Karl Borch. Eects on demand of changes in the distribution of income. Econometrica, 1953.

Rongda Chen, Sheng Ye, , and Xianchao Huang. Risk measure and earlywarning system of china's stock market based on price-earnings ratio and price-to-book ratio. Mathematical Problems in Engineering, 2014.

Jenny Chu, Patricia M. Dechow, Kai Wai Hui, and Annika Yu Wang. The valuation premium for a string of positive earnings surprises: The role of earnings manipulation. Working Paper, 2016.

Chin-Sheng Huang & Jiang-Chuan Huang Chun-Fan You. Abnormal dividend-yield returns and investment strategy. Emerging Markets Finance and Trade, 2017.

Engle, Robert. "GARCH 101: The use of ARCH/GARCH models in applied econometrics." Journal of economic perspectives 15.4 (2001): 157-168.

Robert F Engle. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom ination. Econometrica, 1982.

Devon Erickson, Max Hewitt, and Laureen A. Maines. Do investors perceive low risk when earnings are smooth relative to the volatility of operating cash ows? discerning opportunity and incentive to report smooth earnings. The Accounting Review, 2016.

Kenneth R. Fama, Eugene F.; French. Common risk factors in the returns on stocks and bonds. Journal of Financial Economics, 1993.

M.S. Feldstein. Mean-variance analysis in the theory of liquidity preference and portfolio selection. The Review of Economic Studies, 1969.

Deborah Fuhr. Exchange-traded funds: A primer. Journal of Asset Man- agement, 2001.

C. Granger. Some properties of time series data and their use in econometric model specication. Journal of Econometrics, 1981.

C. Granger and A. Andersen. An Introduction to Bi-Linear Time Series Models. Vandenhoeck and Ruprecht, 1978.

Salvatore Greco, Matthias Ehrgott, and Jose Rui Figueira. Multiple criteria decision analysis: state of the art surveys. Springer, 2016.

Christian M. Hafner and Alexandre R. Lauwers. An augmented taylor rule for the federal reserve's response to asset prices. International Journal of Computational Economics and Econometrics, 2017.

Nils H. Hakansson. Multi-period mean-variance analysis: Toward a general theory of portfolio choice. The Journal of Finance, 1971.

Peter Hansen and Asger Lunde. A forecast comparison of volatility models: does anything beat a garch (1, 1). Journal of applied econometrics, 2005.

Anil K. Jain. Data clustering: 50 years beyond k-means. Pattern recognition letters, 2010.

Daniel Kahneman and Amos Tversky. Handbook of the fundamentals of nancial decision making: Part I, chapter Prospect theory: An analysis of decision under risk, pages 99127. World Scientic, 2013. Charles O. Kile and Mary E. Phillips. Using industry classication codes to sample high-technology: Analysis and recommendations. Journal of Accounting, Auditing & Finance, 2009.

Doo Young Kim and Chris P. Tsokos. Active and dynamic approaches for clustering time dependent information: Lag target time series clustering and multi-factor time series clustering. Journal of Statistical Theory and Applications, 2018.

Narayana R. Kocherlakota. The decentralized central bank: A review essay on the power and independence of the federal reserve by peter conti-brown. Journal of Economic Literature, 2017.

Mark H. Lai, Tsong-Yue; Stohs. Yes, capm is dead. International Journal of Business, 2015.

Harry Markowitz. Portfolio selection. The Journal of Finance, 1952.

Harry Markowitz. Portfolio Selection: Ecient Diversication of Invest- ments. John Wiley, 1959.

Doo Young Kim Michael Kotarinos and Chris P. Tsokos. Multi-level time series clustering: Issues with traditional risk management frameworks. JSM Proceedings, 2019.

J. Von Neumann and O. Morgenstern. Theory of Games and Economic Behavior, 3rd ed. Princeton University Press, 1953.

Alexander Nezlobin, Madhav V. Rajan, and Stefan Reichelstein. Structural properties of the price-to-earnings and price-to-book ratios. Review of Accounting Studies, 2016.

Executive Oce of the President. 2017 North American Industry Classification System Manual. Oce of Management and Budget, 2017.

Antti Petajisto. "ineciencies in the pricing of exchange-traded funds. Financial Analysts Journal, 2017.

R. Roll. A critique of the asset pricing theory's tests. Journal of Financial Economics, 1977.

W.F. Sharpe. Capital asset prices: A theory of market equilibrium under conditions of risk. Journal of Finance, 1964.

William F. Sharpe. Mutual fund performance. The Journal of Business, 1966.

Shou Hsing Shih and Chris P. Tsokos. A weighted moving average approach for forecasting. Journal of Modern Applied Statistical Methods, 2007.

Isah Shittu, Ayoib Che Ahmad, and Zuaini Ishak. Equity valuation multiples and stock returns: Evidence from malaysian listed rms. Working Paper, 2016.

Jeremy J. Siegel. The shiller cape ratio: A new look. Financial Analysts Journal, 2016.

Nassim Nicholas Taleb. The Black Swan: The Impact of the Highly Im- probable. Allen Lane, 2007.

X. et al. Wang. Characteristic-based clustering fortime series data. Data mining and knowledge Discovery, 2006.

Jr. Ward, J. H. Hierarchical grouping to optimize an objective function. Journal of the American Statistical Association, 1963.

Y. Xiong and D. Yeung. Time series clustering with arma mixtures. Pattern Recognition, 2004. 80

Appendix A Brief Definitions of Selected Terms

Accredited Investor: A term the Securities and Exchange Commission uses for high net worth investors, financial institutions and large companies.

Alpha (Finance): A measure of the performance of an asset relative to a benchmark.

Augmented Dickey-Fuller Test: A statistical unit root test for a time series.

Autocovariance: A measure of the volatility of a process over time.

Autoregression: A time series approach that uses past values to predict future estimates.

Autoregressive Conditional Heteroskedasticity (ARCH): A time series technique for estimating volatility or variance over time.

Autoregressive Integrated Moving Average (ARIMA): A time series technique for predicting a value over time. Generalization of ARMA.

Autoregressive Moving Average (ARMA): A time series technique for predicting a value over time.

Bayesian Statistics: A branch of statistics that allows for probabilistic representations to change over time rather than being fixed.

Black Swan Events: low probability high impact events that are difficult to predict and analyze using current data.

Book to Market: The ratio of a company's **book value** to its market value.

Book Value: The value of a company based on its assets and liabilities from financial statements.

Broker-Dealer: A financial institution that engages in the buying and selling of securities either for its own account or on behalf of a client.

Capital Asset Pricing Model (CAPM): A model for pricing an asset in terms of market conditions and its underlying risk.

Cardinal Utility Scale: A scaling process of utility that preserves preference orderings up to an affine transformation.

Central Processing Unit (CPU): A computer hardware component that performs logical operations.

Cloud Computing: The use of web linked computing servers on an as-needed basis.

Clustering: Breaking observations in a data set into groups based on certain characteristics.

Committee on Uniform Security Identification Procedures (CUSIP): A nine digit code used to identify a security for the purpose of data consistency and clearing transactions.

Credit Spread: The difference in yield between two bonds of equivalent maturity but with different levels of default risk.

Decision Maker (DM): In decision theory, an individual making a choice or decision.

Derivative (Finance): An asset that derives its value based on another asset.

Diluted Price to Operations Earnings: The ratio of a company's price to its **operations earnings** after adjusting for **dilution**.

Dilution: The reduction in existing shareholders' ownership through the issuance of new equity.

Dividend: A cash distribution paid to shareholders by a company.

Dividend Payout Ratio: A ratio of net income paid out as **dividends** to shareholders.

Dividend Yield: A measure of the amount of **dividends** paid out relative to a stock's price.

Dominated Portfolio: See Stochastic Dominance.

Dow Jones Industrial Average: An index of 30 of the largest multinational United States based companies in the world.

Earnings Before Interest, Taxes, Depreciation, and Amortization (EBITDA): A measure of a company's pre-tax performance before non-cash expenses.

Econometrics: A branch of empirical economics that uses mathematical and statistical theories in an economic context to draw conclusions.

Enterprise Value Multiple: The value a company is trading at relative to its **EBITDA**. Similar to the **P/E ratio** but uses EBITDA in place of earnings.

Exchange: A regulated financial market for the sale of financial products.

Exchange Traded Fund (ETF): A marketable security that tracks an asset class and trades on exchanges.

Exponential Generalized Autoregressive Conditional Heteroskedasticity (EGARCH): An alternative to the standard **GARCH** model using logged values.

Fast Fourier Transform: A method for decomposing a signal into a combination of Sine and Cosine processes to view trends and patterns.

Federal Discount Rate: The amount the Federal Reserve charges member banks on loans to maintain reserve requirements.

Financial Industry Regulatory Authority (FINRA): A non-profit regulatory and licensing agency that oversees the activities of broker dealers, investment advisers, and other financial professionals.

Financial Institution: Corporations that serve as intermediaries and participants in financial markets.

Free Cash Flow: The cash a company produces after paying for operating expenses and capital expenditures.

Game Theory: A branch of economics that uses mathematical techniques to understand the interactions between participants and the relationship between participant choices and equilibrium conditions.

Generalized Autoregressive Conditional Heteroskedasticity (GARCH): A time series technique for estimating volatility or variance over time. Generalization of **ARCH** process.

Generally Accepted Accounting Principles (GAAP): The accounting standard adopted for use in the United States.

Global Industry Classification Standard (GICS): A framework for classifying companies into sectors based on economic activity.

Gross Domestic Product (GDP): An aggregate measure of all the activity in an economy. If not otherwise specified typically assumed to be referring to the United States.

Graphics Processing Unit (GPU): A computer hardware component designed for the processing of graphical images. GPUs have recently seen numerous scientific computing applications with applications to machine learning and matrix computations.

Hedge Fund: A **mutual fund** that is exempt from registration with the Securities and Exchange Commission (SEC).

Heteroskedasticity: A phenomenon where a process does not exhibit equal variance over time.

Hierarchical Model: A modelling technique where models are expressed across several levels.

Integrated Generalized Autoregressive Conditional Heteroskedasticity (IGARCH): An extension of GARCH for non-stationary processes.

International Financial Reporting Standards (IFRS): The accounting standards typically used in Europe.

Investment Grade Bonds: Bonds that are deemed appropriate for **non-accredited** financial investors.

K-th Moving Average: A new approach for filtering and smoothing a time series process for better prediction.

K-th Moving Average ARIMA: A modification of the ARIMA approach using a K-th Moving Average filter.

K-th Moving Average GARCH: A modification of the **GARCH** procedure using a **K-th Moving Average** filter.

K-th Weighted Moving Average: A new approach for filtering and smoothing a time series process for better prediction.

K-th Weighted Moving Average ARIMA: A modification of the **ARIMA** approach using a **K-th Weighted Moving Average** filter.

K-th Weighted Moving Average GARCH: A modification of the GARCH procedure using a K-th Moving Average filter.

Large Cap Stock: See Market Capitalization.

Liquidity: A measure of the ease and speed in which investors can change market positions and investment structure.

London Inter-bank Offer Rate (LIBOR): A measure of the amount banks charge each other to borrow money.

Long Asset Position: A financial position that results in the purchase or a positive position in an asset.

Mahalanobis Distance: A weighted distance metric.

Margin Trading: The use of funds borrowed from a **broker-dealer** to make investments. **Short selling** requires the use of margin positions.

Market Capitalization: Total value of all of a company's shares of stock. Also called Market Cap.

Measuring Attractiveness by a Categorical Based Evaluation Technique (MACBETH): A Multi-Criteria Decision Analysis implementation.

Modern Portfolio Theory: A new focus on financial methods that concentrate on the performance of a portfolio rather than individual assets.

Modular Programming: A programming framework where pieces of code are written to perform specific tasks and then re-used as part of a larger program.

Momentum Investing Strategies: Choosing assets to invest in based on previous growth trends.

Multi-Criteria Decision Analysis: A field of operations research that focuses on helping individuals, companies, and organizations make and evaluate decisions when multiple diverse criteria are present.

Multi-Level Time Series Clustering: A machine learning technique for clustering designed for use with time series data.

Mutual Fund: A professionally managed investment that pools money from many investors.

Nash Equilibrium: An equilibrium state for participants in a game theory model.

North American Industry Classification Code (NAISC): A six digit company identifier code.

Operations Earnings: Net profit directly tied to a company's operating activity. This metric excludes interest and taxes but includes depreciation.

Panel Data: Multi-Dimensional data collected over time.

Pareto Efficient: A state where no individual could be made better off without making another individual worse off.

Pareto Frontier: A set of allocations that are Pareto Efficient.

Portfolio: A collection of assets.

Portfolio Re-balancing: The process of re-allocating portfolio assets to mimic changes in the portfolio structure since its creation.

Pre-Cardinal Information: See Cardinal Utility Scale.

Price/Earnings Ratio (P/E Ratio): The ratio of a company's price relative to its earnings.

Price to Book: The ratio of a company's price to its **book value**.

Price to Cash Flow: The ratio of a company's price relative to its **free cash flow**.

Prospect Theory: A theory by Tversky and Kahneman on several irrational traits exhibited by individuals making decisions on risk and uncertainty.

Quantitative Finance: A branch of finance that uses advanced mathematical and statistical techniques to make investment decisions.

Robo-investing: Investing programs that automatically make investments and manage assets for a client.

Sharpe Ratio: A financial ratio that balances the risk and return of a portfolio.

Shillers Cyclically Adjusted P/E Ratio: A method for computing the P/E ratio over time.

Short Sale: A financial transaction that creates a negative position in an asset using **margin financing**.

Sixty Forty Portfolio: A portfolio that is invested 60% in stocks and 40% in bonds.

Spectral Analysis: A branch of time series analysis that uses **Fast Fourier Transforms** to analyze patterns in a signal. Often used to estimate the **autocovariance** of a process.

Spatio-Temporal Statistics: A branch of statistics with data exhibiting space and time dimensions.

Standard and Poor's 500 index: An index of approximately 500 **large cap** stocks that are listed in the United States.

Stationarity: See Strong Stationarity and Weak Stationarity.

Stochastic Dominance: Process A is said to stochastically dominate Process B if A performs as well as B in every state and better than B in at least one state.

Stochastic Mapping: A stochastic based clustering technique.

Strong Stationarity: A time series process whose distribution does not change over time.

Suitability: The appropriateness of an asset for a given investor given his or her investment goals and preferences.

Supervised Learning: A machine learning technique for structured or labeled data. **Ticker:** An abbreviated symbol used to identify a stock.

Transitive Preferences: A common assumption for preferences that assumes that the utility between choices exhibits mathematical transitivity.

Treasury Bond: A bond issued by the federal government of the United States. Sometimes referred to as a "Treasury."

Unsupervised Learning: A machine learning technique for unstructured or unlabeled data.

Utility Function: A rank order function of preferences.

Von-Neumann Morgenstern Utility Theory: A theory that individuals will maximize expected utility when dealing with risk.

Weak Stationarity: A time series process with constant mean and variance.

Value Investing: Investing in assets that are believed to be undervalued by financial markets.

Variation at Risk (VaR): A measure of the amount of assets of a portfolio potentially at risk in the current financial environment.

Ward's Method: A method for selecting an appropriate cluster size.