Summer 8-5-2019

Reliability of Technical Stock Price Pattern Predictability

Matthew Lutey
mlutey@uno.edu

Reliability of Technical Stock Price Pattern Predictability

A Dissertation

Submitted to the Graduate Faculty of the
University of New Orleans
in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy
in
Financial Economics

by

Matthew Lutey

B.S. Northern Michigan University, 2011
M.B.A. Northern Michigan University, 2013
M.S. University of New Orleans, 2017

August, 2019

Acknowledgment

Thank you to Neal Maroney for guiding me and working with me on this project. Thank you to Dr. Lane, Dr. Naka, Dr. Zirek, Dr. Davis, and Dr. Hassan for job market preparation and feedback on this project. Thank you to other students, faculty, and classmates who have helped me get to this point. Thank you to Dr. Rayome for introducing me to the paper by Lo, Mamaysky and Wang 2000, and for being a friend and mentor.

Table of Contents

List of Tables and Figures

Abstract

Academic research has shown throughout the years the ability of technical indicators to convey predictive value, informational content, and practical use. The popularity of such studies goes in and out over the years and today is being recognized widely by behavioral economists. Automated technical analysis is said to detect geometric and nonlinear shapes in prices which ordinary time series methods would be unable to detect. Previous papers use smoothing estimators to detect such patterns. Our paper uses local polynomial regressions, digital image processing, and state of the art machine learning tools to detect the patterns. Our results show that they are nonrandom, convey informational value, and have some predictive ability. We validate our results with prior works using stocks from the Dow Jones Industrial Average for a sample period from 1925-2019 using daily price observations.

Keywords: Technical Analysis; Charting; Kernel Density Estimation; Image Processing; Support Vector Machines; Cross Validated; Center for Research in Security Prices; Dow Jones Industrial Average; CRSP

Chapter 1

"Reliability of Technical Stock Price Pattern Predictability"

1. Introduction

This dissertation uses state of the art machine learning techniques to process financial time series. The work explores the possibility of using image processing to recognize non-linear stock prices patterns thereby simulating technical trader behavior.

Investors normally fall into either fundamental investors, or technical investors (Linton 2010), Elder 2002, Pring 2002, Edwards and Magee 2007, Lo, Mamaysky, and Wang 2000, Neeley, Rapach, Zhou 2014). Fundamental investors ignore the timing aspect of buying stocks which is a potential drawback since markets have been shown to move in trends over the past 100 years (Covel 2009).

Technical analysis relies on past prices to identify current price behavior and trends that are expected to continue in the future. Literature has been back and forth on the viability of technical analysis. Literature in support of technical analysis include Fama and Blume 1966, Sweeney 1988, Han, Yang and Zhou 2013, Neeley, Rapach, Tu and Zhou 2014, Avramov, Kaplanski, and Subrahmanyam 2018, Osler and Chang 1995, Lo, Mamaysky, and Wang 2000, Blume, Easley, O'Hara 1994.

Behavioral finance argues that investors have constraints on time, informational, and cognitive ability. This plays a role in the decision-making process. (Simon, 1955).

The efficient market hypothesis doesn't leave any room for technical analysis in current markets. The justification says that markets have already priced in all relevant factors regarding a stock. These include past prices (weak-form), all public information (semi-strong form), and all available information including private and insider information (strong form). The major flaw or drawback is that it relies on the assumption that investors are all rational in their decision-making process and they can obtain all available information prior to making decisions in the markets. Any point in time where this doesn't hold (e.g. a Dentist taking a stock tip from a client, or a college student buying a stock because he wants to gamble his parents money) would lead to some degree of inefficiency in markets at least according to the definitions and assumptions outlined in the efficient market hypothesis by Fama 1970.

There are several papers that outline the theoretical support for technical analysis which relies on markets displaying some form of market efficiency, mainly that they are weak-form inefficient. Treynor and Ferguson 1985, in whether information has been fully incorporated into equity prices, Brown and Jennings 1989 using past prices to gain information in current prices. On the role of volume Blume, Easley, and O'Hara 1994 as well as Grundy and McNichols 1989.

Han, Yang and Zhou 2013 provide justification for profitability in technical analysis studies as being investors not acting on all profitable information. Neeley, Rapach, Tu and Zhou 2014provide justification for why technical analysis outperforms macro fundamentals in forecasting risk premium in that they pick up on some omitted fundamental variable.

Lo, Mamaysky and Wang 2000, Osler and Chang 1995, Chang and Osler 1994, automated technical analysis. Blume, Easley and O'Hara 1994 study the role of volume. Avramov, Kaplanski, and Subrahmanyam 2018 study moving averages and their predictability drowning out the effect of a popular momentum strategy. They also provide potential justifications.

Empirical research documenting the inefficiencies of markets include the January effect (Huag and Hirschey, 2006) and the firm size effect (Van Dijik, 2011). Lo and MacKinlay (1988, 1999) on past prices forecasting future returns. Chang and Osler (1994), Osler and Chang (1995).

Lo, Mamaysky and Wang (2000) said that technical indicators can be found using consecutive extrema in stock prices. They also said that consecutive extrema in stock prices are too noisy thus making it impossible to distinguish between informativeness and noise. Their methodology involves smoothing prices to find extrema.

Our paper extends this work by using local polynomial regressions. We compare our results and find the same outcome with the same patterns in the same places according to their empirical sample.

We extend this work to find extrema from pixel values in digital images. The advantage here is that it doesn't rely on time series and that extrema are conveyed by only one point (a pixel).

Time series methods wouldn’t be able to detect these patterns in data. Lo, Mamaysky and Wang 2000 show that simulated geometric Brownian motion detects far fewer patterns than real stock data.

We match our digital image processing method to their result. We again find extrema in the same places and the same patterns.

We test patterns such as head-and-shoulders, inverted head-and-shoulders, broadening top, broadening bottom, triangle top, triangle bottom, rectangle top, and rectangle bottom. We omit double top and double bottom patterns from our image processing framework and leave that to future work.

We extend our results using digital image classification. This uses state of the art machine learning tools such as support vector machines.

We use a sample of Dow Jones Industrial Average (DJIA) stocks to test whether image processing can uncover the pattern rules given training data of consecutive extrema and pattern labels for non-DJIA stocks.

We show that pattern rules can be uncovered by mapping the association of the true pattern with a series of 5 consecutive extrema, given as few as 1000 training examples. This is with a high degree of accuracy.

We test the profitability of the patterns using one-day returns, three days following pattern completion. We test both goodnesses of fit, and difference in means. We show the patterns are time-varying and have most have statistically significant excess returns.

The results of this research have potential implications for academia and industry. Namely, pattern statistics (we generate confidence intervals around the patterns) and program trading.

Future work may want to aggregate the bias between the output of our classifier and trained professional analyst recommendation. This would create a robust pattern dataset.

A trained professional analyst is noted to be either a CMT designation (Certified Market Technician) (www.cmt.org). Other designations include Market Technicians Association (MTA), the Nippon Technical Analysts Association (NTAA), which formed the International Federation of Technical Analysts (IFTA) with the Society of Technical Analysts (STA) in London. There is also a Canadian Society for Technical Analysts (CST). (Edwards and Magee 1966, Linton 2010). The IFTA sponsors the designations for Certified Financial Technician (CFTe) program. They also have a Master of Financial Technical Analysis (MFTA). Further, the IFTA journal sponsors works that include the use of technical analysis by practitioners.

2. Literature Review.

The literature on this subject notes that investors come from two camps, either fundamental investors, or technical investors. Rarely do the investors borrow from both camps. (Neely et al (2013), Pring (2002), Elder (2002), Linton (2010). Covel (2009) “I have established that trading can be fundamentally or technically based.” The main flaw with fundamental analysis pointed out by technical literature is that it fails to detect entry and exit points. There is a clear distinction between trend following and day trading. The market moves in trends which anyone can see by looking at a chart of the Dow Jones Industrial Average (DJIA) over the last 98 years. Trend trading is what is seen to be more profitable in the long run. Technical analysis studies are broken into those that follow trends or determine the possible start of a new one. Moving averages are those that follow trends. Visual patterns are said to determine the start of a new one. No study tries to pick tops and bottoms (successfully). Covel (2009) notes that mechanical trading systems (generally used by trend followers) are based on an objective and automated set of rules. Traders follow these trading rules using a computer and getting themselves to buy in and sell out of a market. This can make life easier by eliminating the emotional aspect of trading decisions.

There are three pillars that go into a trading system, having the system (strategy), sound money management, and human emotion. Balancing all three is like sitting on a three-legged stool. If one goes, the system (trader) goes. (Elder, 2002).

Covel (2009) discusses that trend following is nothing new and that it goes back decades. It is just discovered by different traders at different times.

Covel (2009) notes the significance of fundamental investors in creating bubbles. Noting that technical traders were overshadowed by fundamental investors in the 1990’s dot com bubble and 2008 real estate and credit bubbles, that so many investors and traders with so little strategy were making money hand over fist that trend following disappeared from everyone's radar, even though they kept on making money.

Covel (2009) shows how trend followers have increased the bottom line when every bubble has popped. Creating a hypothetical index of three longtime trends following firms compared against the S&P 500 stock index. Dunn Capital Management, Campbell and Co., and John W. Henry and Co. into an equally weighted index show 1,000 from 1984-2003 turning into $6,236 on the S&P 500, and a whopping $47,891 for trend following.

The premise for reversal strategies is that the market goes in trends. According to Abraham Trading Company (ATC), (Covel, 2009) "commodity interests will, from time to time, enter into periods of major price change to either a higher or lower level. These price changes are known as trends, which have been observed and recorded since the beginning of market history. There is every reason to believe that prices will continue to trend."

Hypothetical growth of $1,000 in ATG from 1987-2003 (using a trend following strategy) shows $34,051 in 2003, compared to the same $1,000 in the S&P 500 growing to $4,280.

Chesapeake Capital is another example of trend following. In the Q/A session at the annual futures and options expo in Chicago at the height of the dot come bubble Jerry Parker noted the dangers of a buy-and-hold mentality. "The strategy of buy and hold is bad. Hold for what?" On asked about predicting where markets are heading Paker responded "I don't know nor do I care. The system that we use at Chesapeake is about the market knowing where it's going." On counter-trend or day trading: "The reason for it is a lot of traders, as well as clients, don't like trend following. It's not intuitive, not natural, too long term, not exciting enough." Parker noted further that "We have a system in which we do not have to rely on our intellectual capabilities. One of the main reasons why what we do works in the markets is that no one can figure out what is happening." A $1,000 investment in Parkers fund (Chesapeake) would result in $12,633 from 1987-2003. This is compared with $4,114 on the S&P500 over the same horizon.

Donchian is an influential trader. He was born in 1905 in Hartford, Connecticut. He graduated from Yale in 1928 with a BA in economics. He was so fascinated by trading that even after losing his investments in the 1929 crash, he returned to work on Wall Street. In 1930, he borrowed some capital to trade shares in Auburn Auto, which is an article by William Baldwin's article on Donchian is referred to as "the Apple Computer of its day." The moment after he made several thousand dollars on the trade, he became a market "technician," charting prices and formulating buy and sell strategies without concern for an investments basic value.

Many chartists outsource their programming. There is some benefit however to knowing your tools, and what is going on under the hood. Ed Seykota is generally acknowledged to have programmed the first computerized trading system. (Covel, 2009).

Visual technical chart patterns are said to pick up on changes in trend. These are known as reversal patterns. (Edwards and Magee, 2007). While these patterns pick up on changes in trends, studies such as moving averages track trends. Moving averages are simple to compute arithmetically, as they are a simple mean of prices. More studies stem from the moving average such as relative strength index, moving average convergence divergence, stochastic momentum index, Bollinger bands, and more.

Nonlinear technical chart patterns have been of interest to practitioners and academics since the 1940s. Original research discussed these patterns and uncovered the rules from trade journals. The main study was whether nonlinear patterns conveyed information to investors. The tools used varied but original research uncovered technical chart pattern rules from trade journals, testing head and shoulders patterns since it was the most common visual/nonlinear pattern. The purpose was to discuss whether the pattern conveyed information to investors. If so, it would spark future research (Lo, Mamaysky, Wang 2000) if not, it could be written off as a 'Voodoo science.' The original paper found that the head and shoulder did alter the conditional distribution of returns. This sparked later research by Lo, Mamaysky, Wang (LMW, 2000). This paper uncovered 10 nonlinear technical chart patterns in 5 pattern pairs. Each pair consists of a bullish and a bearish pattern. Bullish patterns are said to offer positive information about a stock and an investor should take a long position. Bearish patterns are said to offer negative information about a stock and an investor should take a short position. The intuition behind technical chart patterns is that all available information is already incorporated into the stock price, thus fundamentals won't alter the distribution of returns. Investors think that nonlinearities in the chart may indicator buying behavior and be an indicator of future price movement. There are many studies that discuss past prices conveying information about future returns. Ranging from the Alexander Filter Rule (Fama and Blume (1966), Sweeney (1988)), to arithmetic technical chart patterns (Moving Averages, Relative Strength Index, Moving Average Convergence Divergence), Neely et al. (2013), Han et al. (2013), and nonlinear (more difficult to uncover arithmetically) patterns (Lo, Mamaysky, Wang

(2000)). The Intuition behind all chart patterns is that investors look at an analog stock chart image and use it to make decisions about future returns. It is a visual art ((LMW (2000), Linton (2010)).

There are additional patterns outside of the ten patterns discussed by LMW 2000 and are outlined in some academic literature as well as trade journals (IFTA Journal, Futures Magazine), and print books (Edwards and Magee (2007), Linton (2010), Elder (2002), Pring (2002)). Some of these are discussed in a future section but include variations of the patterns tested previously by LMW (2000), and this paper, as well as new patterns not yet tested.

Technical analysis literature is brought to the forefront of academic literature by LMW (2000) by incorporating visual chart patterns, which are harder (than the former paper by Dunis et al. (2013)) to compute arithmetically. Kulkarni and Harman (KH, 2011) outline some of the newer methods in statistical learning theory which can aid in detecting nonlinear patterns data. LMW (2000) call for such methods. Dunis et al. (2013) show modern methods of statistical learning theory such as the support vector machine (SVM, as in KH (2011)) can detect trading rules such as Moving Average Convergence / Divergence and Relative Strength Index. There is little literature to date on the intersection of computer vision and the field of finance.

Rule-Based Pattern Recognition: The method of LMW (2000) to uncover the patterns is built on kernel smoothing and bandwidth selection. Which is basically a moving average with a normal probability distribution on the current observation. The rules are based on identifying extrema, and the location of extrema. Each pattern is based on the completion of five consecutive extrema in rolling windows. LMW(2000) use 38 days as too long of a window would create many patterns that would be difficult to distinguish from noise.

Technical analysis has been common in the cross-section of stocks since Fama and Blum (Journal of Business, 1966). Lo, Mamaysky, and Wang (Journal of Finance, 2000) show that visual patterns also contain the predictive ability and alter the distribution of returns when conditioned on them.

The use of technical analysis is wide and varies from moving average studies, to volume, to nonlinear visual patterns, to Fibonacci ratios. It’s no wonder that Malkiel (1999) related the use of technical analysis from astronomy to astrology. To anon chartist, it can seem like voodoo, especially when one has only been taught the methods of the random walk.

Technical analysis, specifically moving averages pick up on fundamentals (Neeley et al 2013). Han et al (2013) discuss that profitability with technical analysis may be explained by investors not acting on all available information. Elder (2002) A moving average window reflects the average consensus of value of all market participants during the period of its window. It is like a composite photograph that reflects the major features of the market crowd rather than fleeting moods. Neeley et al (2013) show that these indicators forecast equity risk premium as well as or better than macro fundamentals. Which has been shown by Goyal and Welch (2008) to forecast the risk premium in the cross-section of stocks? Opening and closing prices are among the most important prices of the day (Elder (2002)). The opening price reflects all the pressures that have gathered while the market was closed. Openings are often dominated by amateurs who read their newspapers in the evening and trade in the morning. Professional traders are active throughout the day. (Elder (2002)) Closing prices are especially important because the settlement of trading accounts depends on them.

Edwards and Magee (9^{th} Edition, 2007) (the 4^{th} edition was 1966). The book notes the definition of technical analysis. Discussing that the prices of a stock at a given time is held together by all available information by all people involved in the market.

The market is constantly looking ahead, attempting to discount future development. Balancing al information from many different hues. Past and present information is already considered an old/stale. This includes fundamental information. The going price which is established by the market comprehends all fundamental information which the statistical analyst can hope to learn plus any information known only to him, or to a few insiders. In the language of Fama and Blume (1966) this is the strong-form efficiency. Weak-form being only past prices being reflective of the market information, and semi-strong form being all available public and private information but not an insider. Prices move in trends which tend to continue until something happens to change the supply-demand balance. Certain patterns or formations, levels or areas, appear on the charts which have a meaning and can be interpreted in terms of probable future trend development. They are not infallible, (in the words of Edwards and Magee (9^{th} edition)), but the odds are definitely in their favor. They are more prescient than the best-informed most shrewd of statistics.

Empirically this is supported by Neely, et al. (2013), Lo, Mamaysky, and Wang (2000), Han et al (2013), Fama and Blume (1966), Sweeney (1988), among others. Neeley et al (2013), show that technical indicators are more predictive then fundamental indicators (outlined by Goyal and Welch (2008)) in the cross-section of stocks. They show that technical indicators pick up on some omitted fundamental information. Han et al (2013) show that patterns such as moving averages may be persistently profitable because investors choose not to take action on all information.

A technical analyst may go further into his claims. (Following Elder (2002) we use the pronoun he or his because there are more male technical analysts than female. Although Elder (2002) finds that females are more profitability and more in control of their emotions, there are more than twice as many male analysts as female.) He may offer to interpret the chart of a stock whose name he does not know, so long as the record of trading is accurate and covers a long enough term to enable him to study its market background and habits. He may suggest that he could trade with profit in a stock knowing only its ticker symbol, completely ignorant of the company, the industry, what it manufactures or sells, or how it is capitalized. This is an extreme case.

Edwards and Magee (9^{th}, 2000) go on to discuss that fundamental analysis is based on estimating a company's earnings for both the current year and the next year and recommending stocks on that basis. The record on that bases, as estimated by Barron's is that earnings estimates averaged 18% error in the 30 DJIA stocks for any year already completed and 54% error for the year ahead, choosing the 10 DJIA stocks with the best earnings estimates would have produced a 10-year cumulative gain of 40.5% while choosing the 10-worst would have produced a 10-year cumulative gain of 142.5% this is from the same Barron's article.

Charts are the working tools of the technical analyst. They have been developed in a multitude of forms and styles to represent graphically almost anything and everything that takes place in the market. Charts vary in type and time frame from weekly, daily, hourly, "point and figure", candlestick, etc. They may be constructed arithmetic, logarithmic, or square-root-scale, or projected as "oscillators." They may show moving averages, proportion of trading volume to price movement, average price of "most active" issues, odd-lot transactions, short interest, and an infinitude of other relations, ratios, and indexes all technical in the sense that they are derived, directly or indirectly from what has actually been transacted on the exchanges.

The goal of many of these is to make a combination of indexes which will give a warning of a change in the trend without failing or going wrong. This is the job of a full-time economic analyst.

Technical analysis is mainly concerned with the simplest form of a stock chart – a record of the price range (high and low), closing price, and volume of shares traded each day. These daily graphs can be shown on weekly, or monthly frequencies (or intraday) for which most stocks can be purchased, readily made which are easily generated and commercially available on almost any investment software.

Charts can be made using a piece of paper and a pencil. And some time. On a stock chart, the horizontal axis represents time and the verticle axis represents price. The space between the bars on the horizontal axis represents days, space is usually provided to plot volume, shares traded. For our interpretation, we only need the closing price.

When information is processed by a computer, it takes the price time observations (time-series) and tries to make interpretations arithmetically. When humans process the information ocularly, they are making fuzzy decisions and letting the cortex do a lot of math without them realizing it. It is difficult for an automated system to replicate the visual art of technical analysis by a human.

One of the earliest forms of technical analysis is the Dow Theory. Mapped out by Charles Dow who was born in 1851. He was an American journalist who co-founded the Dow Jones & Company. (Linton 2010).

Dow mapped out some of the basic tenets of technical analysis. He also founded the Wall Street Journal and devised the Dow Jones Index as part of his work in researching market movements.

When Dow died in 1902 the editor of the Wall Street Journal and two other colleagues summarized some 250 of Dow's editorials to produce what we now know as Dow Theory. It can be summarized in six basic tenets. "

1. The market has three movements – primary to the major trend of about a year to several years, the medium swing or intermediate reaction of 10 days to 3 months and generally retracing 33% to 66% of the major trend and the short swing or minor movement which can last from hours to weeks. These movements can all be occurring simultaneously – trends, within trends, within trends.
2. The trends have three phases – an accumulation phase with shrewd investors 'in the know' acting contrary to popular opinion, a public participation phase where the market catches

on and prices move more dramatically and a distribution phase where the astute investors begin to unwind their positions.

3. The stock market discounts all news – prices quickly absorb all new information as soon as it becomes available. This was quite an admission from the editor of the leading newspaper at the time and agrees with what we now know as an efficient market hypothesis.
4. Stock market averages must confirm each other – Dow also devised the Transports Average, which like the better-known Dow Jones Industrial Average, survives to this day. Calculated using rail and industry stocks respectively, Dow argued that they need to confirm each other for any trend in prices to be believed.
5. Trends are confirmed by volume – Dow believed that price moves accompanied by high volume represent the 'true' market view and that price moves on low volume were to be taken less seriously.
6. Trends exist until definitive signals prove that they have ended otherwise – the primary trend should be given the benefit of the doubt during secondary reversals. Which is the foundation for what is known as 'the trend is your friend.' Cloud Charts are especially helpful in knowing when a trend has ended "

Dow's original ideas form the basis of the subject of technical analysis in the West as it is known today, it is not entirely clear how heavily these techniques were being used by the trading community at that time. Jeremy Du Plessis (2005) cites textbook references in the middle of the last century pointing to the use of Point and Figure charts pre-1900. (Linton 2010).

According to Linton (2010), it probably wasn't until Robert Edwards and John Magee (1948) that the subject of Technical analysis gained traction as a method of analysis. Their work is now in its ninth edition and their definition of technical analysis is still one of the most cited, noting "Technical analysis is the science of recording, usually in graphic form, the actual history of trading (meaning price changes, volumes, etc.) in a certain share, or commodity, etc., and then deducting from that pictured history, the probable future trend."

A Dow Theory Line in the chart of one of the Averages may be either a Consolidation or Reversal Formation and is rather more likely to be the former than the latter. A Dow Line is a loose Rectangle. Almost any sort of sideways price pattern is termed "Congestion" or trading area provided trading volume tends to diminish during its construction (and provided it doesn't show broadening tendencies), constitutes as Consolidation. Most areas of consolidation are well defined, taking on a recognizable pattern.

A Discussion of Nonlinear Patterns. A healthy uptrend moves up in steps. (Elder (2002)). Head-and-shoulders tops mark the ends of uptrends. The “head” is a price peak surrounded by two lower peaks, or “shoulders.” A neckline connects the lows of declines from the left shoulder and the head. The neckline does not have to be horizontal – it may be flat, rising, or falling. A downsloping neckline is especially bearish – it shows that bears are becoming strong. Chang and Osler (1994) show the predictability of this pattern in the cross section of forex markets. It is used as a representative set of all nonlinear patterns. It is discussed as being the most common nonlinear visual pattern (Pring (2002), Elder (2002)). An Inverse Head and Shoulders are sometimes referred to as a head-and-shoulders bottom (Elder (2002)). – a mirror image of a head-and-shoulders top. It looks like a silhouette of a person upside down: the head at the lowest point, surrounded by two shoulders. This pattern develops when a downtrend loses its force and gets ready to reverse. A downtrend is summarized by making successively lower lows. The trend stops at the lowest low (Elder (2002)). A strong rally from the head allows you to draw a neckline. When a decline from the neckline fails to reach the level of the head it creates the right shoulder. When prices rally from the right shoulder above the neckline on increased volume, they complete the head-and-shoulders bottom and a new uptrend begins.

Sometimes a head-and-shoulders bottom is followed by a pullback to the neckline on low volume, offering an excellent buying opportunity. Measure the distance from the bottom of the head to the neckline and project it upward from the point where the neckline was broken. This gives you a minimum measurement for a rally, which is frequently exceeded.

Edwards and Magee (9^{th} Edition, 2000) discuss that the Head-and-Shoulders pattern is a relation of Dow Theory. It is an adaption of the Dow Theory to the action of an individual stock. The rally from the decline of the head to the neckline to the top of the right shoulders to the decline that breaks the neckline and eventually starts a new trend. This may be why it is the most frequent and reliable pattern. There are several examples of this in Edwards and Magee (1966), as well as Pring (2002), Elder (2002) and Linton (2010).

All basis of technical analysis starts with Dow theory and the head-and-shoulders is a logical predecessor for visual patterns.

The main difference between a head-and-shoulders top/bottom is that the prices distribute from the top and accumulate from the bottom. Informing both shoulders and eventually breaking out.

The tactics for trading inverse head and shoulders (head-and-shoulders bottom) is similar to head-and-shoulders tops. You risk less money trading at bottoms because prices are less volatile and you can use closer stops.

Additional variations include multiple head-and-shoulders, which is like it sounds. Multiple shoulders on the left, followed by multiple heads and multiple shoulders on the right. Almost like conjoined twins, with both sets of arms. The same can hold true on the reverse with multiple shoulders, and one or more heads. This is shown in Pring (2002), Edwards and Magee (1966,2007).

Other formations, not related to head-and-shoulders include rounding tops and bottoms. (Linton (2010), Edwards and Magee (1966,2007), Pring (2002), Elder (2002)). Typically the rounding in price is accompanied by rounding in volume.

Triangles are another pattern that is of importance. This is accompanied by an important theory on prices (following Edwards and Magee (1966,2007)). The market value of a security is determined solely by the interaction of supply and demand. Supply and demand are governed at any given moment by many hundreds of factors, some rational and some irrational. Information, opinions, moods, guesses. The market weighs these automatically. It would be difficult or impossible for a human to consider all of these factors. Disregarding minor fluctuations, prices move in trends that persist for an appreciable length of time. Changes in trend, which represent an important shift in the balance between supply and demand, however, caused, are detectable sooner or later in the action of the market itself.

Triangles can be either consolidation or reversal patterns. Right angle triangles are more predictable (Edwards and Magee (1966, 2007)) then are symmetric triangles. Symmetrical triangles have no way of knowing whether they point up or down until they are broken out of. What is shown in the literature is that the trend will continue more often than not (i.e., an uptrend will remain an uptrend after price breaks from a triangle)? Right angle triangles are more transparent in that we know the direction of the trend as soon as the triangle is recognized as being a right angle triangle. Rectangles are more similar and resemble symmetrical triangles.

Rectangles are bounded by the top and bottom horizontal lines and consist of sideways price movements. On occasion, they consist of slightly upward or downward sloping parallel lines. As long as the breakout is trivial, it may be treated as a rectangle. There are occasions where the boundaries converge. If this happens, then rectangles may be treated as rectangles or symmetrical triangles. The end result will be the same.

Rounding tops or head-and-shoulders may merge or form into a rectangle. The type of trading involved will be trivial and readily apparent when facing a rectangle or a head-and-shoulders.

A head-and-shoulders reflects strong sellers and weak buyers, which can be seen before the conflict has ended. A rectangle represents a contest between the two groups of approximately equal strength, between owners of the stock who wish to dispose of their shares at a certain price and those who wish to accumulate the stock at a certain lower amount. They go back and forth until one group is exhausted. This happens suddenly and the price breaks out. No one can tell which group is going to win until the breakout happens.

An investment trust, large shareholder, has sufficient or good reasons to sell at the top or “supply line” of the rectangle. Another investment trust or group of insiders has an equally good reason for buying at the bottom of the rectangle. The “Demand Line.” If the “spread” between the top and bottom line is wide enough (8-10% of the market value of the stock), the situation may attract a following of short trades that look to pick off the price at the supply and demand levels. If a stock is moving between 76, and 69 buyers may try and buy at 69 and sell at 76. Or, borrow at 76 (sell short) and cover at 69. This type of activity can extend the rectangle, although the number of shares involved is seldom enough to affect the final outcome. Trading inside a rectangle can be profitable at times, especially if accompanied by protective stops. These stops would go outside of the “supply” and “demand” lines at 76 and 69 respectively. These are for both short and long orders respectively.

Rectangles were more common in the 1920s than they were in the 1950s according to Edwards and Magee (1966). Investors would artificially create them.

The SEC has stepped into no longer permit this type of activity, referred to as "wash sales." The tactic is washed out by the SEC which looks for such behavior. According to Edwards and Magee (1966, 2007), this is probably the main reason there are fewer rectangles in the 1950s than the 1920s.

A clearly defined rectangle is almost as reliable as a head-and-shoulders, although not as powerful. Premature breakouts are slightly more frequent from rectangles than triangles.

Pullbacks which are noted as a return of prices to the boundary of the pattern, subsequent to its initial breakout are more common with rectangles than with symmetric triangles. A pullback (or throwback) occurs within 3 days to 3 weeks in about 40% of all cases. A pullback being for a downside breakout and a throwback for an upside breakout.

A rectangle is more often a consolidation pattern than a reversal. The ratio is about the same as with symmetrical triangles. As reversal patterns, rectangles appear more frequently at bottoms than at tops. Long thin, dull rectangles are not uncommon at primary bottoms, sometimes grading into a flat bottomed saucer or dormancy.

A safe formula for measuring implications is given by the rectangle with. Prices should go at least as far in points beyond the pattern as the difference in points between the top and bottom lines as the pattern itself. They may go further. The brief, wide-surging forms, which are nearly square in shape on the chart with the active turnover, are more dynamic then the long narrower manifestations moves out of the latter almost always hesitate or react at the "minimum" point before carrying on.

The most common form of a triangle is composed of a series of price fluctuations, each of which is smaller than its predecessor, each minor top failing to attain the height of the preceding rally, and each minor recession stepping above the level of the preceding bottom. The result is a sort of contracting "Dow Line" on the chart- a sideways price area of the trading range whose top can be more or less accurately defined by a down-slanting boundary line and whose bottom can be similarly bounded by an up-slanting line. This type of triangle is called a symmetric triangle. In the language of geometry, it might be called an acute triangle, since it is not all necessary that is top and bottom boundaries be of equal length or, in other words, make the same angle with the horizontal axis. The pattern has a strong tendency to approximate the symmetrical form, but an additional name it the “coil.” This is also discussed in Elder (2002). Elder likely draws from Edwards and Magee (1966,2007). Edwards and Magee (2007) note that while the process of contraction or coiling, which make up the price action of the symmetrical triangle pattern, is going on, trading activity exhibits a diminished trend, which may be irregular, and persistent as time goes on. The converging upper and lower boundary lines of the price formation come together somewhere out to the right (the future in the time series) of the chart, at the apex of the triangle.

A compact, clean-cut triangle is a fascinating picture (Edwards and Magee 1966), but it has tricky features. Elder (2002) denotes a true pattern reaching up and jumping out at you. If you have to go looking for it, it isn’t there. Beginners tend to go looking for these patterns. Breaking out of the apex is a false move.

A rectangle is a chart pattern that contains price movements between two parallel lines. They are usually horizontal but can sometimes slant up or down (see “Lines and Flags,” later in this section). Rectangles and triangles can serve as continuation or reversal patterns. You need four points to draw a rectangle: The upper line connects two rally tops, and the lower line connects two bottoms. These lines should be drawn through the edges of congestion areas rather than across the extreme highs and lows.

The upper line of a rectangle identifies resistance, while the lower line identifies support. The upper line shows where bulls run out of steam; the lower line shows where bears become exhausted. A rectangle shows that bulls and bears are evenly matched. The key question is whether bulls or bears will eventually win the battle within this pattern. If volume swells when prices approach the upper border of a rectangle, an upside breakout is more likely. If volume increases when prices approach the lower border, a downside breakout is more likely. A valid breakout from a rectangle is usually confirmed by an increase in volume – a one-third to one-half higher than the average of the previous five days. If the volume is thin, it is likely to be a false breakout. Rectangles tend to be wider in uptrends and narrower in downtrends. The longer a rectangle, the more significant a breakout. Breakouts from rectangles on weekly charts are especially important because they mark important victories for bulls or bears. There are several techniques for projecting how far a breakout is likely to go. Measure the height of a rectangle and project it from the broken wall in the direction of the breakout. This is the minimum target. The maximum target is obtained by taking the length of the rectangle and projecting it vertically from the broken wall in the direction of a breakout. Tony Plummer writes that a rectangle is a part of a spiral-like development of a trend. He recommends measuring the height of a rectangle, multiplying it by three Fibonacci ratios (1.618,2.618, and 4.236), and projecting those measurements in the direction of the breakout to obtain a price target. Floor traders can profit from trading short-term swings between a rectangle's walls, but the big money is made by trading in the direction of a breakout. When trading within a rectangle, buy at the lower boundary and sell short at the upper boundary. Oscillators help you decide when prices are ready to reverse within a rectangle. Stochastic, the Relative Strength Index, and Williams %R mark price reversals within rectangles when they hit their reference lines and change directions. To find out whether upside or a downside breakout is more likely, analyze the market in a longer timeframe than the one you are trading. If you want to catch a breakout on a daily chart, identify the weekly trend because a breakout is more likely to go in its direction. Once you buy an upside breakout or sell short a downside breakout, place your protective stop slightly inside the rectangle. There may be a pullback to the rectangle wall on light volume, but prices should not return into a rectangle after a valid breakout.

Lines and Flags: A line is a kind of a rectangle – a lengthy congestion area. In Dow theory, a line is a correction against the primary trend. It is a congestion zone whose height is approximately 3 percent of the current stock market value. When the stock market "draws a line" instead of reacting more deeply against its major trend it shows a particularly strong primary trend. A flag is a rectangle whose boundaries are parallel but slant up or down. Breakouts tend to go against the slope of the flag. If a flag slants upward, a downside breakout is more likely. If the flag slants down, an upside breakout is more likely.

If you see a downsloping flag in an uptrend, place a buy order above the latest peak of the flag to catch an upside breakout. A rising flag in an uptrend marks distribution, and a downside breakout is more likely. Place an order to sell short below the latest low of that flag. Reverse the procedure in downtrends.

Triangles: A triangle is a congestion area whose upper and lower boundaries converge on the right. It can serve either as a reversal or, more often, as a continuation pattern. Some technicians call triangles coils. The market winds up and the energy of traders becomes compressed, ready to spring from a triangle.

A small triangle whose height is 10 to 15 percent of the preceding trend is more likely to serve as a continuation pattern. Many uptrends and downtrends are punctuated by these triangles, as sentences are punctuated by commas. Large triangles whose height equals a third or more of the preceding trend are more likely to serve as reversal patterns. Finally, some triangles simply fizzle out into listless trading ranges.

Triangles can be divided into three major groups, depending on their angles. The upper and lower lines of symmetrical triangles converge at the same angles. If the upper line is inclined 30 degrees to the horizontal, then the lower line is also inclined 30 degrees. Symmetrical triangles reflect a fair balance of power between bulls and bears and are more likely to serve as continuation patterns. An ascending triangle has a relatively flat upper boundary and a rising lower boundary. Its flat upper boundary shows that bulls are maintaining their strength and can lift prices to the same level, while bears are losing their ability to drive prices lower. An ascending triangle is more likely to result in an upside breakout. A descending triangle has a relatively flat lower boundary, while its upper boundary slants down. It's flat lower boundary shows that bears are maintaining their strength and continue to drive prices down, while bulls are losing their capacity to lift prices. A descending triangle is more likely to lead to a downside breakout. Volume tends to shrink as triangles get older. If volume jumps on a rally toward the upper boundary, an upside breakout is more likely. If volume becomes heavier when prices fall toward the lower boundary, a downside breakout is more likely. Valid breakouts are accompanied by a burst of volume – at least 50 percent above the average for the past 5 days. Valid breakouts occur during the first two-thirds of a triangle. It is better not to trade breakouts from the last third of a triangle. If prices stagnate all the way into the apex, they are likely to remain flat. A triangle is like a fight between two tired boxers who keep leaning on each other. An early breakout shows that one of the fighters is stronger. If prices stay within a triangle all the way into the apex, that shows that both boxers are exhausted and no trend is likely to emerge. Charts of related markets often show triangles at the same time. If gold, silver, and platinum all trace triangles and gold break out to the upside, then platinum and silver are likely to follow. This approach works well with currencies, especially with closely related ones, such as the German Mark and Swiss Franc. It also works with stocks in the same group – compare General Motors to Ford but not to IBM. Triangles provide a minimum target for a move following a breakout. Measure the height of a triangle at its base and project vertically from the point where the triangle was broken. If you are dealing with a small triangle in the midst of a dynamic trend, that minimum measurement is likely to be exceeded. You can also the Fibonacci projections mentioned earlier. Trading Rules: It’s better not to trade minor swings within a triangle unless that triangle is very large. As a triangle grows older, the swings become narrower. Profit potential shrinks, while slippage and commissions continue to take just as bad a bite from your account as before. If you trade inside a triangle, use oscillators such as Stochastic and Elder-ray. They can help you catch minor swings. In trying to decide whether a triangle on a daily chart is likely to lead to the upside or a downside breakout, look at the weekly chart. If the weekly trend is up, then a triangle on the daily chart is more likely to break out to the upside, and vice versa. When you want to buy an upside breakout, place a buy order slightly above the upper boundary of a triangle. Keep lowering your order as the triangle becomes narrower. If you want to short a downside breakout, place a sell order slightly below the lower boundary. Keep raising it as the triangle becomes narrower. Once you are in a trade, place a protective stop slightly inside the triangle. Prices may pull back to the wall, but they should not return deep inside a triangle following a valid breakout. When a breakout from a triangle is followed by a pullback, pay attention to volume. A pullback on heavy volume threatens to abort the breakout, but a pullback on light volume offers a good opportunity to add to your position. When prices approach the last third of a triangle, cancel your buy or sell orders. Breakouts from the last third of a triangle are very unreliable.

Atypical Triangles: A pennant is a small triangle whose lines are slanted in the same direction. Pennants that slant against the trend serve as continuation patterns. There is an old saying “The pennant flies at half-mast” – a rally is likely to travel as far after the pennant as it did before. A pennant that slants in the direction of the trend indicates exhaustion – a trend is nearing a reversal. A widening triangle occurs when prices set a series of higher highs and lower lows. This pattern shows that the market is becoming hysterically volatile, with bulls and bears pouring in. The fight between bulls and bears becomes too hot for the uptrend to continue – a widening triangle kills an uptrend. A diamond starts out as a widening triangle and ends as a symmetrical triangle. You have to squint very hard to recognize it. Diamonds are prime examples of Rorschach-type patterns for chartists. If you look hard enough, you will find them, but their trading usefulness is minimal. Double Tops and Double Bottoms are rare. Triple tops and bottoms are even rarer. Most often they are misinterpreted as another reversal form. They can seldom be determined until prices have gone a long way from them.

A double top is formed when a stock advances to a certain level with, usually high volume and approaching the top figure, then retracts with diminishing activity, then comes up again to the same (or practically the same) top price as before with some pickup in turnover, but not as much as the first peak, and then finally turns down a second time for a major or consequential intermediate decline. A double bottom is the same picture, upside down. The triple type makes three tops (or bottoms) instead of two. Most often they are misinterpreted as another reversal form. They can seldom be determined. Double tops occur when prices rally to the area of the previous high. Double bottoms occur when prices fall near the previous low. The second top or bottom can be slightly above or below the first. This often confuses beginning analysts. Savvy traders use technical indicators to identify double tops and bottoms. They are often market by bullish and bearish divergences. Buying at double bottoms and selling short at double tops offer some of the best trading opportunities.

Double bottoms and double tops can be extended to show triple bottoms and triple tops. They have the same form but multiple tops from before the breakout.

For double tops, if two tops are more than a month apart, they are not likely to belong to the same consolidation or congestion formation. If the reaction between the first and second high reduces prices by 20% of their top value, the odds favor a double top interpretation. There are cases where the two peaks occur 2 or 3 weeks apart, and the valley between them only descends about 15%. Most true double tops develop 2 or 3 months apart. Generally speaking, the time element is more critical than the depth of the reaction between the tops.

Double bottoms are the same as double tops but on the reverse. Logically if there are double tops and double bottoms there are triple tops and triple bottoms. Triple tops are wide enough space with deep and usually rounding reactions between them.

The above paragraphs discuss most reversal patterns. Additional pattern types include broadening formations. They are sometimes called “inverted triangles” because, starting with very narrow fluctuations, they widen out between diverging rather than converging boundary lines. They are classified in Edwards and Magee (1966,2007) as broadening patterns because except for the inverted resemblance to a triangle they have different implications.

If symmetrical triangles are areas of “double” between groups of investors, and rectangles are “conflict” Broadening formations is a situation when the public is misinformed and overreacting to rumors. Broadening formations have multiple forms.

The “symmetrical” form consists of a series of price fluctuations across a horizontal axis, with each minor top higher and each minor bottom lower than its predecessor. The pattern may be roughly marked off by two diverging lines, the upper sloping (from left to right) and the lower sloping down. These are loose and irregular. As symmetrical triangles are compact and regular. The tops and bottoms within the formation tend to fall with fair precision on those boundary lines. In a broadening formation, the rallies and declines usually do not step at clearly marked boundary lines.

Three Major Groups of Indicators: Indicators can help you identify trends and their turning points. They can provide a deeper insight into the balance of power between bulls and bears. Indicators are more objective than chart patterns. The trouble with indicators is that they often contradict one another.

Types of Divergences: Oscillators, as well as other indicators, give their best trading signals when they diverge from prices. Bullish divergences occur when prices fall to a new low while an oscillator refuses to decline to a new low. They show that bears are losing power, prices are falling out of inertia, and bulls are ready to seize control. Bullish divergences often mark the end of downtrends.

Bearish divergences occur in uptrends – they identify market tops. They emerge when prices rally to a new high while an oscillator refuses to rise to a new peak. A bearish divergence shows that bulls are running out of steam, prices are rising out of inertia, and bears are ready to take control. There are three classes of bullish and bearish divergences. Class A divergences identify important turning points – the best trading opportunities. Class B divergences are less strong, and class C divergences are least important. Valid divergences are clearly visible – they seem to jump from the charts. If you need a ruler to tell whether there is a divergence, assume there is none. Class A bearish divergence occurs when prices reach a new high but an oscillator reaches a lower high than it did on a previous rally. Class A bearish divergences usually lead to sharp breaks. Class A bullish divergences occur when prices reach a new low but an oscillator traces a higher bottom than during its previous decline. They often precede sharp rallies.

Class B bearish divergences occur when prices make a double top but an oscillator traces a lower second top. Class B bullish divergences occur when prices make a double bottom but an oscillator traces a higher second bottom.

Class C bearish divergences occur when prices rise to a new high but an indicator stops at the same level it reached during the previous rally. It shows that bulls are becoming neither stronger nor weaker. Class C bullish divergences occur when prices fall to a new low but the indicator traces a double bottom.

Class A divergences almost always identify good trades. Class B and C divergences more often lead to whipsaws. It is best to ignore them unless they are strongly confirmed by other indicators.

Triple Bullish or Bearish Divergences consist of three price bottoms and three oscillator bottoms or three price tops and three oscillator tops. They are even stronger than regular divergences. In order for a triple divergence to occur, a regular bullish or bearish divergence first has to abort. That’s another good reason to practice tight money management! If you lose only a little on a whipsaw, you will not suffer – and you will have both the money and psychological strength to re-enter a trade.

Candlestick charts originated in Japan several centuries ago, but have recently gained a following in other countries. This is an alternative to the familiar bar chart (Pring, 2002).

A typical candle consists of two parts: the real body, that is, the rectangular part, and the shadow or wick, that is, the two vertical extensions. The top and bottom of the rectangle are determined by the opening and closing prices for the day. If the closing price is above the opening (the real body), it is plotted in white. When the close is below the opening, it is plotted in black. The top of the real body represents the opening price, the bottom the close. This is reversed in the case of a white rectangle where the close is plotted at the top and the open at the bottom.

The thin, vertical shadow lines from the real body reflect the high and low for the day. Since the closing and opening prices can be identical, or identical with the high or low, there are a number of possible combination that needs to be represented.

Candlesticks provide essentially the same information as bar charts, but their more pronounced visual representation of the material enables technicians to identify characteristics that are less obvious on bar charts. Certain phenomena in bar charts have been given their own names, such as key reversal days or island reversal days, likewise, with candles. Because of the large number of potential variations for both individual days and price formations on encompassing several days, it is common practice to give exotic names to the various possibilities.

These are outside the scope of this research but would serve as a compliment or extension. Nielson (2001) covers these, as well as Pring (2002). Candlesticks may be used on their own or combined with reversal patterns. They hold their own on identifying short-term reversals and continuation situations.

Lo, Mamaysky, and Wang (2000) discuss the predictability of five pattern pairs previous discussed, but not tested by Chang, Osler (1994). These patterns include Head and Shoulders, Inverse Head and Shoulders, Broadening Top, Broadening Bottom, Triangle Top, Triangle Bottom, Rectangle Top, Rectangle Bottom, Double Top, Double Bottom. Chang, Osler (1994). Nonlinear visual technical patterns are used widely by practitioners (LMW (2000), Chang and Osler (1994), Pring (1988). Volume is widely accepted by both practitioners and academics as being useful in forecasting returns. Practitioners and academics have been at odds over the validity of the technical analysis. (Blume, Easley, O'Hara (2000)). Nonlinear visual patterns are shown to be predictive when the volume is increasing (Lo, Mamaysky, Wang (2000)), but the evidence is inconclusive when the volume is decreasing.

de. Malkiel(1999) the author of a Random Walk Down Wall Street refers to technical analysis as about as useful as astrology. In response to this Lo, Mamaysky, and Wang (2000) note in their paper that Lo and Mackinlay reject the random walk for a large portion of the aggregate stock market (1967-1999).

LMW (2000) notes "Technical analysis, known as 'charting', has been a part of financial practice for many decades but this discipline has not received the same level of academic scrutiny and acceptance as more traditional approaches such as fundamental analysis."

"It has been argued that the difference between fundamental analysis and technical analysis is not unlike the difference between astronomy and astrology. Among some circles, technical analysis is known as 'voodoo finance.'" Lo and Mackinlay (1988, 1999) have shown that past prices may be used to forecast future returns to some degree, in rejecting the Random Walk Hypothesis for weekly U.S. stock indexes.

Campbell, Lo, and MacKinlay(1997, 43-44) provide an example of the linguistic barriers between technical analysts and academic finance:"The presence of clearly identified support and resistance levels, couple with a one-third retracement parameter when prices lie between them, suggests the presence of strong buying and selling opportunities in the near-term." "The magnitudes and decay pattern of the first twelve autocorrelations and the statistical significance of the Box-Pierce Q-statistic suggest the presence of a high-frequency predictable component in stock returns." Our findings suggest a high degree of nonrandomness in uncovering the nonlinear visual patterns. Over 9500 daily price observations on a single security, we find that at most 68 patterns are uncovered (Head and Shoulders). Chang and Osler (1994) weren't wrong to test Head and Shoulders as a subset of the full pattern group. This leads into part two where we use the patterns from chapter 1 as a benchmark for digital image processing. We test whether image processing can detect the patterns and benchmark for speed and accuracy. Recently CJ Neely (2010) ties in the use of both fundamental and technical analysis in studying the equity risk premium. Finding that technical indicator perform just as well, if not better than traditional macroeconomic variables. Jasemi, Milad, and Ali M. Kimiagari. (2012), note that moving averages are one of the most popular and easy to use tools available for technical analysts. They form the building blocks for other technical indicators and overlays. Many studies consider the use of technical indicators in event study from Fama and Blume (1966), Sweeny (1988), and more recently Neely (2010) and Han, Yufeng, Ke Yang, and Guofu Zhou (2013) study technical indicators on the equity risk premium, and Fama French 3 factors, and capital asset pricing model (CAPM).

Fama and Blume (1966) note in recent literature there has been considerable interest in the theory of random walks in the stock-market prices. The basic hypothesis of the theory is that successive price changes in individual securities are independent random variables. Independence implies, that the history of a series of changes cannot be used to predict future changes in any "meaningful" way. The first order coefficients for the daily price changes of the individual securities are positive in twenty-two out of thirty cases, and the average value of the coefficients is 0.026. The results are consistent with the small degree of persistence on a very short-term basis that was uncovered by the filter tests. The discussion suggests that for measuring the direction and degree of dependence in price changes, the standard tools are probably as powerful as the Alexandrian filter rules. Sweeney (1988) notes an application of ½ of 1 percent filter rules remains profitable for floor traders. Why the markets seem weak-form inefficient at their level of transaction costs. One answer is that the cost of a seat on the exchange is the present value of the profits that could be made. This does not explain why current seat holders have not competed for these profits to zero; there are too many of them to argue that a successful conspiracy is at work. Han, Yufeng, Ke Yang, and Guofu Zhou (2013) document that an application of a moving average timing strategy of technical analysis to portfolios *sorted* by volatility generates investment timing portfolios that substantially outperform the buy and hold strategy. For high-volatility portfolios, the abnormal returns, relative to the capital asset pricing model (CAPM) and the Fama-French 3-factor models, are of great economic significance and are greater than those form the well-known momentum strategy. Recent papers by Neely, Rapach, Tu, and Zhou (2013) confirm the predictability of technical indicators, in the face of macroeconomic variables. Noting that macroeconomic indicators outline by Goyal and Welch (2008) are actually less predictive than technical indicators in the cross-section of stocks. Avramov, Kaplanski, and Subrahmanyam (2019) suggest that a modern application a moving average strategy (named Moving Average Deviation) shows predictability, drowns out momentum and profitability (before transaction costs) in the U.S. aggregate stock market. NYSE, Nasdaq, AMEX (CRSP) data. Other technical indicators (Relative Strength Index, Moving Average Convergence Divergence, among others) are used but are based on the moving average. The above indicators are the ones that are tested in Zhou et al (2013) and are found to be highly predictive in the cross-section of stocks. Nonlinear visual chart patterns are different from moving average based strategies in that they are largely in the eye of the beholder. Lo, Mamaysky, Wang (2000) provide a framework for identifying nonlinear visual patterns objectively. Visual nonlinear patterns were first discussed academically by Chang and Osler 1994. They test the head and shoulders pattern as a subset of all visual patterns, noting patterns later tested by Lo, Mamaysky, Wang (2000) as well as patterns to be tested outlined in Pring (1998). Lo et al (2000) find that stock returns can be predicted by visual chart patterns and the patterns may contain some practical value. They are careful to note predicted does not necessarily mean profitable. McLean and Pontiff (2016) the findings point to mispricing as the source of predictability. Post-publication, stocks in characteristic portfolios experience higher volume, variance, and short interest, and higher correlations with portfolios that are based on published characteristics. Neely further discusses four types of theoretical models in the literature that seek to explain technical indicators in asset pricing. They are noted below.

The first type of theoretical model recognizes differences in the time for investors to receive information. Under this friction, Treynor and Ferguson (1985) show that technical analysis is useful to assess whether information has been fully incorporated into equity prices, while Brown and Jennings (1989) demonstrate that past prices enable investors to make better inferences about price signals. In addition, Neely points out that Grundy and McNichols (1989) and Blume, Easley, and O'Hara (1994) show that trading volume can provide useful information beyond prices.

An additional type of model is Cespa and Vives (2012) who show that asset prices can deviate from their fundamental values if there is a positive level of asset residual payoff uncertainty and/or persistence in liquidity trading. In this setting, rational long-term investors follow a trend. In the real world, different responses to information are more likely during recessions, due to consumption smoothing asset sales by households that experience job losses and liquidation sales of margined assets by some investors (Neely). These factors help to explain why we find that technical indicators display enhanced predictive ability during recessions.

The next type of model allows for underreaction and overreaction to information. Due to behavioral biases, Hong and Stein (1999) explain that, at the start of a trend, investors underreact to news; as the market rises, investors subsequently overreact, leading to even higher prices. Similarly, positive feedback traders – who buy (sell) after asset prices rise (fall) – can create price trends that technical indicators detect. George Soros (2003) argues that positive feedback can alter firm fundamentals, thereby justifying to a certain extent the price trends. Edmans, Goldstein, and Jiang (2012) show that such feedback trading can occur in a rational model of investors with private information.

The last model outlined in Neely (2010) shows that models of investor sentiment shed light on the efficacy of technical analysis. Since Keynes (1936), researchers have analyzed how investor sentiment can drive asset prices away from fundamental value. DeLong, Shleifer, Summers, and Waldmann (1990) show that, in the presence of limits to arbitrage, noise traders with irrational sentiment can cause prices to deviate from fundamentals, even when informed traders recognize the mispricing.

These papers seek to explain why technical indicators may forecast the equity premium. Neely outlines that theoretical models based on information frictions help to explain the predictive value of the technical indicators. Empirically, Moskowitz, Ooi, and Pedersen (2012) find that pervasive price trends exist across commonly traded equity index, currency, commodity, and bond futures. Neely notes that insofar as the stock market is not a pure random walk and exhibits periodic trends, technical indicators should prove informative, as they are primarily designed to detect trends.

Lo, Mamaysky, Wang (2000) state that technical analysis, also known as “charting” has been a part of financial practice for many decades, but the discipline has not received the same level of academic scrutiny and acceptance as more traditional approaches, such as fundamental analysis. The article tests subjective chart patterns such as head and shoulders and double bottom which confirms some information may be obtained. This is backed up by literature in the late 1960s and 1980s by Fama and Blume (1966), and later RJ Sweeney (1988) which show the predictive power of filter rules in cross-sectional stock returns.

Chang and Osler (1994) discuss the first study of nonlinear visual chart patterns in academic literature. Pulling the rules of the head and shoulders pattern from trade journals. The pattern was found to be predictive in the cross-section of forex markets. Chang and Osler (1994) note that there are more nonlinear visual patterns than just the Head and Shoulders pattern which are outlined in Pring (1988). The patterns are said to be highly predictive according to practitioners. Chang and Osler left additional patterns for future work. Some of the patterns are picked up in Lo, Mamaysky, and Wang (2000).

From Chang and Osler (1994) Results show the head-and-shoulders trading rule appears to have some predictive power for the German mark and yen but not for the Canadian dollar, Swiss franc, French franc, or pound. Nonetheless, if one had speculated in all six currencies simultaneously, profits would have been both statistically and economically significant. Taken individually, profits in the markets for yen and marks are also substantial when adjusted for transactions costs, interest differentials, or risk. These results are robust to changes in the parameters of the head-and-shoulders identification algorithm, changes in the sample period, and the assumption that exchange rates follow a GARCH process rather than a random walk. These results are inconsistent with virtually all standard exchange rate models and could indicate the presence of market inefficiencies.

Technical analysis, the prediction of price movements based on past price movements, has been shown to generate statistically significant profits despite its incompatibility with most economists' notions of "efficient markets." In the stock market, excess profits based on technical trading rules are documented by Brock, Lakonishok, and LeBaron (1992), and in the foreign exchange market such excess profits are found by Dooley and Shafer (1984); Logue, Sweeney, and Willett (1978); Sweeney (1986)l; and Levich and Thomas (1993).

Tests of technical analysis have largely limited their attention to techniques that are easily expressed algebraically, namely filter rules and moving average. Practitioners, however, rely heavily on many other techniques, including a broad category of exclusively visual patterns. Typically known by fanciful names, this category includes "head and shoulders," "rounded tops" and "bottoms," "flags," "pennants," and "wedges." Highly nonlinear and complex, trading rules based on these patterns normally cannot be expressed algebraically.

The purpose of this paper is to begin evaluating this large set of visual nonlinear trading rules by focusing on one of the best-known patterns, head, and shoulders. Technical analysts claim that this pattern identified when the second of a series of three peaks is higher than the first and the third, presages a trend reversal. The computer-based identification algorithm locates such patterns using local maxima and minima.

Since the head-and-shoulders pattern is considered by practitioners to be one of the most, if not the most, reliable of all chart patterns, it represents a natural point of departure for empirical research. If trading based on this pattern generates excess profits, investigating other patterns may prove interesting. Conversely, if profits are insignificant, then this entire branch of visually based technical analysis may be called into question.

The authors test the head-and-shoulders rule on daily spot rates for six currencies against the dollar: the yen, German mark, Canadian dollar, Swiss franc, French franc, and pound. Their data cover the entire floating rate period (from March 19, 1973, to June 13, 1994), a twenty-one-year span that provided the authors with more than 5,500 daily observations. Currency markets seem especially appropriate for testing technical signals, as they are characterized by very high liquidity, low bid-ask spreads, and round-the-clock decentralized trading. Furthermore, because of their size, these markets are relatively immune to insider trading. In any event, technical analysts claim that "the principles that underlie the analysis of currencies from a technical aspect are basically the same as those used in any other financial market or for individual stocks" (Pring 1985, p. 466). The authors are aware of only two studies that evaluate – for any market – the visual, nonlinear patterns that are the focus of this paper. The two studies come to different conclusions: Levy (1971) tests the predictive power of all thirty-two possible five-point chart patterns, including the head and shoulders. He finds no evidence of profitable forecasting ability. As discussed later, the authors question the validity of the results. Brock et al. (1992), find that breakouts from observed trading ranges are meaningful predictors of short-term returns in the Dow Jones Index during 1897-1986, a result corroborating technicians' claims regarding "support" and "resistance" levels. In short, research on these visual trading patterns is both scarce and inconclusive; thus, as Neftci (1991) notes, these visually based strategies are currently "a broad class of prediction rules with unknown statistical properties."

The study can be viewed as contributing to a growing body of research testing for nonlinear dependence in financial prices. Early tests for the presence of nonlinearities, testing the null hypothesis of i.i.d. behavior, indicate that nonlinearities are indeed present in stock markets Hsieh 1991 and in floating exchange rates (Hsieh 1989). The form of these nonlinearities remains unclear. Modeling financial prices as a GARCH process seems to capture some of the nonlinearities indicated by more general tests; more specifically, it is helpful for predicting volatilities (Hsieh 1989). Other sources of nonlinearity are also consistent with the data. Another potential source of nonlinearity is chaos, although a few available tests fail to confirm its existence in exchange rate data. These tests may be helpful in identifying another specific form of exchange rate nonlinearity that is consistent with data on floating exchange rates.

There are three parts to the methodology by which the authors calculate and interpret the profits earned by taking foreign exchange positions once the pattern is recognized. The first part is an objective, computerized, identification of the head-and-shoulders pattern itself. The second part is a strategy, replicable in real time without the knowledge of the future, for entering and exiting speculative positions after recognizing such patterns. The third part concerns evaluating whether the profits obtained from this trading rule imply that there were predictable profit opportunities in the data.

The approach in this paper is to evaluate whether these profits are statistically greater than those that would have been found had there been no intertemporal dependence in exchange rate changes.

They identify reliable confidence intervals via the bootstrap methodology, implemented by constructing 10,000 new exchange rate series. In each simulated series, daily changes are determined by drawing randomly, with replacement, from the original series of exchange rate changes and applying these changes consecutively to the exchange rate's actually starting value. When we apply our trading rules on these constructed series, we obtain a distribution for profits under the null hypothesis that there are no predictable profit opportunities. From this distribution, we calculate confidence intervals.

Technical nonlinear patterns are different from moving averages (and patterns comprised of moving averages) in that they are harder to test arithmetically. Lo, Mamaysky and Wang (2000) provide a framework for testing such patterns, using the pattern rules from Edwards and McGee (1966). These are largely based on kernel density estimation and polynomial regressions. This is discussed in the next section. A subset of the visual patterns (discussed later) are tested by Lo, Mamaysky, and Wang (LMW, 2000) for both predictability and informational value. Noting both the discrimination from academic literature on technical studies and how difficult it is to compute them arithmetically. The authors further discuss the ease at which to rebalance a portfolio using modern quantitative finance techniques compared with the difficulty to objectively obtain and test a visual pattern such as the head-and-shoulders. Despite their validity and viability, they have a bad wrap in the literature. Often being referred to as being as useful as astrology, or similar to alchemy. (LMW, 2000), Covel (2009).

LMW (2000) find out that the patterns have predictive ability, and contain some informational value. Citing previous studies by Blume, Easley, and O'Hara (1999) which consider the joint distribution between prices and volume containing predictive power. Academia seems to be okay with prices and volume being predictive, but not prices and past prices (LMW, 2000).

The paper by LMW (2000) was groundbreaking in that being published in a top finance journal, it showed that visual patterns are worth their salt and are predictive in the cross-section of both the NYSE/AMEX and Nasdaq. The authors are careful to point out that predictability does not necessarily mean profitability.

Future work by LMW (2000) is noted to come up with more rigorous ways to develop the patterns. The authors rely on smoothing estimators (discussed in a later section) and call for the use of a local polynomial kernel. They also call for new ways to obtain the patterns. We address both.

Our paper picks up where LMW (2000) left off in first computing an objective method for finding the patterns and by contributing new methods to uncover the patterns (i.e. digital image processing) which is discussed in further sections. The first methodology deals with detecting the patterns objectively (LMW (2000), relied heavily on visual inspection). The second part discusses whether image processing can detect the patterns in data which is the cornerstone for our research question (can image processing be a useful tool for economists?). Once we know what the patterns look like, and how to detect them objectively and arithmetically, we can use them for benchmarking on new methods.

3. Nonparametric Estimation

Hurvich and Simonoff (1998) discuss nonparametric estimation of an unknown smooth regression. The methodology was popular in the literature around the time of Lo, Mamaysky, and Wang (2000).

Nonparametric data assumes we have data $y = (y_1, \ldots, y_n)'$ generated by the model

$$y_i = m(x_i) + \epsilon_i \ \forall \ \text{i=1,…,n} \tag{1}$$

Where m(.) is an unknown smooth function, the $x_i$are given real numbers in the interval [a,b] and the ϵ_i are independent random variables with mean 0 and variance σ_0^2. Either the predictor vector x is non-random, or the y is conditional on the observed values if it is random. In our case, x is a state variable for prices.

Hurvich and Simonoff (1998) note that many different estimators of m have been proposed. In research by Lo, Mamaysky, and Wang (2000) the authors call for the use of local polynomial estimators in finding smooth data from prices.

A pth-order local polynomial estimator is defined as the constant term $\beta^{\wedge}{}_0$ of the minimizer of

$$\sum_{i=1}^{n} \{y_i - \beta_0 - \ldots - B_p(x - x_i)^p\}^2 K(\frac{x - x_i}{h}) \tag{2}$$

Where k is the kernel function, generally taken to be a symmetric probability density function with finite second derivative (Hurvich and Simonoff (1988)). Typical choices of p are 0,1,2 and 3. We use local constant (p=0). There are said to be asymptotic and boundary bias correction advantages with local linear (p=1) and local cubic (p=3) estimators over local constant (p=0) and local quadratic (p=2). We use a local constant estimator following Lo, Mamaysky, and Wang (2000).

Other estimators include a Gasser-Muller convolution kernel estimator (Gasser and Muller, 1979) and smoothing splines.

A crucial step in estimating m is choosing the smoothing parameter (h for the local polynomial and kernel estimators), which controls the smoothness of the resultant estimate. Automatic smoothing parameter selectors generally fall into two broad classes of methods: classical and plug-in approaches. Classical methods are based on the minimization of an approximately unbiased estimator of either the mean average squared error

$$MASE = \frac{1}{n}E[(m^{\wedge}{}_{h} - m)'(m^{\wedge}{}_{h} - m)] \quad (3)$$

(e.g. generalized cross-validation (GCV); Craven and Wahba (1979)) or Akaike information criterion (AIC); Akaike (1973)).

There are many kernel estimators to use when doing this. A Gaussian kernel is the most common approach. In contributing to their work we use a local polynomial regression. This is discussed below.

Smoothing data to find the patterns involves using kernel density estimators to find extrema on the prices. Local polynomial regression is what we use to carry this out. There are many different kernels to choose from, the Gaussian kernel is a popular choice. Other kernels include cosine, rectangular, sigmoid to name a few. (From Kedd (2015)). To adjust the degree of smoothness we set a bandwidth (h) in our function. Härdle (1990). These are referred to also as penalty functions. Penalty functions and bandwidth selection are also discussed in Hurvich and Simonoff (1998), and Kedd(2015).

Turluch - Bandwidth Selectors for KDE - popular approach is to visually select the best one. Härdle, Linton (1994)- Kernel Density Estimators, different kernels for smoothing. Gaussian is more popular.

Choice of H is more important than the kernel density estimator itself. We don't know the bandwidth because we don't know the density of our function. So we have to compute the optimal value. This is the tradeoff between computing optimal bandwidth and plug-in methods. When we plug in we inspect the data for how smooth we want it. So, if we had a window where we know there is a chart pattern we can visually inspect h to fit the data so that it meets our rule criteria. Turlach (1993) discusses how this is not replicable. Which is what we find when we try and use just cross-validated (objective) bandwidth. Our estimates are off by a few days etc. LMW (2000) calls for updated methods to uncover the patterns objectively, mainly select h in a replicable way. When we do this, our price becomes an approximation of the price using kernel smoothers and an error term.

Lo, et al (2000) directly call for new more rigorous methods for uncovering the patterns based on their methodology in their introduction. They also call for a new methodology for uncovering the patterns. We discuss both. The former in this chapter, the later in the next.

Minimization Functions Golden-section finds extremum by narrowing the range of values inside which the extremum is known to exist. The name comes from maintaining the function values for triples of points with distance from the golden ratio. This is the limit of the Fibonacci search. Newton's method in optimization is derived from calculus in finding the zeros of a known function, starting at an initial guess. This follows the methodology above and makes sure that our objective bandwidth can match the data from the original paper.

Following Lo, Mamaysky, and Wang (2000) we rescale our variable K(u) to follow the probability density function.

$$K_h(u) \equiv \frac{1}{h} K(\frac{u}{h}), \int K_h(u) du = 1 \tag{4}$$

If h is very small, the averaging will be done to a small neighborhood (small number of observations) around $X_t s$. If h is very large, the averaging will be done over larger neighborhoods of the $X_t s$.Therefore controlling the amount of averaging reflects the degree of smoothness, and thus adjusting the smoothing parameter h, this is referred in the literature (Scott and Sain 2005, Lo Mamaysky Wang (2000)) as the bandwidth.

Some kernel estimators we use in this paper are in the table below. The equations come from Applied Statistics Lecture Notes by M. de Carvalho and J. Blanchet. Institute of Mathematics, Analysis, and Applications. EPF Lausanne.

A kernel estimate $m^{\wedge}{}_{h}(x)$ will converge to m(x) with a small enough bandwidth. This is shown by Härdle (1990). This holds true for many estimators. The most popular choice of the kernel is the Gaussian Kernel and is what we use in this paper.

Popular bandwidths are shown to be based on the trace of the smoothing parameter h. These are summarized in the table below.

AICC	$1+\frac{2(trH[h]+1)}{n-trH[h]-2}$
GCV	$-2Log[1-\frac{trH[h]}{n}]$
AIC	$2\frac{trH[h]}{n}$
T	$-Log[1-2\frac{trH[h]}{n}]$

Table 1 Possible Bandwidths

Plug in estimators can also be used. These involve visually inspecting the data and then fine-tuning the parameter h for how smooth it should be. These are probably the best for these estimators followed by AICC. According to Hurvich and Simonoff (1998). The drawback of plug-in estimators is that it isn't automated.

These are used in the variety of kernel functions. A popular choice being a Gaussian kernel because it simulates a normal distribution on each observation x_i. Other popular choices are the Epanechnikov kernel.

Kernel	Formula
Biweight	$K(u) = \frac{15}{16}(1-u^2)^2 I\{\lvert u\rvert \le 1\}$
Triweight	$K(u) = \frac{35}{32}(1-u^2)^3 I\{\lvert u\rvert \le 1\},$
Cosine	$K(u) = \frac{\pi}{4} cos(\frac{\pi}{2} u)$
Epanechnikov	$K(u) = \frac{3}{4}(1-u^2) I\{\lvert u\rvert \le 1\},$
Epan2	$K(u) = \frac{3}{4}(1-u^2) I\{\lvert u\rvert \le 1\},$
Gaussian	$K(u) = \frac{1}{2\sqrt{2\pi}} e^{-\frac{u^2}{2}}, u \in R$
Rectangular	$K(u) = \frac{1}{2}\{\lvert u\rvert \le 1\}$
Triangular	$K(u) = (1-\lvert u\rvert) I\{\lvert u\rvert \le 1\},$
Sigmoid	$K(u) = \frac{2}{\pi}\frac{1}{e^u + e^{-u}}$

Table 2 Possible Kernels

Selecting the bandwidth is as important, if not more important than choosing the kernel (Scott and Sain (2005), Lo, Mamaysky, Wang (2000)). The success of the smoothing parameter is dependent on h. If it is too small, it will fit exactly. If it is too large it will be too smooth and be a poor estimator. The penalty for too smooth is that there will be no extrema found. On the extreme another end, it will be too many extrema which according to Lo, Mamaysky, Wang (2000) the reason for not finding extrema from price is that it has too many false positives.

We show the ability of our local polynomial kernel, and AIC bandwidth on simulated data. We use an approximation Z_t for a normal distribution, with mean zero and $\sigma = 1$. We use $X_t$to approximate a uniform distribution between 0 and 2π. We take 500 draws of each and use E[Y|X] ≡ $Y_t = Sin(X_t) + 0.5Z_t$. We show simulated data for our sample in the plot below.

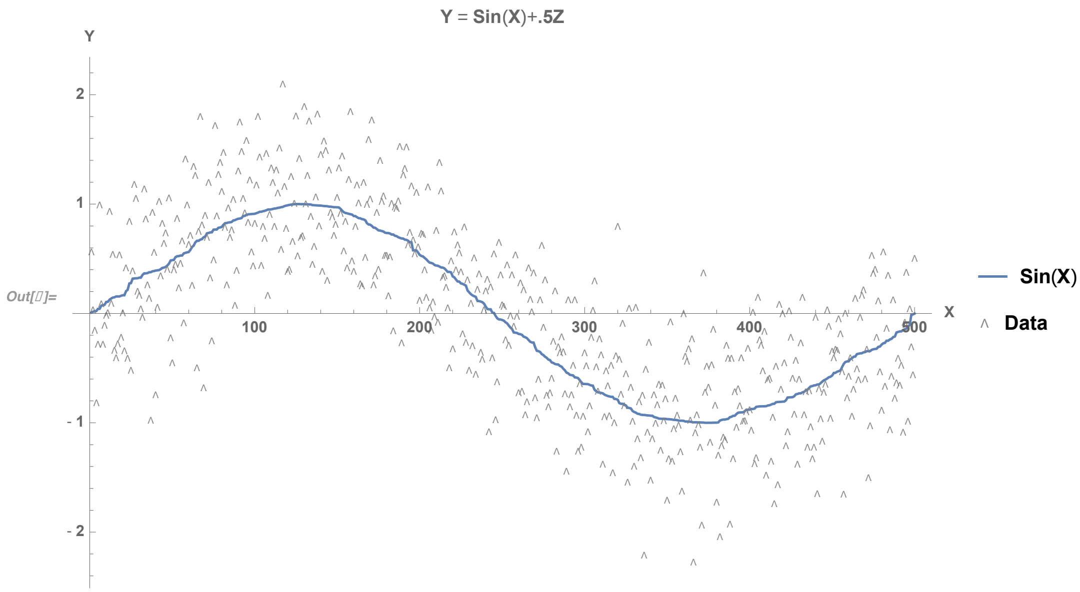


Figure 1. Random data following a normal distribution with sin approximation.

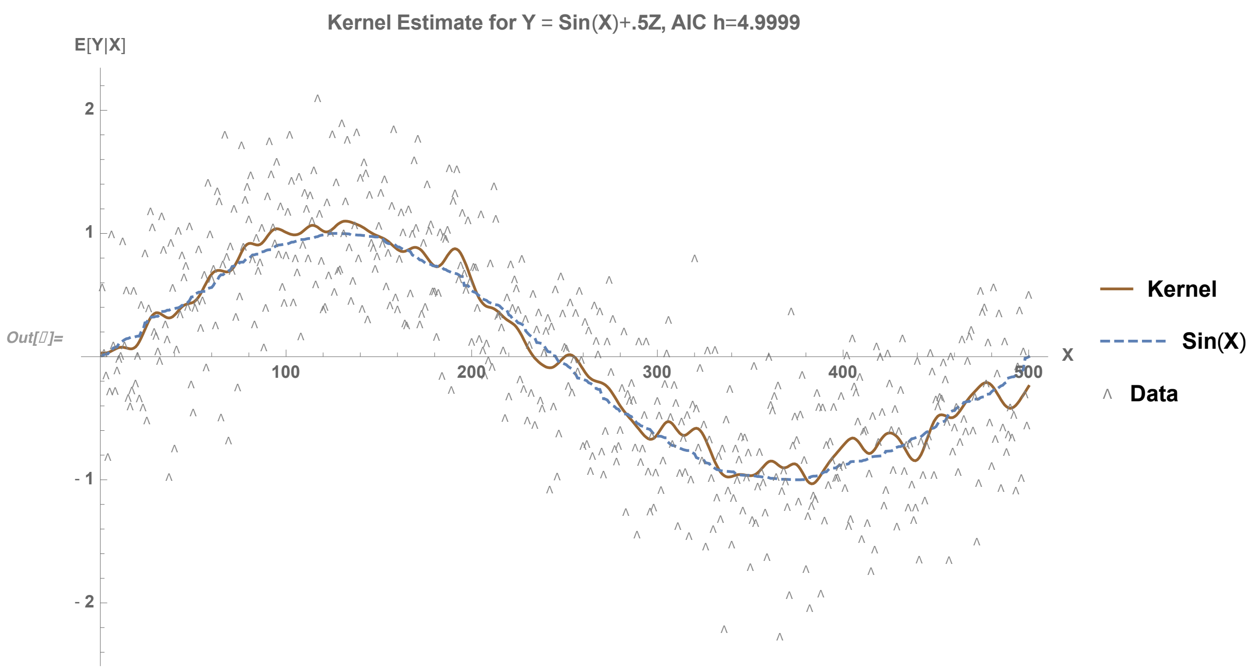


Figure 2. Random data with sin approximation and AIC bandwidth

Akaike Information Criteria (Akaike, 1973) sets the bandwidth to be too small. It fits both the "noise" $0.5\epsilon Z_t$ and also the "signal" Sin(.).

The normal reference rule using a normal kernel is h = 1.06 $\sigma n^{\frac{-1}{5}}$ for univariate data. (Scott (2010)) this is referred to as Silverman's rule of thumb and yields a larger bandwidth and fits the function better.

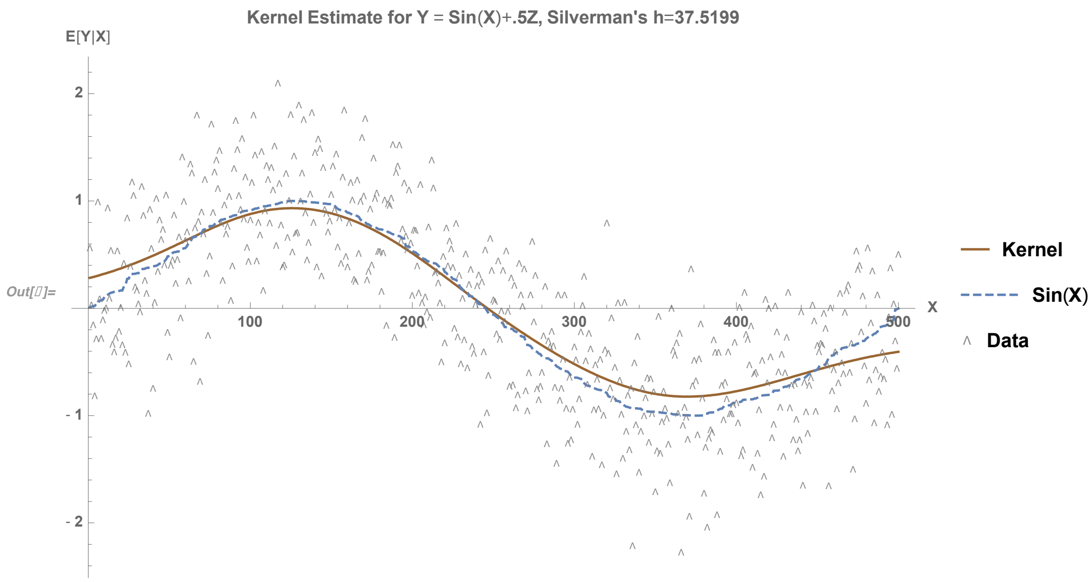


Figure 3. Random data with sin approximation and Silverman's bandwidth.

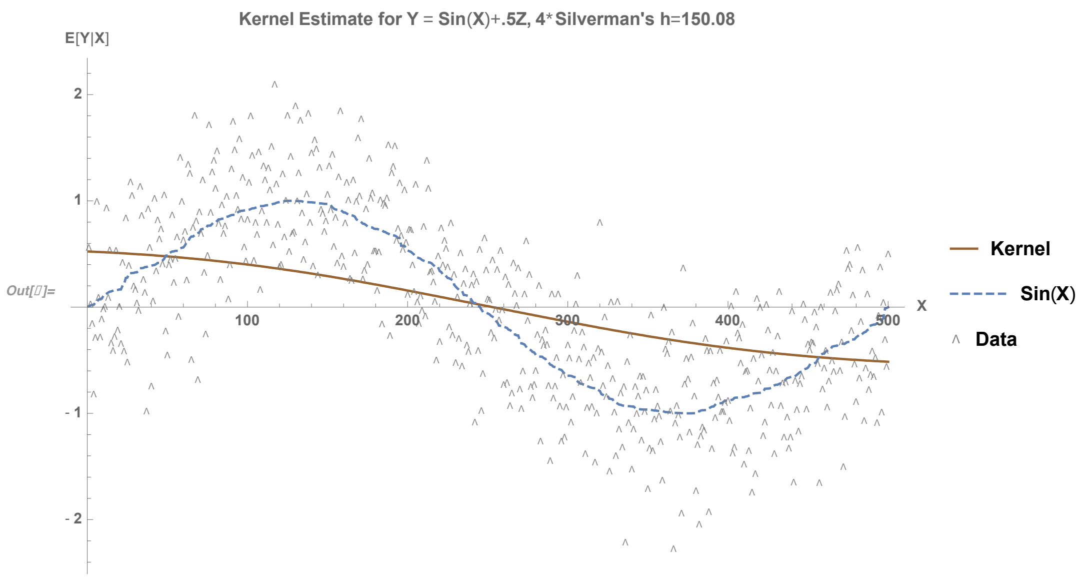


Figure 4. Random data with Silverman's bandwidth and Sin approximation.

If we take too large of a bandwidth, say by increasing it to a factor of 4 * Silverman's Rule of thumb we get too smooth of an estimator.

These are referred to Silverman's rule of thumb. A popular method is a cross-validated function.

$$CV(h) = \frac{1}{T}\sum_{t=1}^{T}(P_t - m^{\wedge}{}_{h,t})^2, \tag{5}$$

$$m^{\wedge}{}_{h,t} \equiv \frac{1}{T}\sum_{T\neq t}^{T}\omega_{T,h}Y_T \tag{6}$$

The estimator $m^{\wedge}{}_{h,t}$ is the kernel regression estimator applied to the price history $\{P_t\}$ with the tth observation omitted, and the summands are the squared errors (Following LMW (2000)), of the $m^{\wedge}{}_{h,t}$s following each omitted observation. For a given bandwidth parameter h, the cross-validated function shows how well the kernel estimator can fit the data observation $P_t$when that observation is not used to construct the kernel (it is omitted). By selecting the bandwidth that minimizes this function we obtain asymptotic properties such as minimum mean-squared error. According to Lo, Mamaysky, Wang (2000) the bandwidth used by a Cross-Validated approach is too smooth according to the bias of trained technical analysts. They found that 0.3*h remedies this problem but note there are other bandwidths that fix the solution, however, change the end result overall (Härdle (1990)).

Upon reconstruction, we find that 0.3*h doesn't match the results from their paper entirely. Lo, Mamaysky, Wang (2000) note that a promising direction for future work is to consider alternatives to kernel regression. They discuss the viability of local polynomial regression, in that the former lack local variability in the degree of smoothing. Local polynomial regression provides local averaging of polynomials to obtain the estimator m(x). They discuss that these may yield important improvements in the pattern recognition problem. We address this in our paper.

For minimization functions, we use the Golden Section. This is based on the Fibonacci sequence for the search for global maxima and minima in a function. Rather than using a rate of change as in Newton's Method.

The Fibonacci numbers 1,1,2,3,5,8,13 have used in areas of mathematics, natural science, nature, and technical analysis. Their formation is such that the current number is the rest of the previous number, plus the one before it. Starting with 0 and 1.

$$F_n = f_{n-1} + f_{n-2} forn > 2 \tag{7}$$

$$f_2 = f_1 = 1 \tag{8}$$

The ratio of two consecutive Fibonacci numbers $\frac{f_{n+1}}{f_n}$approaches the golden ratio for large n.

With this formula, the first Fibonacci ratios can be calculated.

$$lim_{n\to\infty} \frac{f_n}{f_{n+k}} = lim_{n\to\infty} \frac{f_n}{f_{n+1}} \frac{f_{n+1}}{f_{n+2}} \dots \frac{f_{n+k-1}}{f_{n+k}} = \phi^{-k} = (\frac{1+\sqrt{5}}{2})^{-k} \tag{9}$$

$$f_0 = (\frac{1+\sqrt{5}}{2})^{-0} = 1 \tag{10}$$

$$f_1 = (\frac{1+\sqrt{5}}{2})^{-1} = 0.6180 \tag{11}$$

$$f_2 = (\frac{1+\sqrt{5}}{2})^{-2} = 0.3820 \tag{12}$$

Lo, Mamaysky, and Wang (2000) discuss that finding extrema off of price is too noisy and resultant patterns would be impossible to distinguish from noise.

4. Automating Technical Analysis.

Lo, Mamaysky, and Wang (2000) define patterns in terms of their ability to meet rules based on local extrema (maxima and minima). They obtain kernel estimators $m^{\wedge}(.)$ of the price for a given time series so that the extrema can be determined numerically. They then analyze the smoothed estimate $m^{\wedge}(.)$of the price for the occurrence of technical patterns.

Lo, Mamaysky, Wang (2000) note that the last two steps are straightforward applications of kernel regressions. It is the first step which is estimating to be the most controversial because here it is mimicking the judgment of a trained professional analyst. We do not have a trained professional analyst handy so our best estimate is to find a solution that replicates Lo, Mamaysky, Wang (2000) as best as possible. If they do not meet their scrutiny, (i.e. the bandwidth selected does not yield good patterns) it may not be a reliable estimate of what they would see in practice and thus make decisions from.

Pattern recognition is known to identify handwritten digits, faces, distinguish between animals, breed of dogs, objects, hands, fingerprints, self-driving cars, and other human activities, it is difficult to fully capture what a trained analyst would see in a stock chart.

We propose definitions of 10 technical patterns based on their extrema. These are from Lo, Mamaysky, and Wang (2000) and ultimately from Edwards and Magee (1966, Chaps. VII-X): Other print literature (Pring, (2002), Elder (2002) discuss the patterns and show examples of various forms but do not develop objective rules for testing. It is possible that these other patterns, along with additional patterns outlined in Chang and Osler (1995) might make their way into literature at some point.

Using this procedure, and armed with our Gaussian Kernel in the section above with local polynomial regression of order 0 we find results that match with the original samples in Lo,

Mamaysky, Wang (2000). These are shown below, with the original results for reference. Black and white images are from Lo, Mamsky, Wang (2000).

We first smooth the price data in rolling 38-day windows. We use 35 days for recognizing extrema and three additional days for recognizing a completed pattern.

A computer would only need 36 days to recognize a pattern as only one day on either side of the last extrema is needed. The three additional days is for a ‘human’ factor. This follows LMW (2000).

We smooth price using the local polynomial kernel of order 0 and a Gaussian kernel. We use a modified form of the cross-validated bandwidth, named Maximum Price Deviation.

We first smooth the data, check for extrema, then check whether there are enough extrema to complete a pattern. If there are, we check whether they fit any of the pattern definitions defined above. If they do, we log the location of extrema, kernel values, price values, and plot both for observation.

Pattern windows without enough extrema to complete a pattern definition, or windows with enough extrema but do not meet the pattern rules are discarded. We have roughly 4800 non pattern windows and fewer than 70 completed pattern windows for the CTX security which spans a decade short of the full data set (1967-2014). We use this security to match to the original paper by Lo, Mamaysky, Wang (2000). The patterns in rolling windows are shown below. Patterns are found using maximum price deviation, with maxmiss set between 0.01 and 0.02.

What Lo was doing was saying look, this pattern meets the rules. Let’s not be concerned with hitting every possible example of every pattern but populate enough examples to test if they have predictive power. He finds they do.

What we want to do is take that and say let’s use computer vision techniques to teach a computer to better recognize the patterns the way a human would see them.

We cross check everything in excel. We basically fit our windows with a kernel smoother using our bandwidth and we say look, here’s the pattern where’s the extrema. And we put each pattern through the rules in the paper to see if it meets them, and we see that it does. The key here is that it was all automatic, we didn’t have to visually inspect our bandwidth (h).

Head and Shoulders:	Inverse Head and Shoulders:
E1 is a maximum E3 > E1, E3 > E5 E1 and E5 are within 1.5 percent of their average E2 and E5 are within 1.5 percent of their average,	E1 is a minimum E3 < E1, E3 < E5 E1 and E5 are within 1.5 percent of their average E2 and E4 are within 1.5 percent of their average.
Broadening Top:	Broadening Bottom:
E1 is a maximum E1 < E3 < E5 E2 > E4	E1 is a minimum E1 > E3 > E5 E2 < E4
Triangle Top:	Triangle Bottom:
E1 is a maximum E1 > E3 > E5 E2 < E4	E1 is a minimum E1 < E3 < E5 E2 > E4
Rectangle Top:	Rectangle Bottom:
E1 is a maximum tops are within 0.75 percent of their average bottoms are within 0.75 percent of their average lowest top > highest bottom,	E1 is a minimum tops are within 0.75 percent of their average bottoms are within 0.75 percent of their average lowest top > highest bottom

Table 3 Pattern Rules

The definition for Double Tops / Bottoms is slightly more involved. There are only two required tops/bottoms, E1, and E alpha. These must occur at least 22 days apart.

Double Top:	Double Bottom:
E1 is a maximum E1 and Ea are within 1.5 percent of their average ta* - t1* > 22	E1 is a minimum E1 and Ea are within 1.5 percent of their average ta* - t1* > 22

Table 3 Pattern Rules

Following Lo, Mamaysky, Wang (2000) our algorithm uses a sample of prices $\{P_1, \ldots, P_T\}$. We fit kernel regression in these windows, one for each subsample or window from t to t+l+d-1, where t varies from 1 to T-l-d+1, and l and d are fixed. Lo, Mamaysky and Wang (2000) use l=35 and d=3, so each window consists of 38 trading days.

We fit each window with a kernel regression to narrow our attention to just patterns that are completed within the span of the window -l+d trading days. If we fit a single kernel to the entire dataset, many patterns of various durations would emerge, and without additional constraint, it would be impossible to distinguish signal from noise. The window length is fixed at l+d, but kernel regressions are estimated on a rolling basis and we search for patterns in each window, following Lo, Mamaysky, and Wang (2000).

For a fixed window, we can only find patterns that are completed within l+d trading days. Future work is still left to discover patterns on a longer horizon. We focus on 35 days to follow the paper by Lo, Mamaysky, and Wang (2000).

The parameter d controls for the human effect of not recognizing patterns exactly when they occur. We allow for a lag between pattern completion and the time it takes to detect the pattern. To do this we require that the final extremum that completes a pattern occurs on day t+l-1; d is the lag time that takes place between the pattern being completed and recognizing it. In a later section we compute post pattern returns on day t+l+d, that is, one day after the pattern has completed. An example being if a head-and-shoulders pattern has completed on day t+l-1 (having used prices from time t through time t+l+d-1, we compute the conditional one-day gross return as $Z_1 \equiv$

$\frac{Y_{t+l+d+1}}{Y_{t+l+d}}$. We do not use any forward information in computing returns conditional on pattern completion. The lag d ensures that we are computing our conditional returns out-of-sample and without any "look-ahead" bias. This follows (Lo, Mamaysky, Wang (2000)).

Within each window, we estimate a kernel regression using the prices in that window, hence:

$$m^{\wedge}{}_{h,t} \equiv \frac{\sum_{s=t}^{t+l+d-1} K_h(\tau - s)P_s}{\sum_{s=t}^{t+l+d-1} K_h(\tau - s)}, t = 1, \ldots, T - l - d + 1 \tag{13}$$

$K_h(z)$ is given earlier and h is the bandwidth parameter. $m^{\wedge}{}_h(\tau)$is a differentiable function of τ.

Once the function has been approximated, we can compute the extrema in the rolling window. We use the golden section which is a search for minima. Newton's method is also an alternative for finding extrema. Lo, Mamaysky, and Wang (2000) find extrema using derivatives. For derivatives, we find the times τ such that Sgn($m^{\wedge\prime}{}_h(\tau)$)=-Sgn($m'^{\wedge}{}_h(\tau + 1)$) where $m'^{\wedge}{}_h(\tau)$ is the derivative of $m^{\wedge}{}_h$ with respect to τ and Sgn(.) is the signum function. If the signs of $m'^{\wedge}{}_h(\tau)$ and $m'^{\wedge}{}_h(\tau + 1)$ are +1 and -1 respectively, then we have found a local maximum, and if they are =1 and +1, respectively, then we have identified a local minimum. The same holds for any method of finding extrema.

Once we identify extrema we identify whether it is maxima or minima in the price series P_t in the window range [t-1,t+1], and the original price is used to determine whether a pattern has occurred according to our definitions from Edwards and Magee (1966).

If $m'^{\wedge}{}_h(\tau)$=0 for a given τ, which occurs when closing prices stay the same for several consecutive days, we check whether the price we have is a local maximum or minimum. We do this by looking for the date s such that s = inf{s> τ: $m^{\wedge\prime}{}_h(s)0$}. We ignore the similar prices and move to the next logical consecutive price. Here we apply the same methodology as above, except we compare Sgn($m'^{\wedge}{}_h(\tau - 1)$ $and\ Sgn(m^{\wedge\prime}{}_h(s))$.

One consequence of this algorithm as noted by Lo, Mamaysky, Wang (2000) is that the series of extrema that it identifies contains alternating minima and maxima. That is if the kth extrema is a maximum, then the (k+1) will always be a minimum and vice versa.

The advantage of using this approach is that it ignores noisy price data. For example, we could identify extrema every time price changes direction. For example, a price P_t is identified as a local extremum when $P_{t-1} < P_t$ and P_t>P_{t+1}. Notice how we need one day on either side of the extrema for it to be recognized. Using this methodology would yield too many extrema and too many patterns to meet be visually consistent with the patterns that technical analysts find compelling.

Once we have identified all of the local extrema in the window [t,t+l+d-1], we check whether the window has enough extrema to complete a pattern, if it does we proceed to check whether it meets any of the pattern definitions above. We then move to the next window [t+1,t+l+d] and we continue until the end of the sample [T-l-d+1,T].

Empirical Examples (Patterns Matched to Lo, Mamaysky, Wang (2000)).

To see how our algorithm matches up to the patterns from prior work, we use bandwidths mentioned above and the local polynomial kernel of order 0. We use an AICC bandwidth to find the patterns.

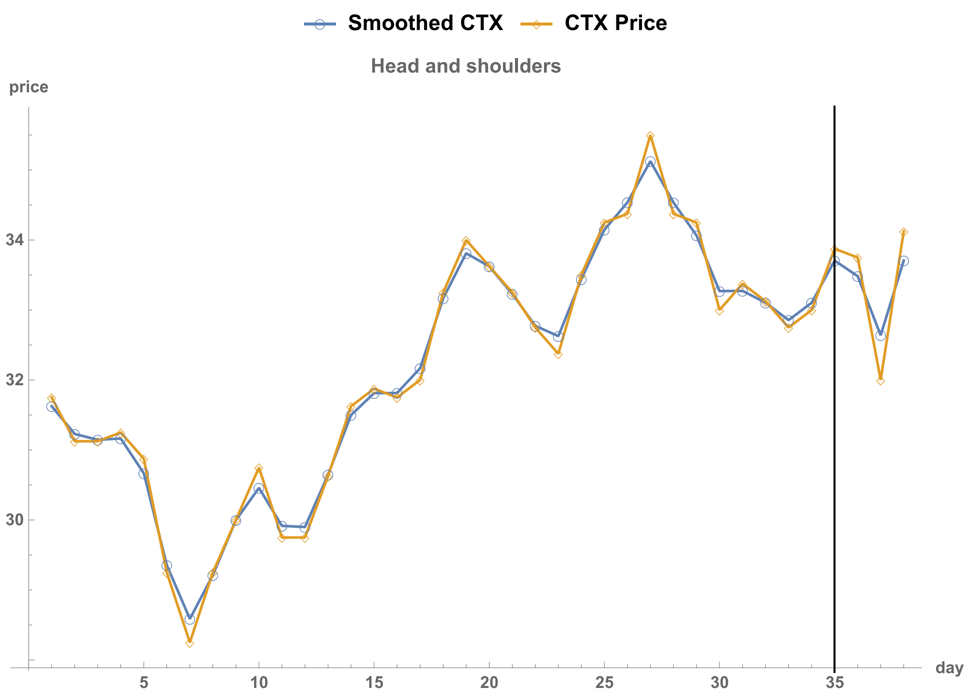
We follow the methodology of Lo, Mamaysky, and Wang (2000) and apply the algorithm to the daily returns of a single security, CTX, during the five-year period from 1992-1996. Inspection of the paper by Lo, Mamaysky, and Wang(2000) will show similar patterns as what we find below.

In each graph, the blue line is the raw price and the orange line is the kernel estimator (smoothed price) $m^{\wedge}{}_{h}\,(.)$, the circles locate price observations and the diamonds denote kernel estimations of prices P_t. The verticle line denotes the 35^{th} day, which is the last day for finding extrema. Day t+l-1.

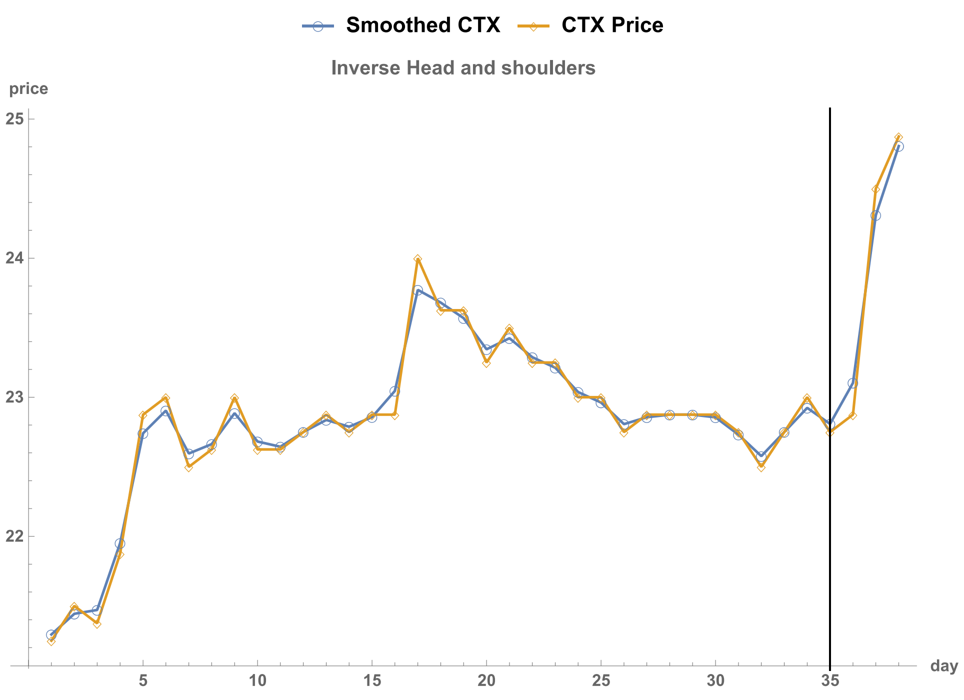
5. Empirical Examples

Lo, Mamaysky, and Wang (2000) confirm the ability of their process to meet their recommendation which is the human judgment in identifying the patterns. Our algorithm is able to match theirs. Using local polynomial regression, Gaussian Kernel, and AICC bandwidth. This meets part one of their recommendation for future work. Our patterns are checked and crossreferenced for meeting the same extrema dates. We obtain the same result through different methodology.

We cross check our patterns in excel. We show that the extreme points on the kernels do result in extrema and meet the pattern rules. Double tops and bottoms are harder to test this way. The first 8 patterns match up. It follows that our kernel values to check out with the extrema as provided in Mathematica. We show the pattern matching from the 5 pattern pairs outlined in Lo, Mamaysky, Wang (2000) below.

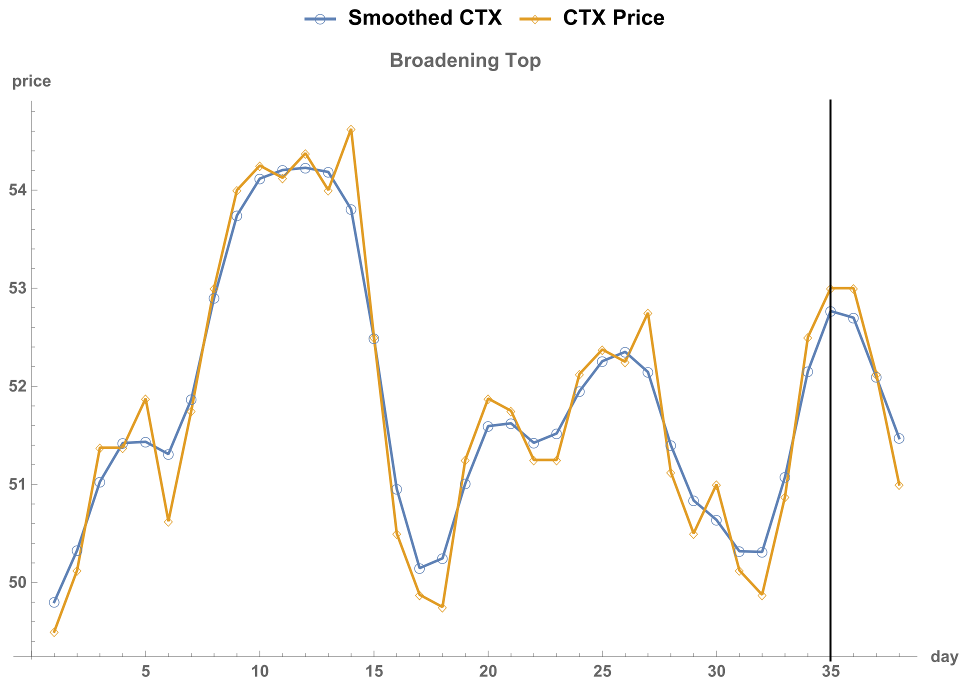


(a)

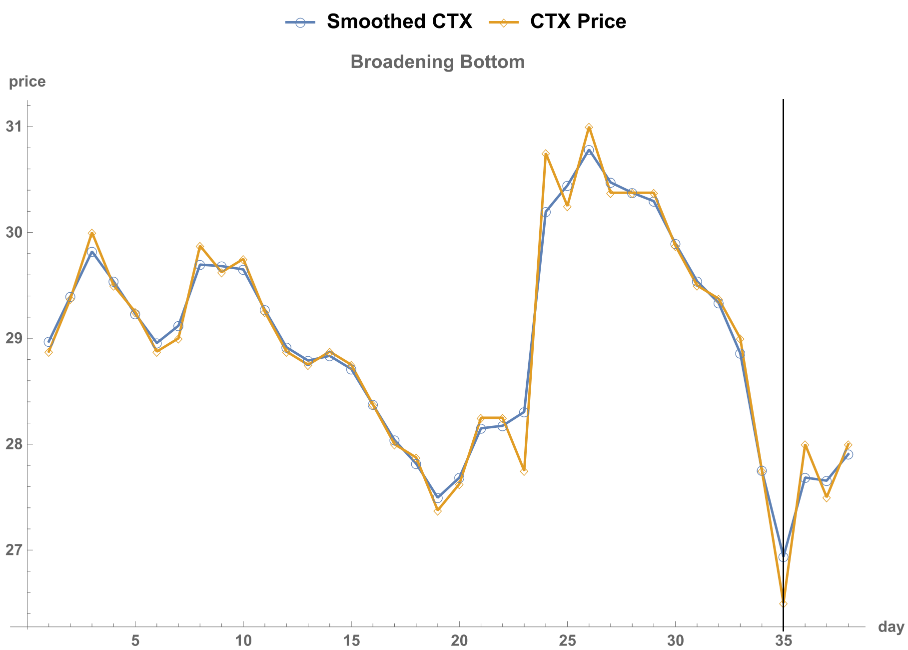


(b)

Figure 5.Head-and-Shoulders and Inverse Head-and-Shoulders

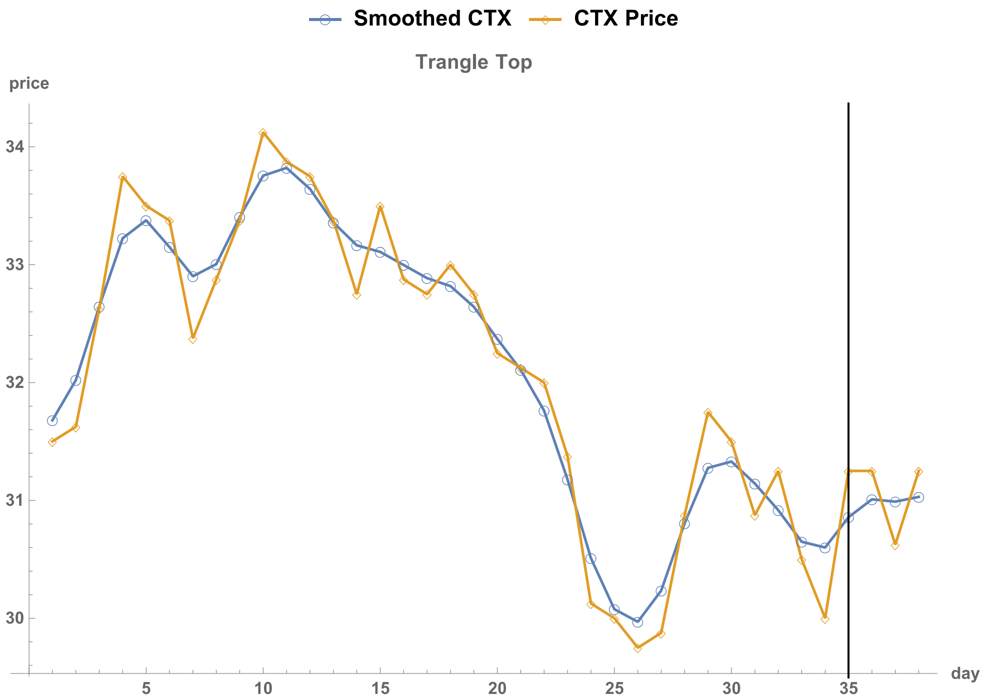


(a)

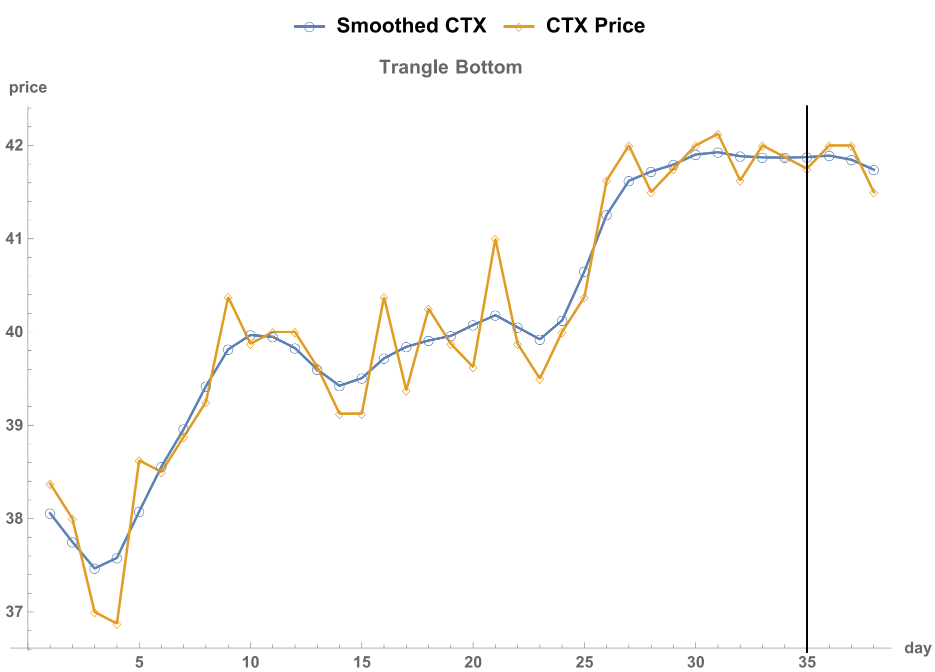


(b)

Figure 6. Broadening Top and Broadening Bottom

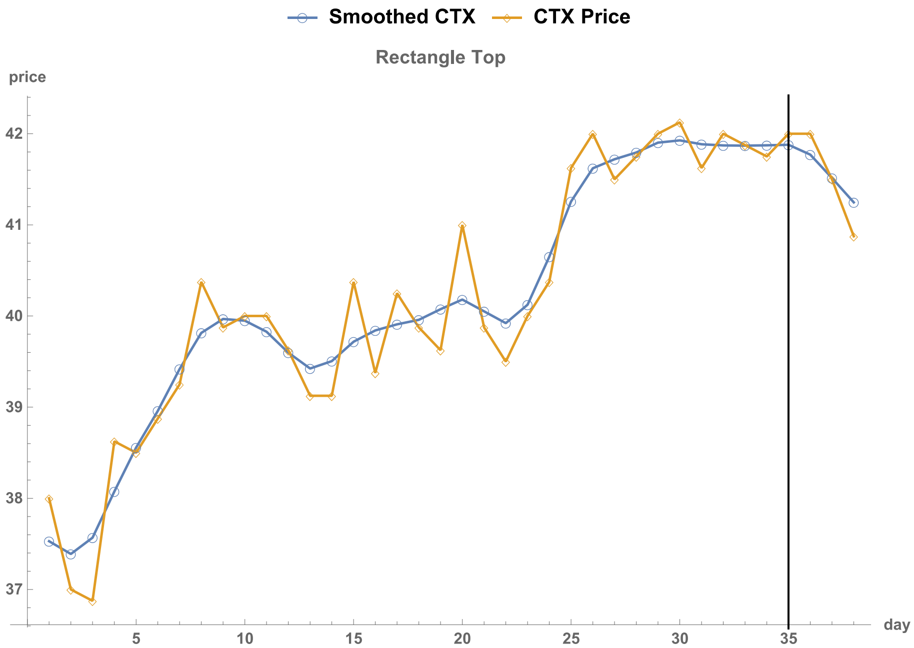


(a)

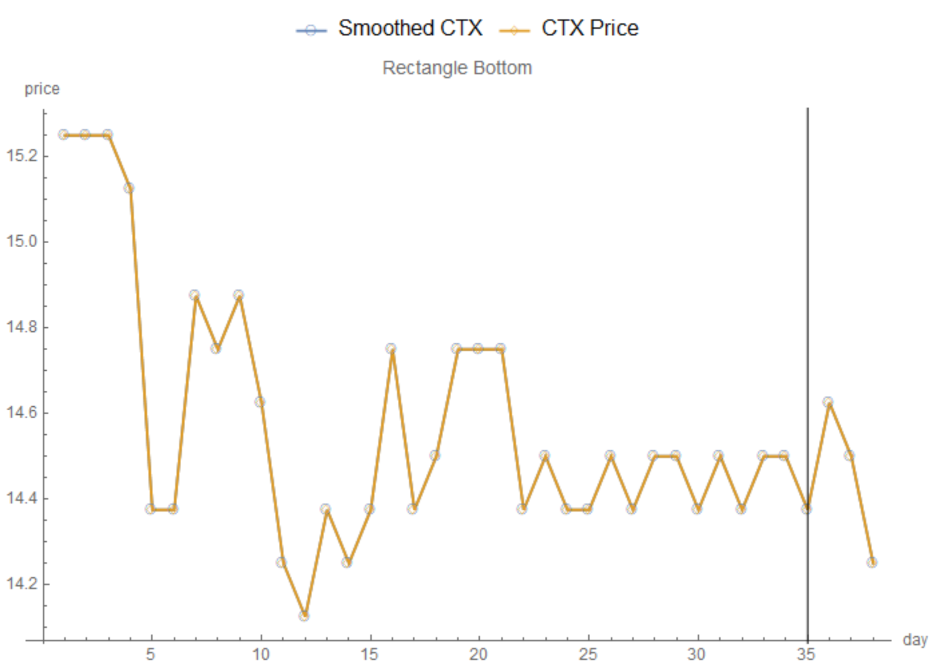


(b)

Figure 8.Triangle Top and Triangle Top

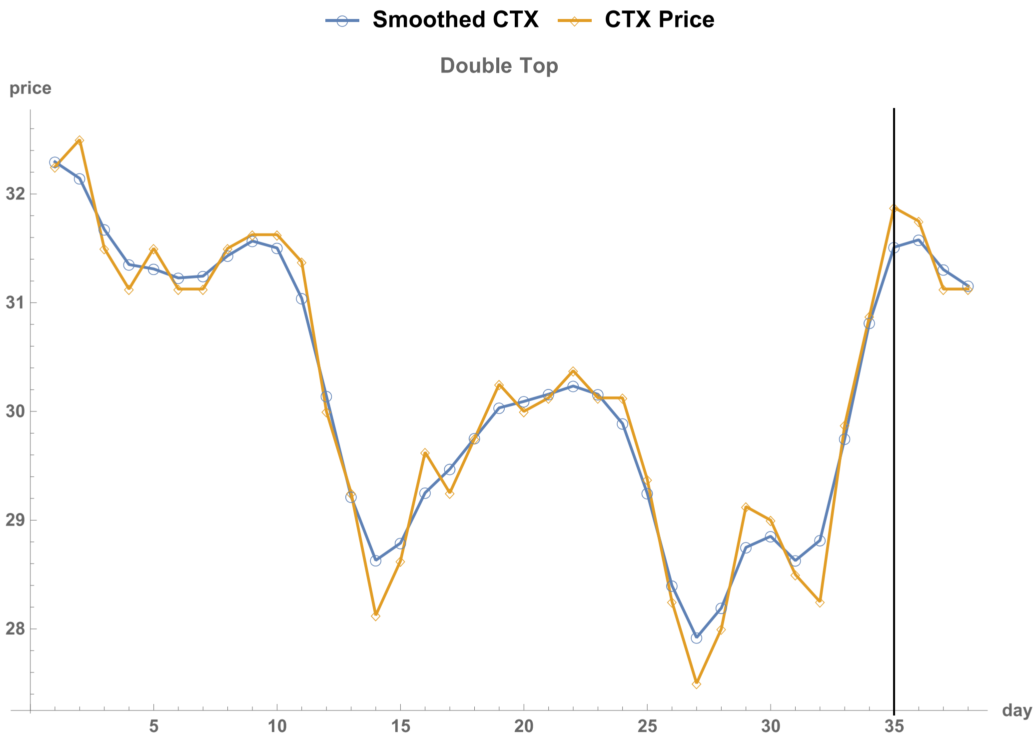


(a)

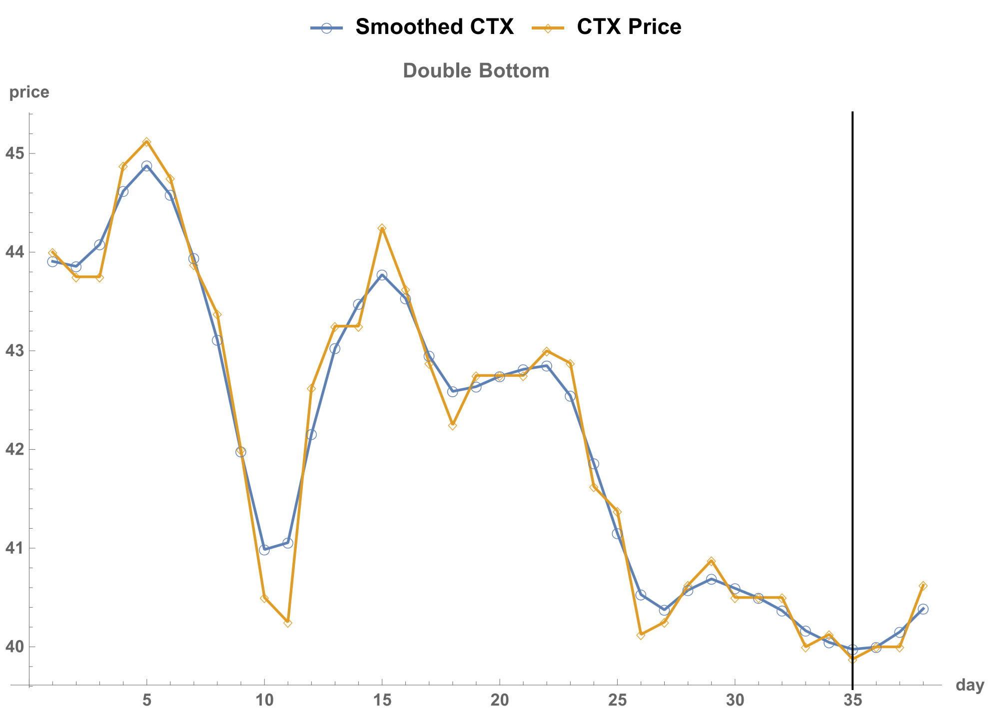


(b)

Figure 9.Rectangle Top and Rectangle Bottom



(a)



(b)

Figure 10.Double Top and Double Bottom

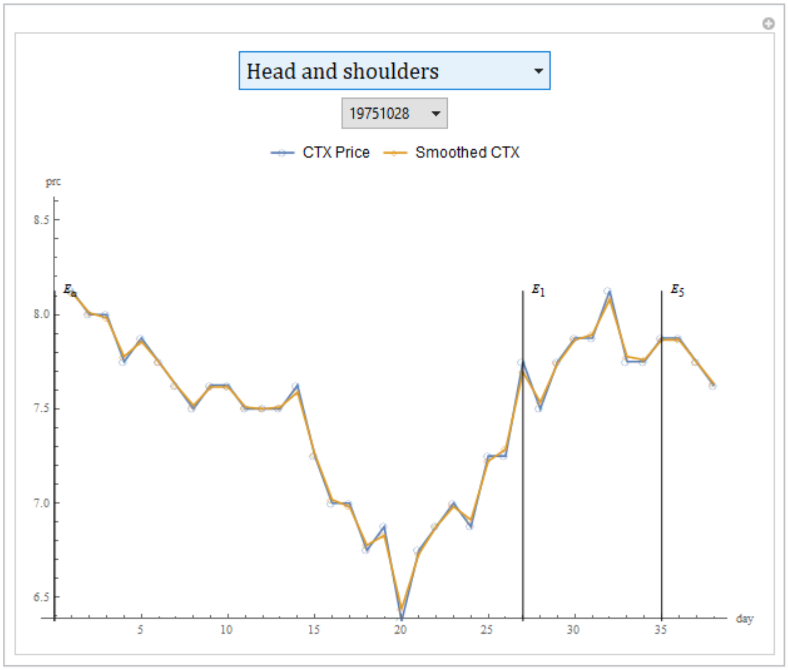
Knowing that our algorithm is able to match Lo, Mamaysky, and Wang (2000) and thus the visual inspection of their technical analysts in the year 1999 we know that our algorithm is on par for extending their work. For curiosity, we extend the sample to [T-l-d+1:T] for CTX.

For comparing patterns with a trained technical analyst, we note some of the professional societies below:

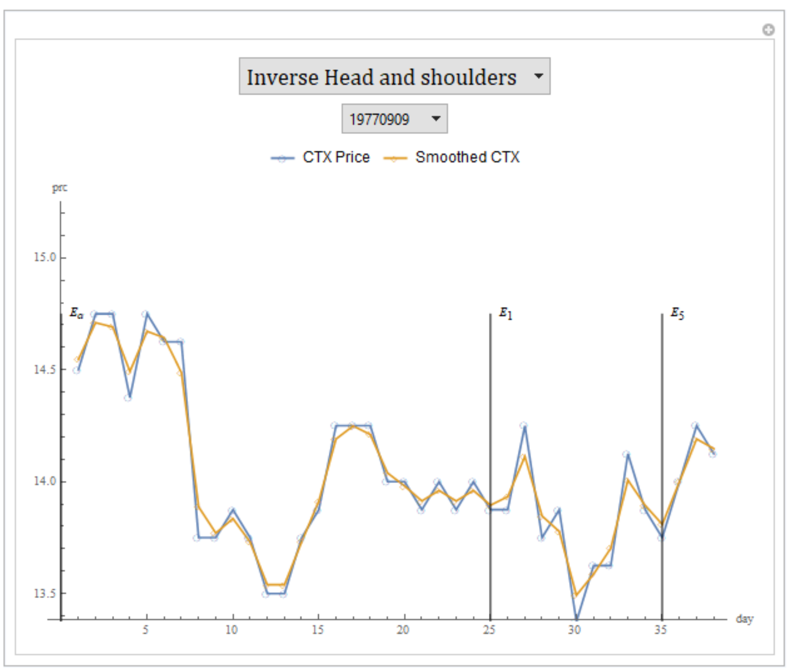
From 9500 daily price observations we have the following:	
Head and Shoulders (HS):	87
Inverse Head and Shoulders (HIS):	75
Broadening Top (BTOP):	33
Broadening Bottom (BBOT):	26
Triangle Top (TTOP):	42
Triangle Bottom (TBOT):	37
Rectangle Top (RTOP):	36
Rectangle Bottom (RBOT):	34
Double Top (DTOP):	41
Double Bottom (DBOT):	42

Table 4. Patterns Found

The patterns we find are shown below. Note the head-and-shoulders is the most common pattern found on the dataset. This is consistent with literature by Osler and Chang (1995), Linton (2010), Pring (2002), and Elder (2002). We have examples of all patterns above available. A few are shown below for proof of concept. Theses are out of sample as they were not used to fit the original bandwidth and local polynomial regression.

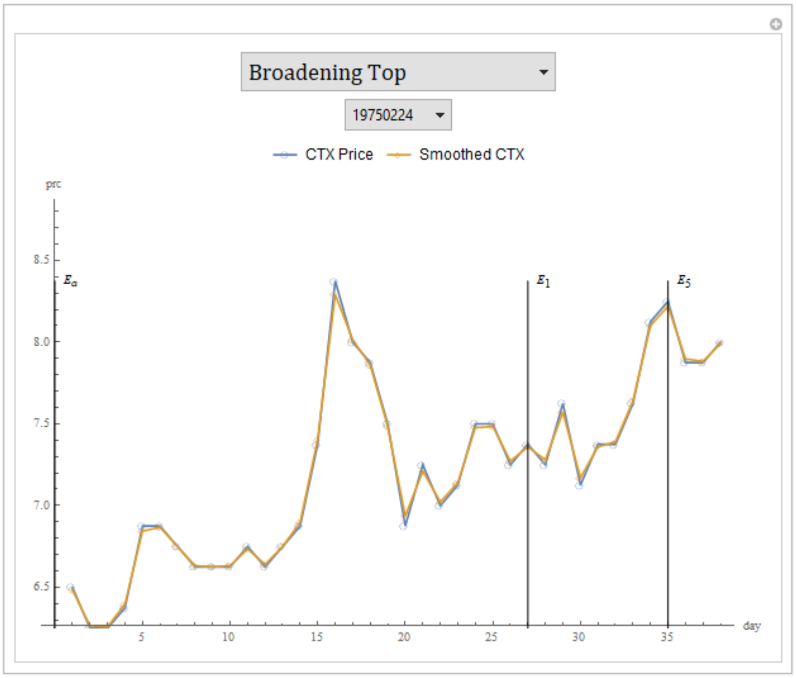


(a)

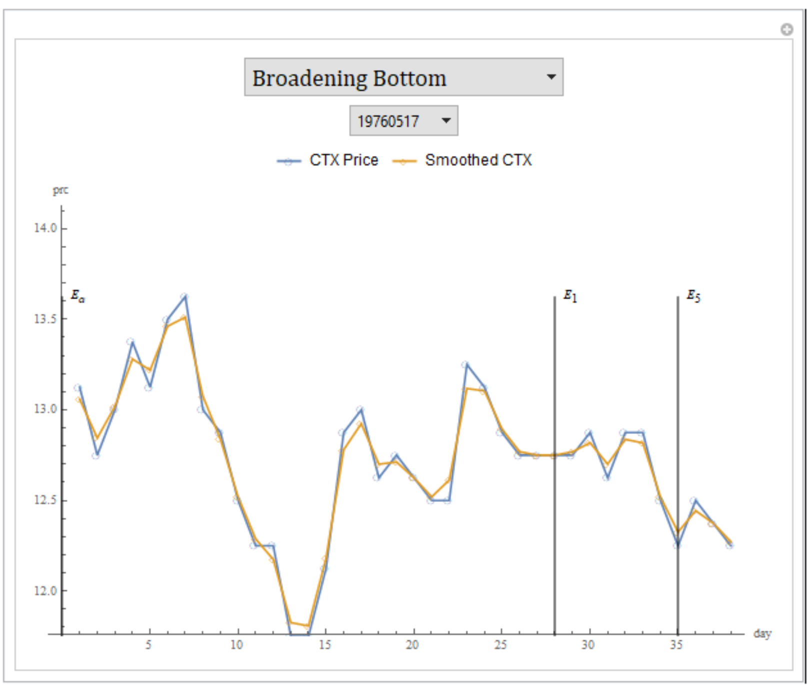


(b)

Figure 11. Head-and-Shoulders and Inverse Head-and-Shoulders Out of Sample

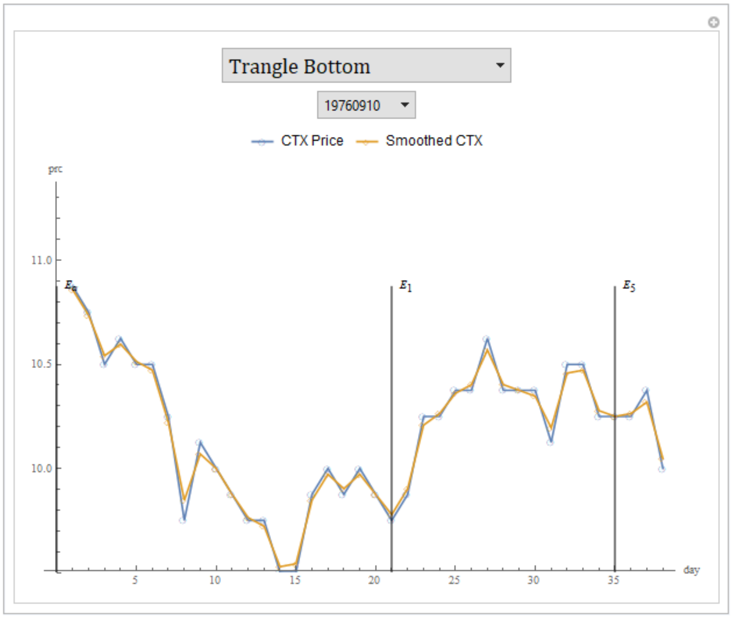


(a)

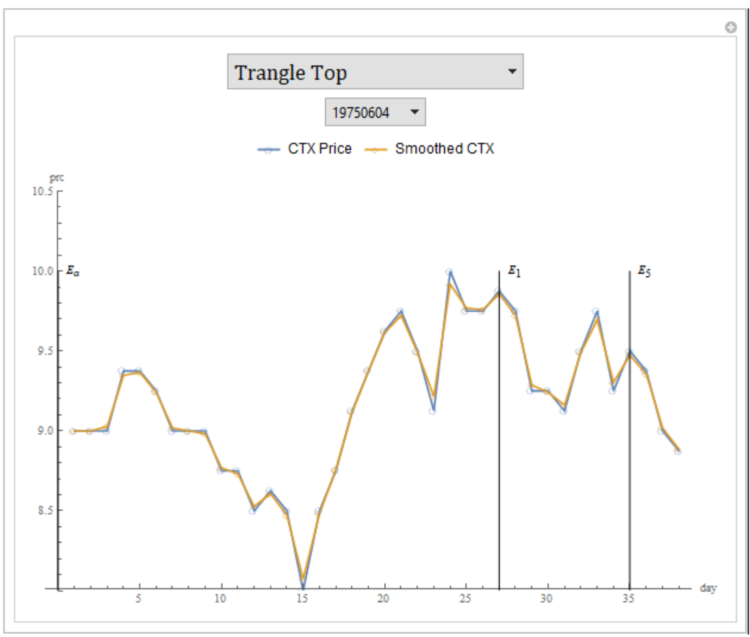


(b)

Figure 12. Broadening Top and Broadening Bottom Out of Sample

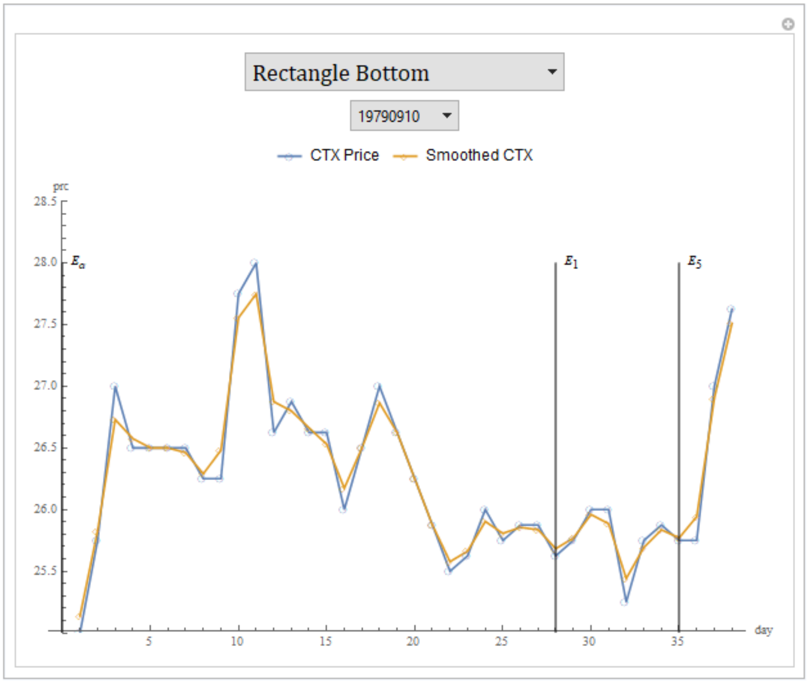


(a)

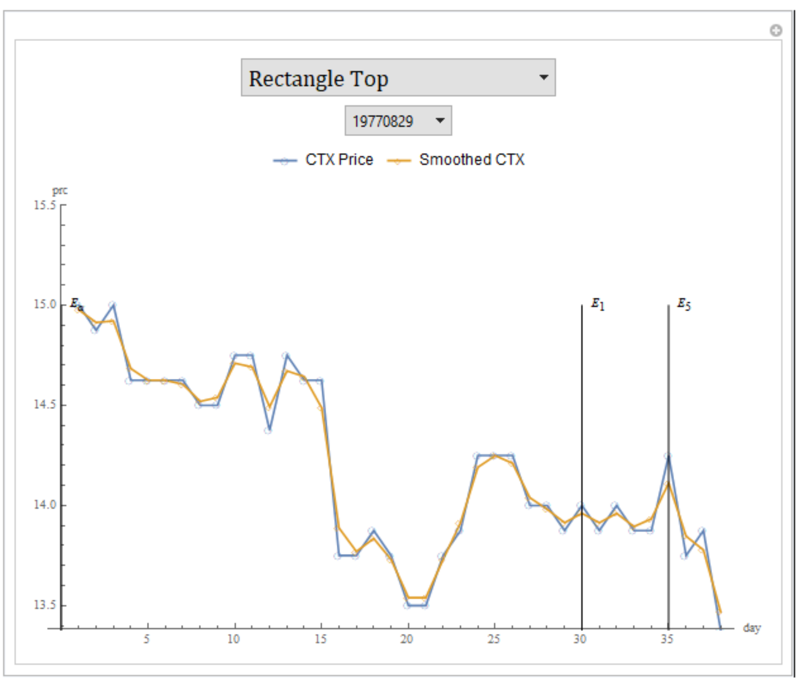


(b)

Figure 13: Triangle Top and Triangle Bottom Out of Sample

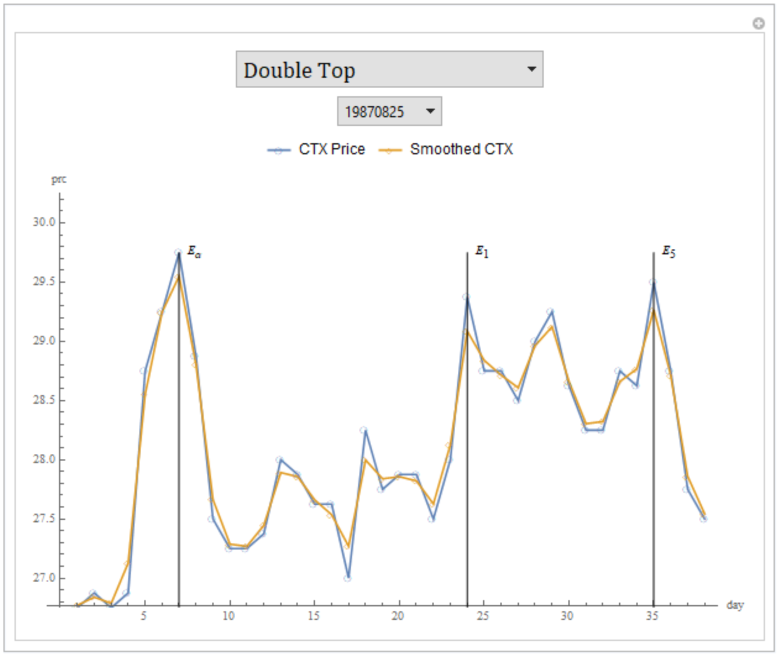


(a)

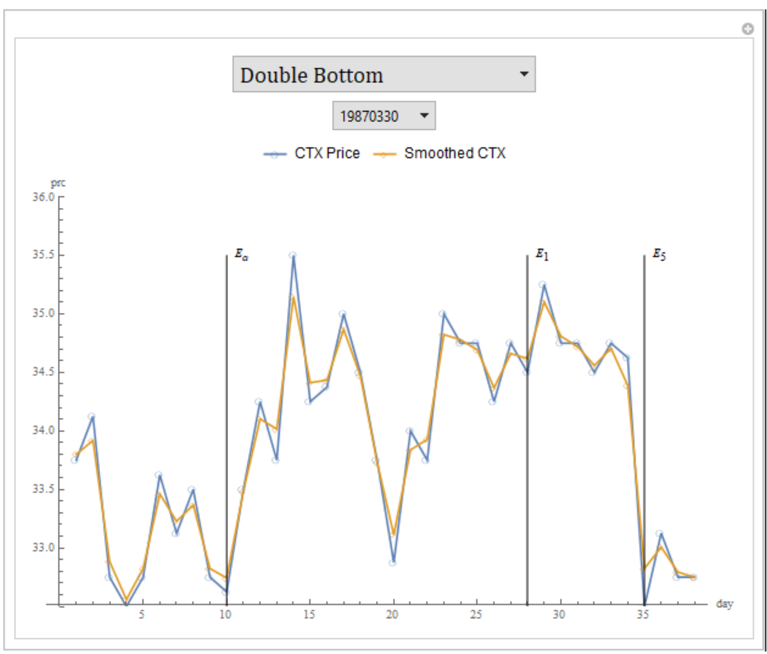


(b)

Figure 14. Rectangle Top and Rectangle Bottom Out of Sample



(a)



(b)

Figure 15. Double Top and Double Bottom Out of Sample

Chapter 2.

"Reliability of Technical Stock Price Patterns and Conditional Returns"

1. Introduction

In this section, we take the pattern rules from the above and develop a novel set up for extracting extrema from pixel values. These are not time dependent. We make a thin list plot of the prices and smoothed prices in each rolling window.

In answering part two of Lo, Mamaysky, and Wang (2000)'s call for an extension we come up with additional ways to obtain the patterns. We do this by finding extrema from digital images of the patterns, rather than on time series observations. We use this by making a very small plot of the pattern values. Remove redundant information and obtain extreme values. We then feed this information into a digital image classifier in part 3. We use the above methodology (methodology I) to obtain our pattern windows. What we do is make a very thin plot of our price data and smoothed data. We take the pixel values and delete all of the duplicate information. We then find the extrema from these values.

Analog image processing is the method technical analysts make decisions in stock charts. They process the information ocularly. They make fuzzy decisions based on how certain they are the image they are looking at contains a stock price pattern. This is not accounted for in prior technical analysis literature. This chapter aims at mimicking the way a trained technical analyst sees and processes information in stock charts. It may have a potential speed and accuracy advantage over the previous method discussed in chapter one.

In this section, we take our pattern data and plot images of them to find the extrema from pixel values. We test whether the returns are dependent on the size deciles using a goodness of fit. We also test a difference in means using a z-test.

We then take the pattern data and see how well a digital image classifier can uncover the patterns in the next section. We use held out data not included on the DJIA 30 and use it for training purposes this is done in part 3. We then test the ability of a digital image classifier to uncover the patterns on the DJIA 30. We show that with a high degree of accuracy a classifier can uncover both in sample and testing data.

2. Data

We use the 1925-2019 CRSP data for this section and the section after to extend our pattern sample.

We plot the survival of all CRSP components showing the average number of days a firm has survived. We find a high number of patterns the longer a firm has survived.

We use 98 years for our training data to encompass all market types. The figure below shows the number of days a firm has survived sorted from largest to smallest.

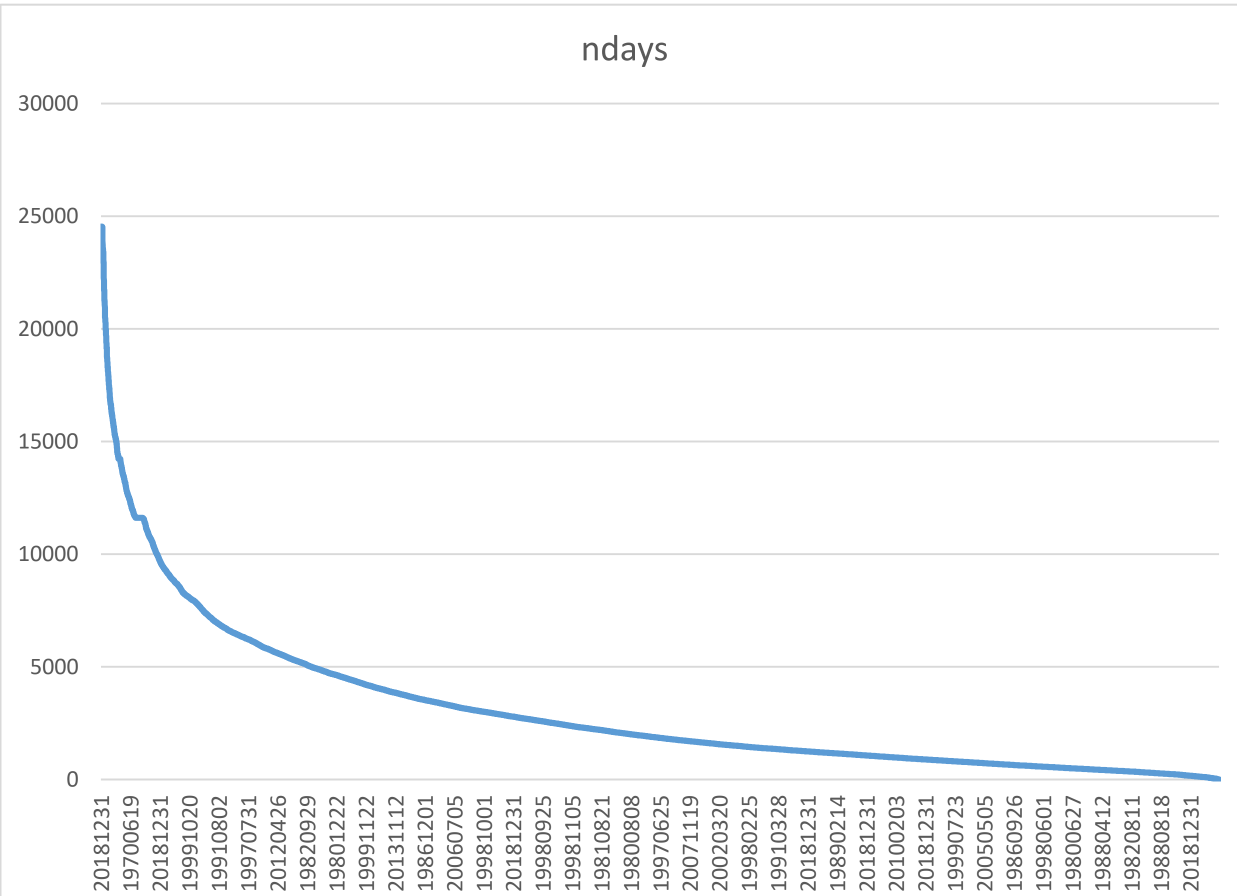


Figure 16 Number of Days

The Y axis is days survived and the X-axis is the end date. This is for all firms in the database, which is around 24,000. The average survival is 7.38 years. We narrow our database to the current Dow Jones Industrial Average since they haven’t been tested by literature and it gives us a nice sample for finding patterns. The average years here is 56.32.

permno	hcomnam	htsymbol	startdate	enddate	ndays	nyears
10107	MICROSOFT CORP	MSFT	19860313	20181231	8269	33.076
10284	AMERICAN EXPRESS CO	AXP	19251231	19390731	4050	16.2
11308	COCA COLA CO	KO	19251231	20181231	24541	98.164
11850	EXXON MOBIL CORP	XOM	19251231	20181231	24541	98.164
12490	INTERNATIONAL BUSINESS MACHS COR	IBM	19251231	20181231	24541	98.164
14541	CHEVRON CORP NEW	CVX	19251231	20181231	24541	98.164
14593	APPLE INC	AAPL	19801212	20181231	9595	38.38
16851	DOWDUPONT INC	DWDP	20170901	20181231	334	1.336
17830	UNITED TECHNOLOGIES CORP	UTX	19290411	20181231	23566	94.264
18163	PROCTER & GAMBLE CO	PG	19290812	20181231	23463	93.852
18542	CATERPILLAR INC	CAT	19291202	20181231	23379	93.516
19502	WALGREENS BOOTS ALLIANCE INC	WBA	19340215	20181231	22132	88.528
19561	BOEING CO	BA	19340905	20181231	21964	87.856
21936	PFIZER INC	PFE	19440117	20181231	19149	76.596
22111	JOHNSON & JOHNSON	JNJ	19440925	20181231	18942	75.768
22592	3M CO	MMM	19460114	20181231	18567	74.268
22752	MERCK & CO INC NEW	MRK	19460515	20181231	18467	73.868
43449	MCDONALDS CORP	MCD	19660705	20181231	13213	52.852
47896	JPMORGAN CHASE & CO	JPM	19690305	20181231	12569	50.276
55976	WALMART INC	WMT	19721120	20181231	11630	46.52
57665	NIKE INC	NKE	19801202	20181231	9603	38.412
59328	INTEL CORP	INTC	19721214	20181231	11613	46.452
59459	TRAVELERS COMPANIES INC	TRV	19721214	20181231	11613	46.452

Table 5 Dow Jones Industrial Average Components.

65875	VERIZON COMMUNICATIONS INC	VZ	19840216	20181231	8791	35.164
76076	CISCO SYSTEMS INC	CSCO	19900216	20181231	7274	29.096
86868	GOLDMAN SACHS GROUP INC	GS	19990504	20181231	4948	19.792
87436	DISNEY WALT CO	DIS	19991118	20010319	335	1.34
92611	VISA INC	V	20080319	20181231	2716	10.864
92655	UNITEDHEALTH GROUP INC	UNH	19841017	20181231	8622	34.488

Table 5. Dow Jones Industrial Average (DJIA) Components.

The unconditional returns for the DJIA are shown below. This matches up with the summary statistics shown in the next section.

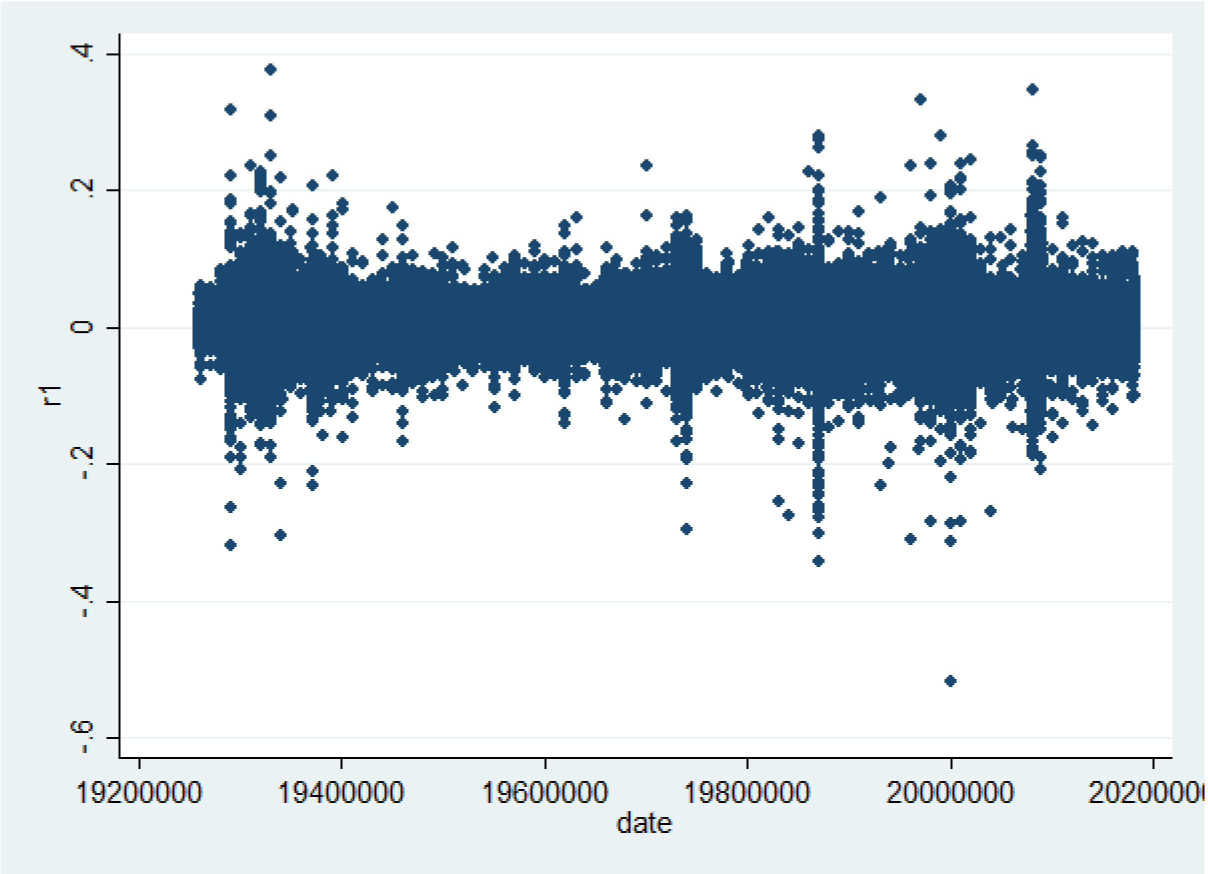


Figure 17. Unconditional Returns

3. **Methodology**

We focus on the patterns that are a series of five consecutive extrema. Thus leaving out DBOT and DTOP.

In this section, we focus on finding extrema from pixel values which adds to prior work (i.e. finding new ways to uncover the patterns). We show that the patterns are found on both methods (kernel smoothing with local polynomial regression) and image data the same.

The main difference with the image data is that the extrema consist of just pixel values, rather than time series. This is of interest in showing the viability of a digital image classifier to be useful in the field of finance.

We test our patterns for returns using the methodology from LMW, (2000). Mainly we look at what percent of the returns conditioned on each pattern fit into the 10 deciles of the unconditional returns.

After that, we train a digital image classifier to recognize the patterns from extrema. Our goal is to have both accuracies both in sample and out of sample.

We use the same methodology outlined in Chapter 1, but rather than finding our extrema from a time-series we find our extrema from pixel values. We then extend the sample to include the Dow Jones components from CRSP as outlined above.

We discuss the implications of the return in the sections below, before concluding. The main results show that the pattern classifier can add value without taking away from the returns.

We use the extreme values and match the pattern name found by the methodology in this section. It involves smoothing the data with local polynomial regression. Making a very thin plot of every point for both the kernel and the price. Deleting duplicate information. Finding extrema. We then use the 5 extrema in the rules from part one. The result is the same, we find the same patterns as what is in LMW, (2000). We showed the results of the patterns on the DJIA in the section above. In this section, we use digital image classification to identify the patterns.

For enacting the returns we have two implications. 1) for bottom patterns we earn the market return. 2) for top patterns we earn -1 * the market return.

We have short entry on the market when we have a top pattern and long entry on a bottom. Short entry requires borrowing a security and then covering it later, thus -1*the market return. We are betting on the prices to go down.

4. Results

Here we show summary statistics for the patterns found using image processing on the DJIA stocks. We show both the frequency of patterns found and the pattern statistics including all moments.

Frequency counts for 8 technical indicators detected among the DJIA stocks from 1925-2019. Each stock's price history is scanned for the occurrence of the following 8 technical indicators: head-and-shoulders (HS), inverted head-and-shoulders (HIS), broadening top (BTOP), broadening bottom (BBOT), triangle top (TTOP), triangle bottom (TBOT), rectangle top (RTOP), rectangle bottom (RBOT). The 'Raw' column shows the number of total daily prices observations over the sample.

Sample	Raw	BBOT	BTOP	HS	HIS	RBOT	RTOP	TBOT	TTOP	Conditional
DJIA	428,835	228	583	1488	651	262	599	286	523	4620
	100%	4.94%	12.62%	32.21%	14.09%	5.67%	12.97%	6.19%	11.32%	100%

Table 6. Frequency of Patterns

This table shows the pattern distribution across the dataset for the DJIA. 32% of the patterns are head and shoulders. 14% of the patterns are reverse head and shoulders. 13% of the patterns are rectangle tops, 13% of the patterns are broadening tops. 11% of the patterns are triangle tops. With the exception of the reverse head and shoulders there seem to be a bias towards top reversal patterns compared to bottoms. 6% of the patterns are triangle bottoms. 6% of the patterns are rectangle bottoms. 5% of the patterns are broadening bottoms., the majority of these patterns are tops. Slightly more than 1% of the overall data result in patterns.

The conditional returns for all 8 patterns are shown in the figure below. This is then analyzed in the summary statistics table with more moments.

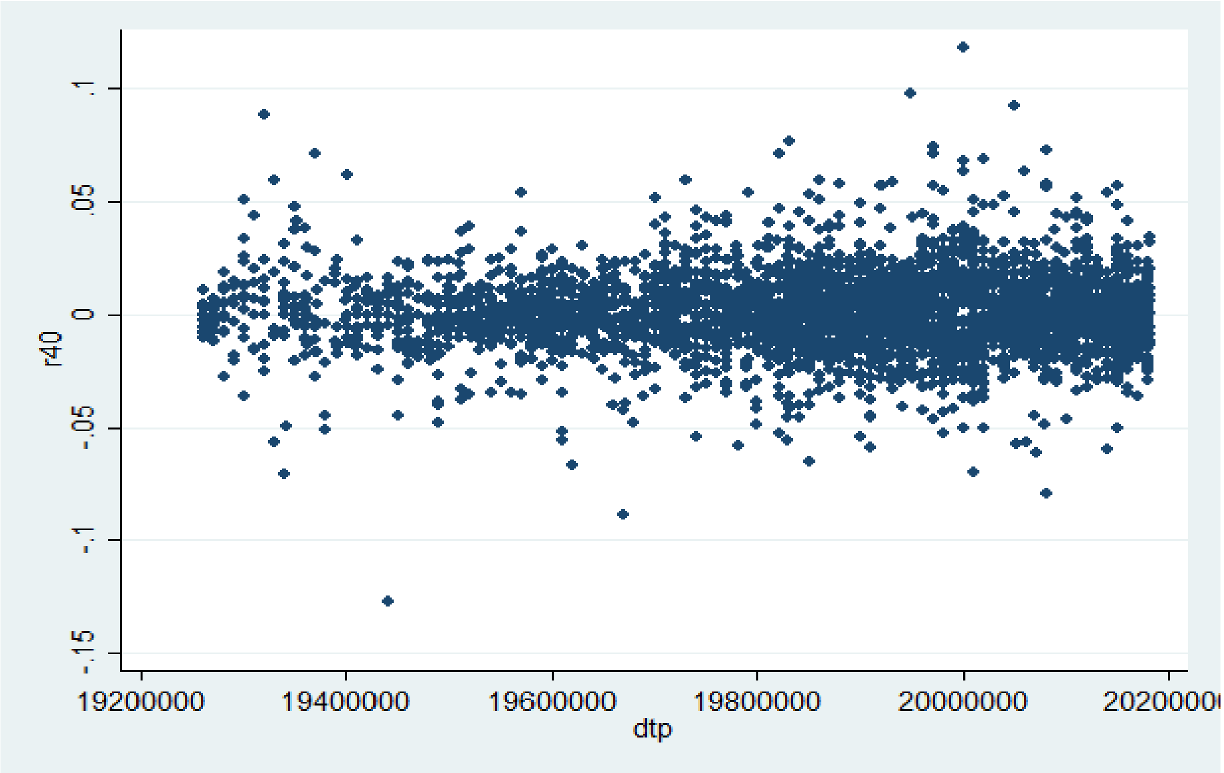


Figure 18. Conditional Returns

This figure shows that patterns appear to be more concentrated in recent years and that there are fewer patterns early in the history of the current DJIA sample.

The return statistics are analyzed for both conditional returns, unconditional returns and by a pattern in the table below. We analyze all moments of returns.

Summary statistics for 8 technical indicators detected among the DJIA stocks from 1925-2019. Each stock's price history is scanned for the occurrence of the following 8 technical indicators: head-and-shoulders (HS), inverted head-and-shoulders (HIS), broadening top (BTOP), broadening bottom (BBOT), triangle top (TTOP), triangle bottom (TBOT), rectangle top (RTOP), rectangle bottom (RBOT). The 'Raw' column shows the summary statistics for the total daily prices observations over the sample.									
Moment	Raw	HS	HIS	RTOP	RBOT	BBOT	BTOP	TBOT	TTOP
Mean	.0006816	-.0013	-.0007	-.0009	.0009	.0020	.0009	.0012	-.0011
Standard Deviation	.0187	.0155	.0152	0.1392	.0124	.0176	.0172	.0142	.0180
Median	.0000	-.0004	.0000	-.0005	.0000	.0011	.0001	.0011	.0000
Skewness	.1556	-.6076	-.2247	-1.0880	.5647	.1916	1.0581	-.0960	-.6830
Kurtosis	18.0992	6.4783	5.1681	12.5760	6.4959	6.7729	10.4434	4.4076	6.4230

Table 7. Summary statistics for patterns.

This shows that the raw data has the highest and lowest overall returns with the highest variance. The average return is lower than 6/8 patterns. The overall conditional return is 100 basis points higher than the raw return. The conditional returns have lower standard deviation and variance.

We normalize the returns following:

$$X_{it} = \frac{R_{it} - Mean[R_{it}]}{SD[R_{it}]} \tag{14}$$

Recall that patterns are completed at t+l-1. We take the conditional return such that $R^P = log(1 + R_{t+l+d+1})$. For each stock, we have 8 sets of conditional returns, each conditioned on one of the 8/10 patterns discussed previously.

There are 4,620 patterns int he Dow from 1925-2019 (98 years). All patterns are tested on the current 30 Dow Jones (DJIA) components. 228 Broadening Bottoms, 583 Broadening Tops, 1488 Head and Shoulders, 651 Inverse Head and Shoulders, 262 Rectangle Bottoms, 599 Rectangle Tops, 286 Triangle Bottoms, 523 Triangle Tops. This makes 4,620 patterns. There are 429,620 daily price observations across all 30 stocks.

Patterns returns are tabulated across all size deciles. The table below shows the percent of each pattern returns that fit into the range of returns for each decile. From LMW, 2000 if conditioning on the return adds informational value (incremental value) there should be more than a 10% fit.

The goodness of fit follows LMW (2000) where the relative frequency $\delta^{\wedge}{}_{j}$ of conditional returns falling into decile j of the unconditional returns, j = 1,…,10:

$$\delta^{\wedge}{}_{j} = \frac{number of conditional returns in decile j}{total number of conditional returns} \tag{15}$$

Under the null hypothesis that the returns are independently and identically distributed (IID) and the conditional and unconditional return distributions are identical, the asymptotic return distributions of $\delta^{\wedge}{}_{j}$ are given by:

$$\sqrt{n}(\delta^{\wedge}{}_{j} - 0.1) \sim^{a} N(0{,}0.1(1-0.1)). \tag{16}$$

For analyzing the goodness of fit, and the dependence of the patterns to fit in the 10 size deciles we use the following measure for our test statistic:

$$Q \equiv \sum_{j=1}^{10} \frac{(n_j - 0.1n)^2}{0.1n} \sim^{a} \chi_9^2 \tag{17}$$

Testing against the null hypothesis, the asymptotic z-test is in the table below for each value. The standard errors and number of observations are recorded from the population.

We consider a goodness of fit test diagnostic in the table below. Our results suggest that the patterns found depend on the decile.

This has statistical implications but not economic implications. Economically they shouldn’t since all of the DJIA components are large cap stocks.

We look at the deciles of the DJIA and where the patterns are found proceeding the table. This suggests that the deciles are time varying.

Goodness-of-fit statistics for conditional one-day normalized returns, conditional on 8 technical indicators, for a sample of 30 DJIA stocks from 1925-2019. For each pattern, the percentage of conditional returns that falls within each of the 10 unconditional return deciles is tabulated. If conditioning on the pattern provides no information, the expected percentage falling in each decile is 10%. Asymptotic z-statistics for this null hypothesis are reported in parentheses below the statistic. The 8 technical indicators are as follows: head-and-shoulders (HS), inverted head-and-shoulders (IHS), broadening top (BTOP), broadening bottom (BBOT), triangle top (TTOP), triangle bottom (TBOT), rectangle top (TTOP), rectangle bottom (RBOT), rectangle top (RTOP).

Pattern	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	Q (p-value)
BBOT	6.58 (6.83)	10.96 1.89	7.02 (5.79)	13.16 (6.51)	8.33 (3.61)	9.65 (0.74)	8.77 (2.65)	11.40 2.60	14.04 8.51	10.09 0.23	12.35 (0.19)
BTOP	7.20 (5.58)	8.06 (3.80)	12.01 3.90	9.26 (1.52)	8.23 (3.83)	9.78 (0.47)	10.98 2.11	9.95 (0.10)	12.52 5.32	12.01 5.15	588.60 (0.00)
HS	8.33 (3.33)	10.55 1.08	9.68 (0.63)	9.21 (1.63)	11.22 2.65	9.48 (1.11)	11.36 2.93	9.07 (1.72)	9.41 (1.25)	11.69 4.35	6694.54 (0.00)
IHS	8.91 (2.18)	9.98 (0.03)	10.45 0.87	10.45 0.92	10.91 1.96	8.60 (2.95)	10.91 1.96	9.52 (0.88)	9.83 (0.36)	10.45 1.14	759.94 (0.00)
RBOT	11.83 3.66	11.45 2.84	8.02 (3.85)	9.92 (0.16)	13.36 7.28	9.92 (0.16)	9.92 (0.16)	8.78 (2.26)	5.73 (9.02)	11.07 2.74	17.21 (0.05)
RTOP	9.68 (0.63)	13.02 5.92	10.85 1.65	10.18 0.38	8.18 (3.94)	9.52 (1.02)	10.85 1.84	8.85 (2.13)	9.02 (2.08)	9.85 (0.39)	596.12 (0.00)
TBOT	7.34 (5.31)	8.04 (3.84)	7.34 (5.16)	15.03 10.38	10.49 1.06	8.04 (4.13)	13.99 8.60	8.39 (2.98)	9.44 (1.18)	11.89 4.84	39.74 (0.00)
TTOP	10.33 0.65	8.80 (2.36)	11.85 3.60	9.94 (0.12)	9.56 (0.95)	8.60 (2.94)	9.18 (1.77)	9.75 (0.46)	11.47 3.11	10.52 1.32	369.98 (0.00)

Table 8. The goodness of fit diagnostics.

We also test a Kolmogorov-Smirnov p-value for the difference in the conditional and unconditional distributions. The formulae behind this are shown below.

The Kolmogorov-Smirnov test is useful for hypotheses that involve lumpness or clustering in data. This is what is used in Lo, Mamaysky and Wang (2000).

The directional hypotheses are evaluated with the statistics

$$D^{+} = \max_{x}\{F(x) - G(x)\} \tag{18}$$

$$D^{-} = \min_{x}\{F(x) - G(x)\} \tag{19}$$

Where F(x) and G(x) are the empirical distribution function for the sample being compared. The combined statistic is

$$D = \max(|D^{+}|, |D^{-}|) \tag{20}$$

The p-value for this may be obtained by evaluating the asymptotic limiting distribution. Let m be the sample size for the first sample, and let n be the sample size for the second sample.

Smirnov (1933) shows that

$$\lim_{m,n\to\infty} \Pr\{\sqrt{\mathrm{mn}}/(m+n)\, D_{m,n} \leq z\} = 1 - 2\sum_{i=1}^{\infty}(-1)^{i-1}\mathrm{e}^{-2i^{2}z^{2}} \tag{21}$$

The first five terms form the approximation P_a. The exact p-value is calculated by a counting algorithm. A corrected p-value was obtained by modifying the asymptotic p-value by using a numerical approximation technique:

$$Z = \phi^{-1}(P_a) + 1.04/\min(m,n) + 2.09/\max(m,n) - 1.35/\sqrt{\mathrm{mn}/(m+n)} \tag{22}$$

$$p - \text{value} = \phi(Z) \tag{23}$$

where $\phi(.)$ is the cumulative normal distribution.

Our results are as follows.

Kolomgrov-Smirnov test of the equality of conditional and unconditional one-day return distributions for DJIA stocks from 1925-2019. Conditional returns are the daily return three days following the conclusion of an occurrence of one of the 8 technical indicators: head-and-shoulders (HS), inverted head-and-shoulders (HIS), broadening top (BTOP), broadening bottom (BBOT), triangle top (TTOP), triangle bottom (TBOT), rectangle top (RTOP), rectangle bottom (RBOT). All returns are normalized by subtracting their means and dividing by their standard deviations. P-values are with respect to the asymptotic distribution of the Kolomgorov-Smirnov test statistic.

Kolmogorov-Smirnov	HS	HIS	BTOP	BBOT	TTOP	TBOT	RTOP	RBOT	All
Combined γ	1.0327	1.0308	1.0300	1.0097	1.0256	1.0168	1.0168	1.0171	1.0398
p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Table 9 Kolmogorov-Smirnov

This makes it kind of interesting to note the patterns and where they are found when broken into size deciles. We note the time-varying effect of the patterns below.

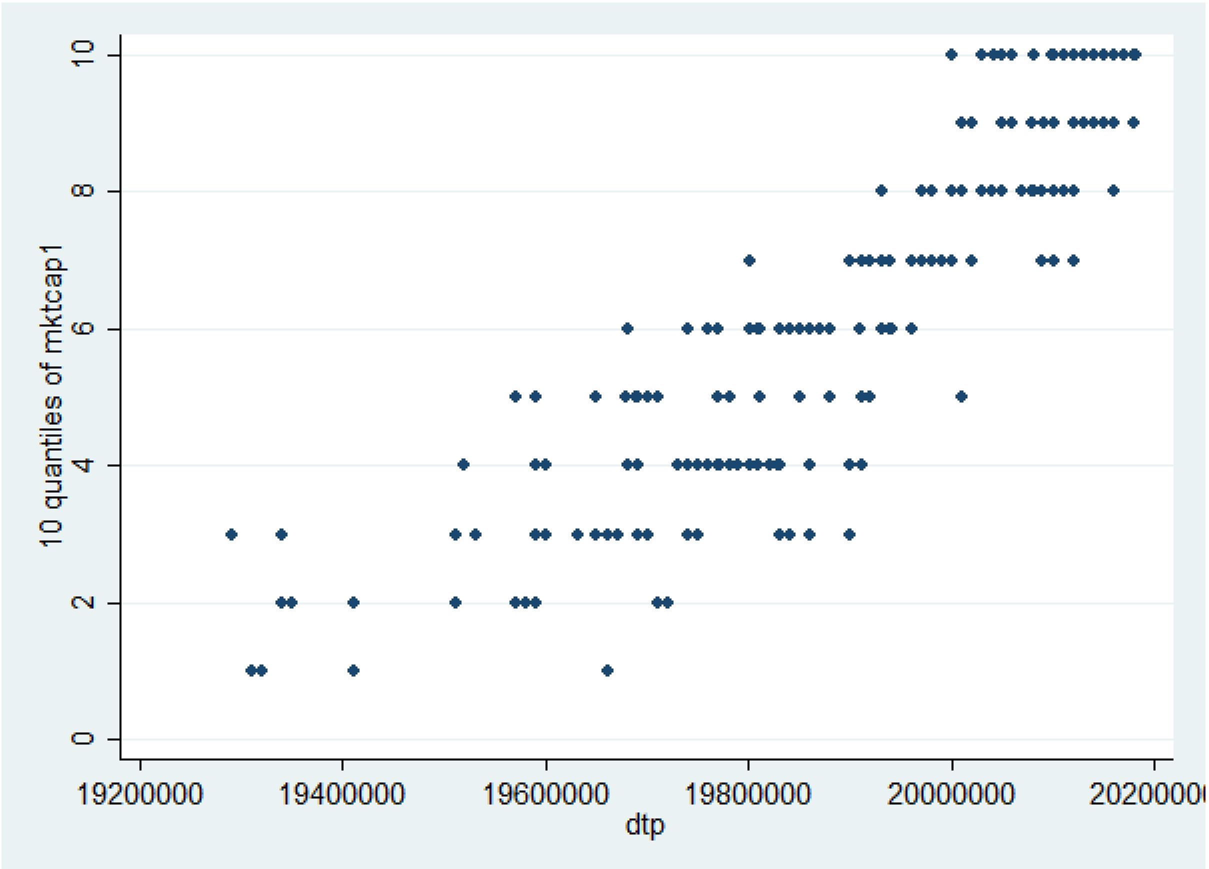


Figure 19. BBOT time-varying returns

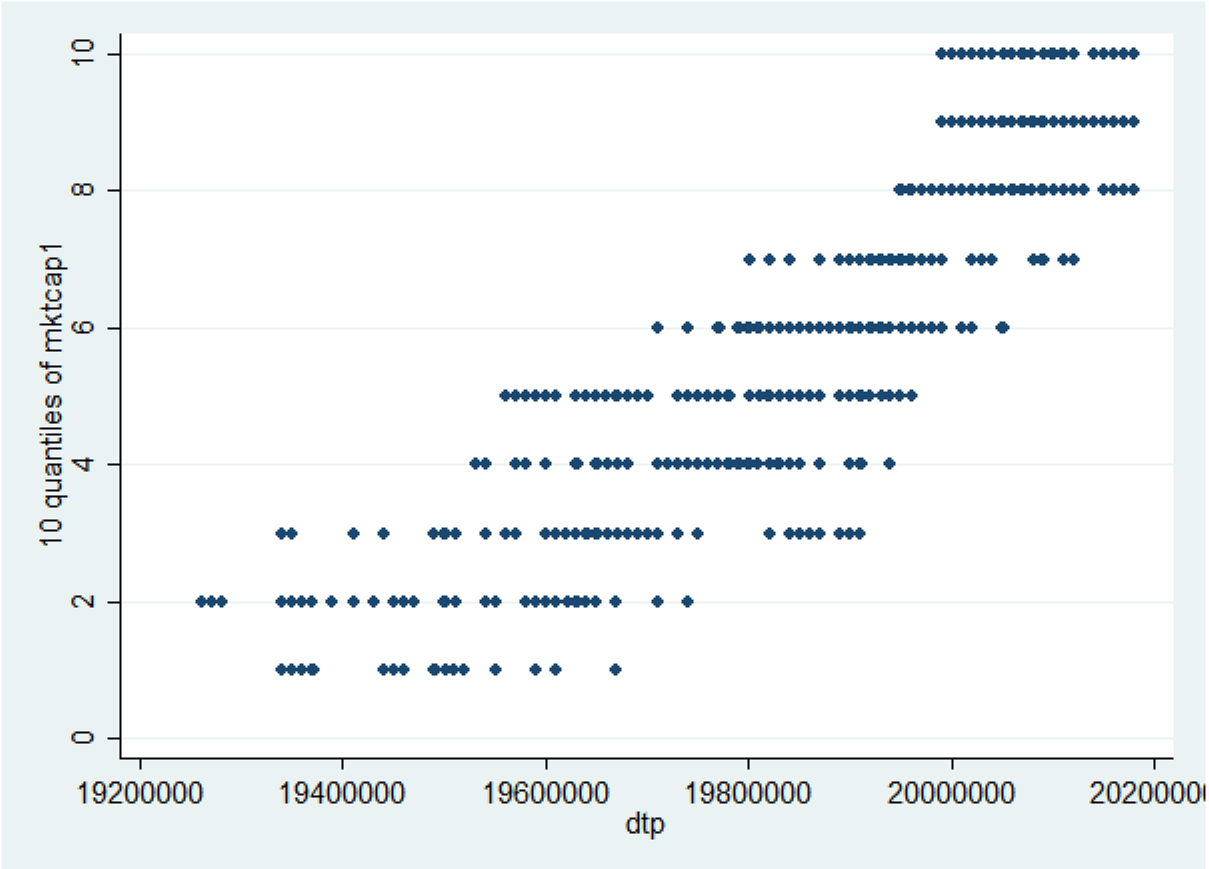


Figure 20. BTOP Time-Varying Observations.

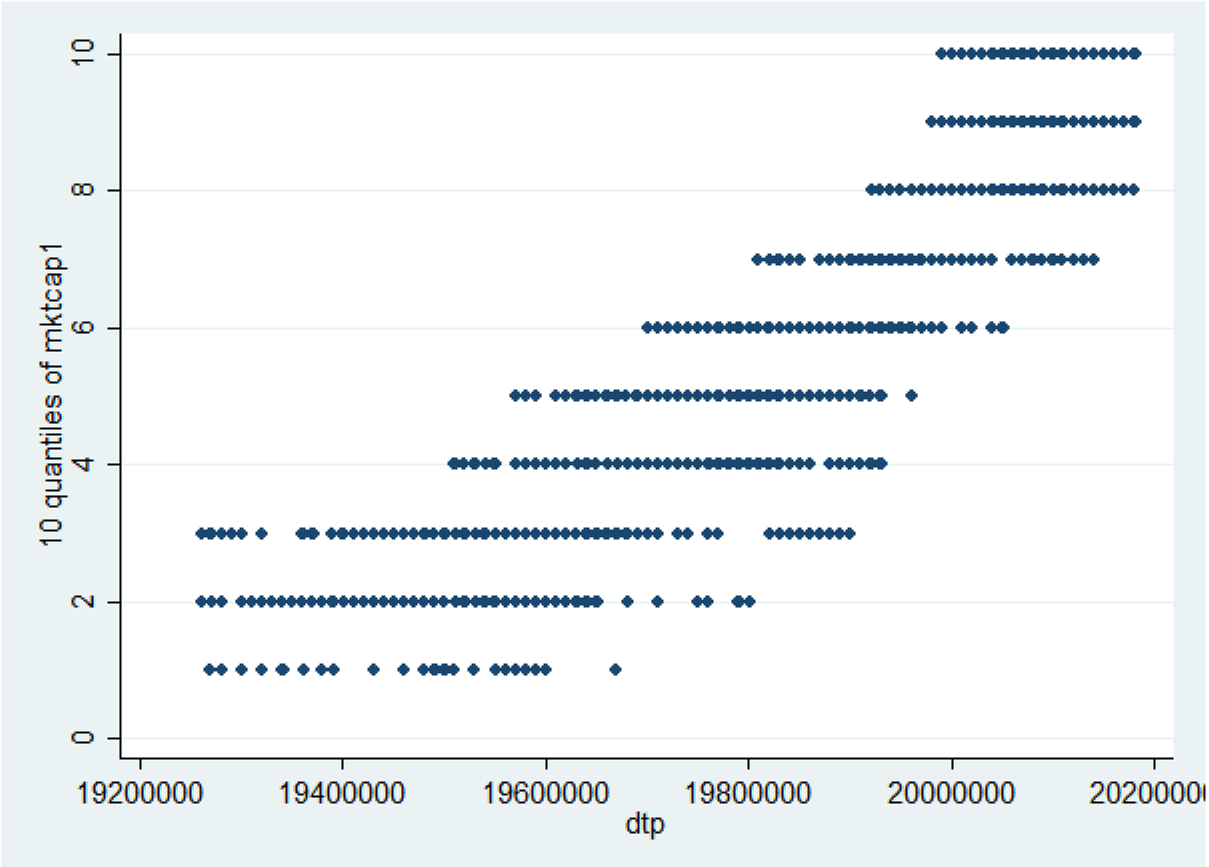


Figure 21. Head and Shoulders Time-Varying Observations

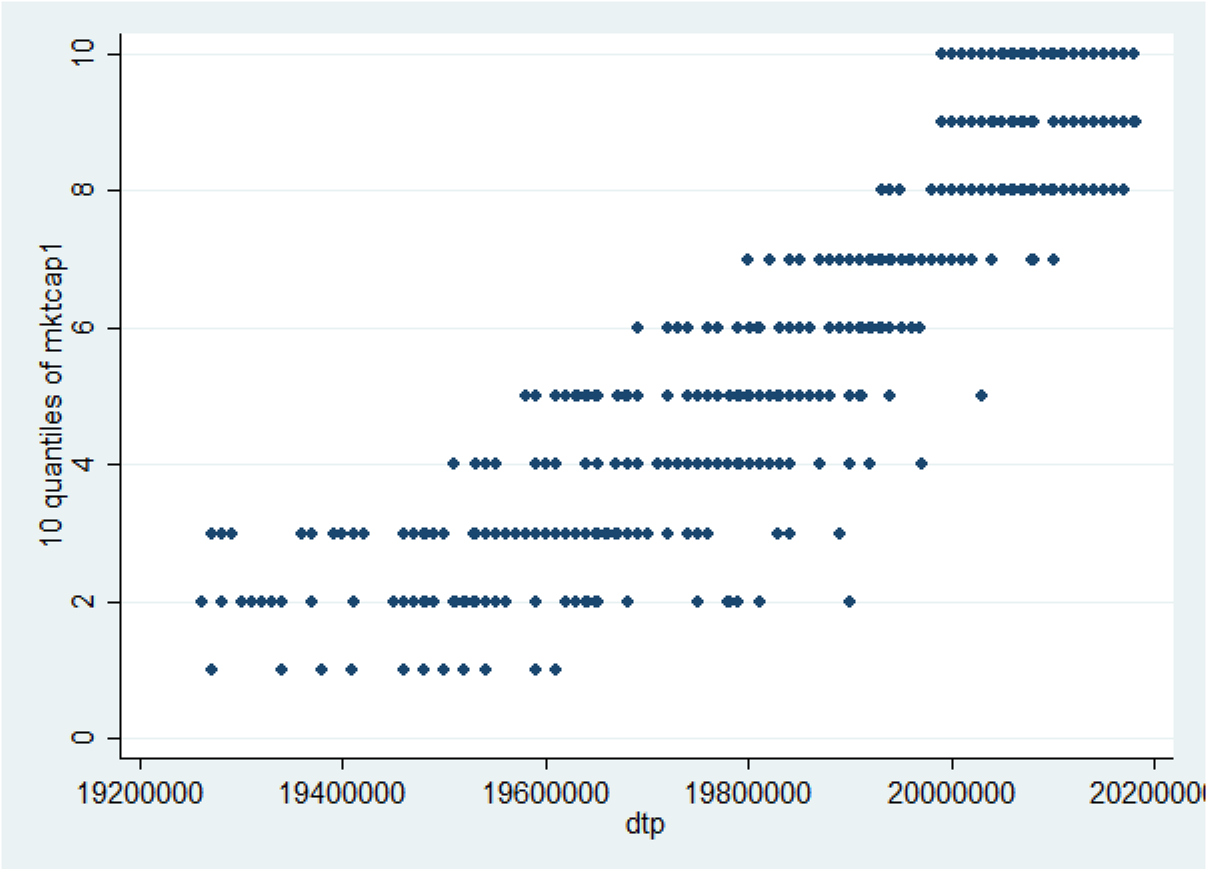


Figure 22. Inverted Head-and-Shoulders Time-Varying Observations.

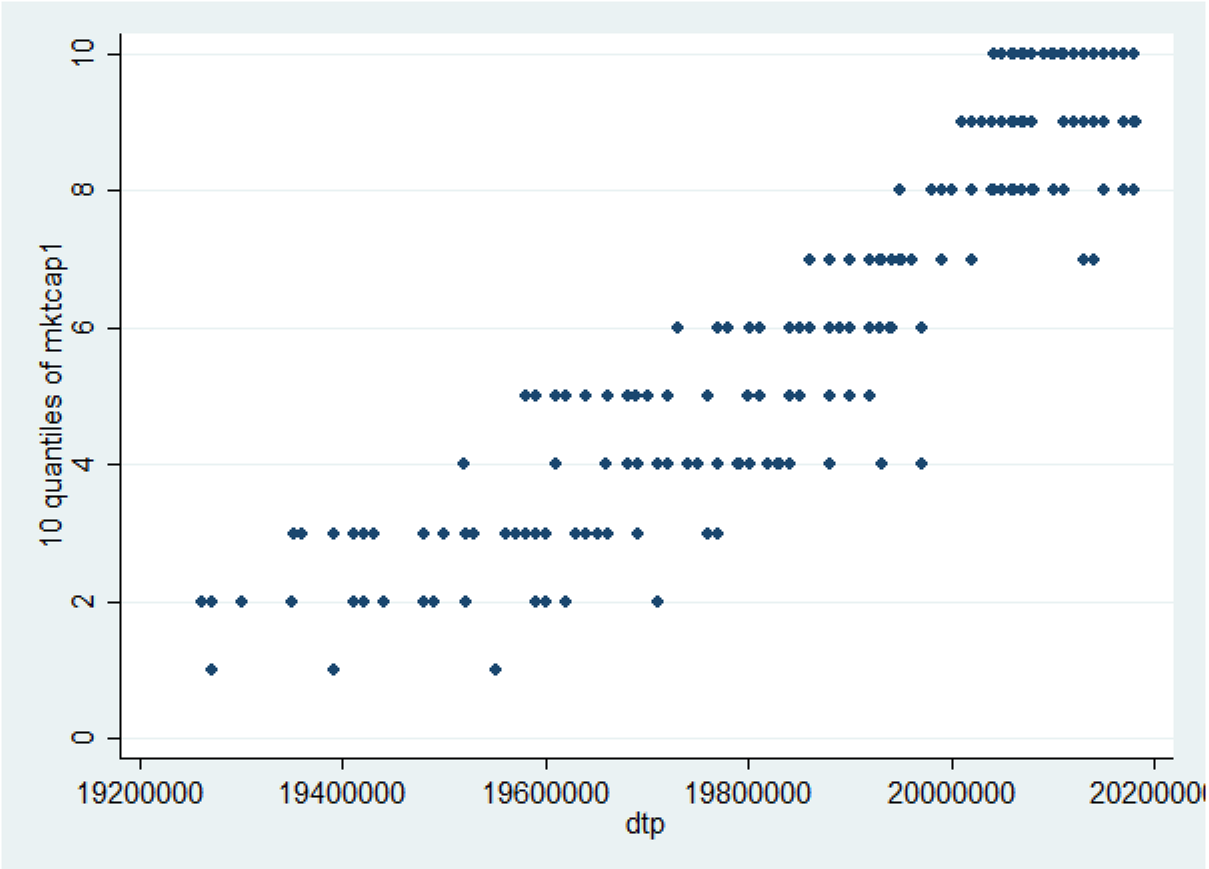


Figure 23. Rectangle Bottom Time-Varying Observations.

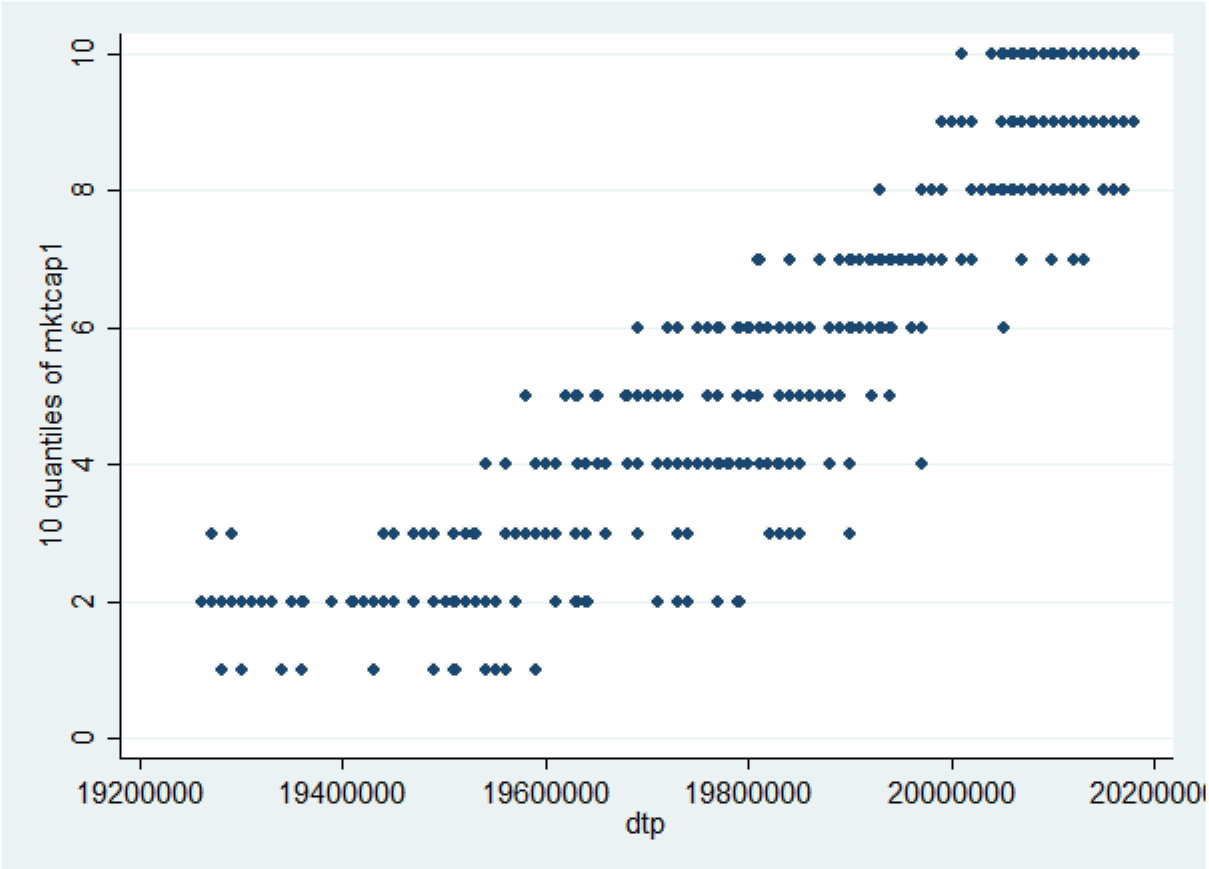


Figure 24. Rectangle Top Time-Varying Observations

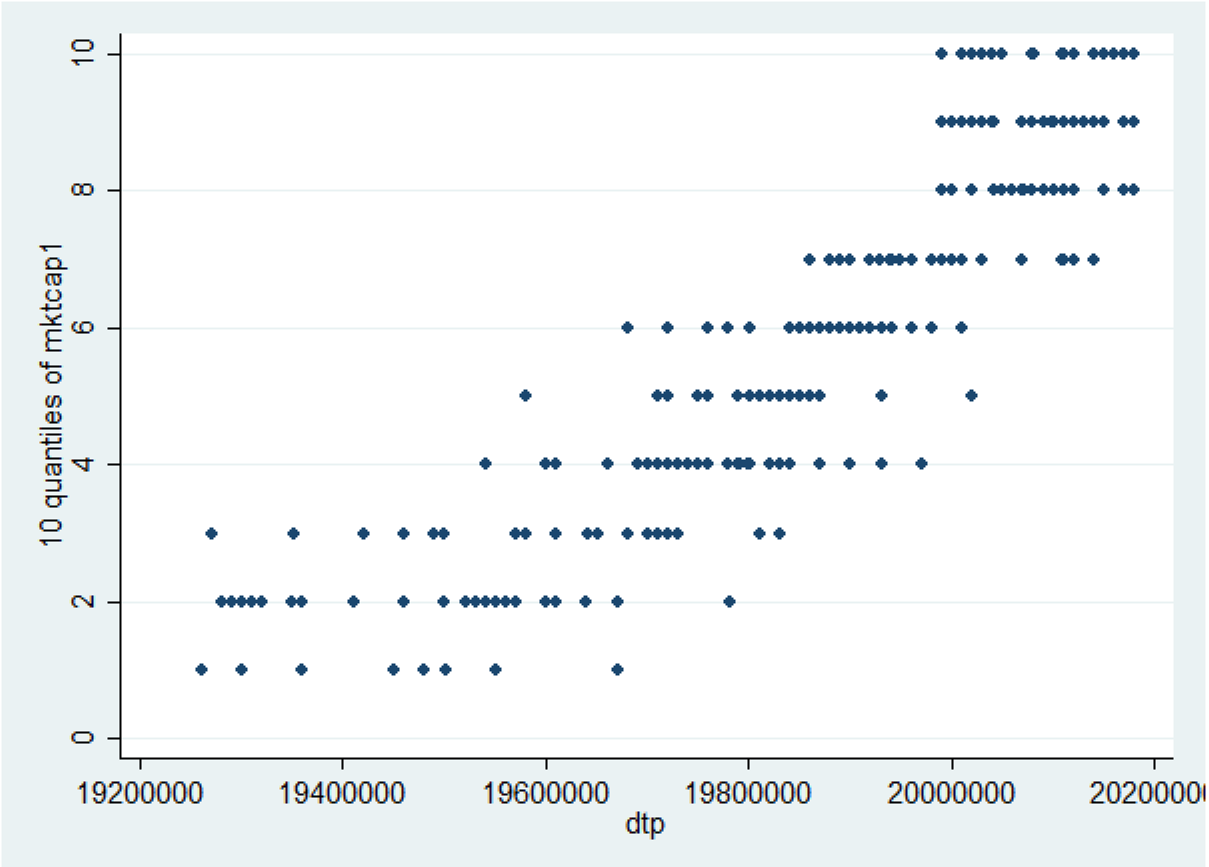


Figure 25. Triangle Bottom Time-Varying Observations

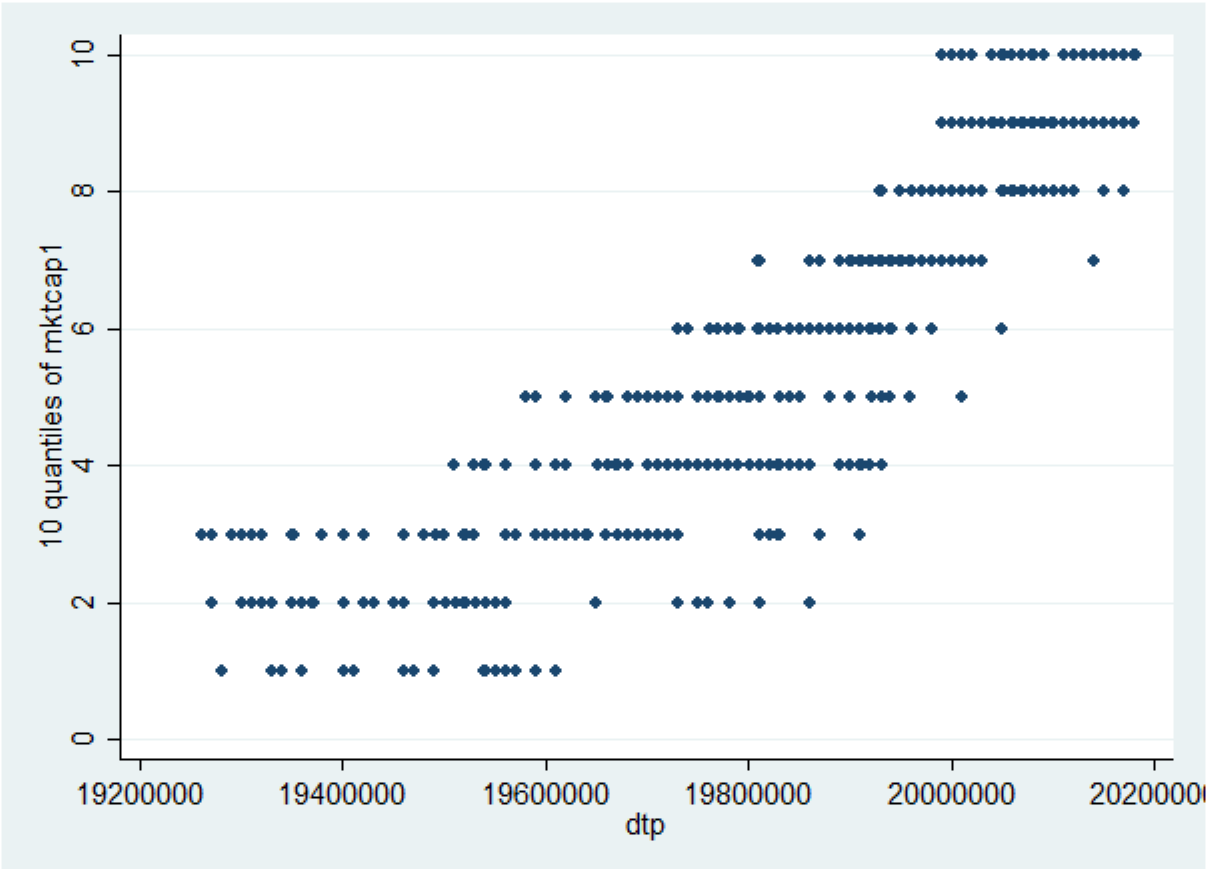


Figure 26. Triangle Top Time-Varying Observations

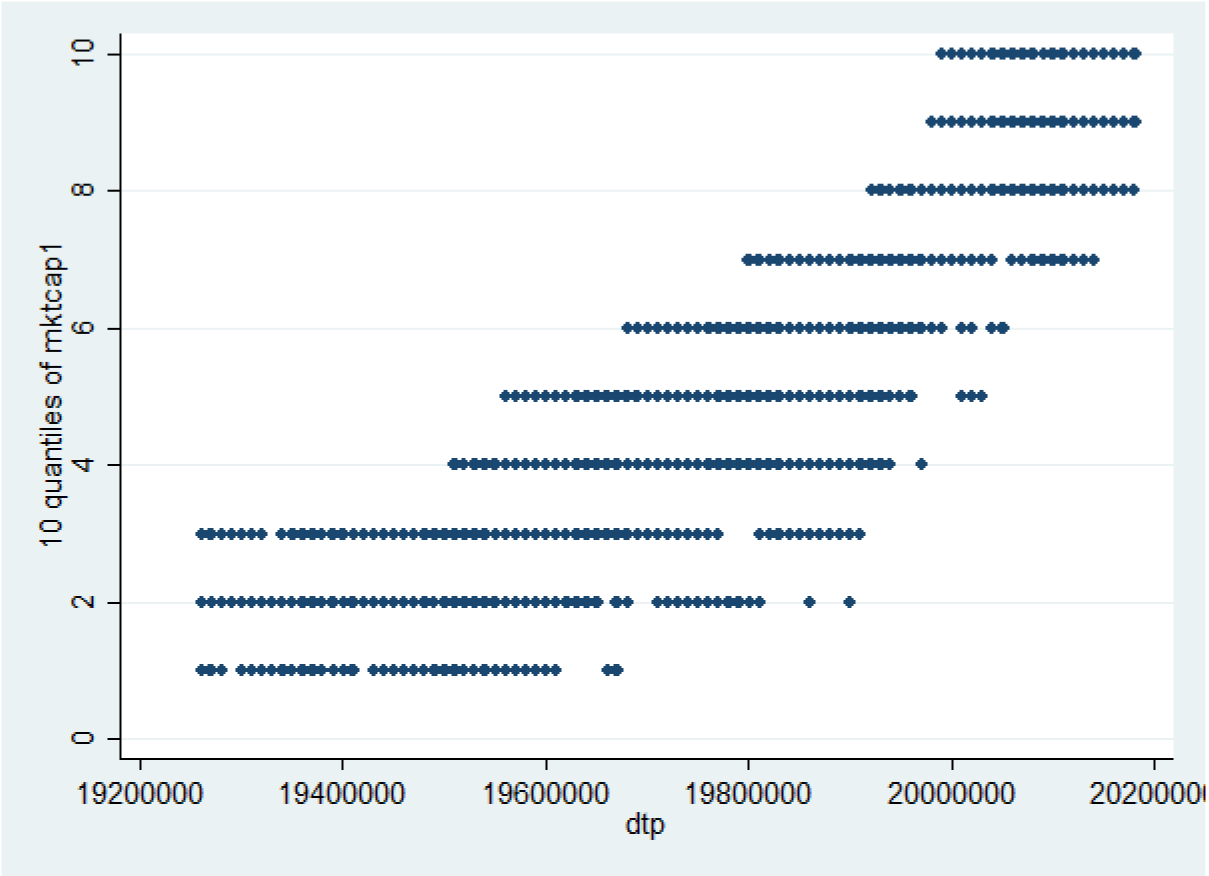


Figure 27. All Time-Varying Pattern Observations

The patterns seem to be persistent across time. The larger deciles appear to be more frequent in more recent years.

This tells a different story economically, that the returns are dependent on time since the sizes are all large gap. Economically it would make sense that a large cap has a larger market capitalization today than in 1935.

Chapter 3 "Predictability of Technical Stock Price Pattern Classification."

1. Introduction

This section discusses the viability of digital image processing methods to uncover nonrandom patterns in financial data. Ordinary time series methods would not otherwise be able to uncover the patterns discussed in chapter one (i.e. autoregressive ar1). These patterns are fit with a kernel density estimator (smoothing parameter) derived from stock prices. Digital image processing may be able to replicate this process by first converting a digital image into binary values (1 and 0). These values represent a pixel being on (black) or off (white). If a pixel is off it contains no information.

This study has potential benefits for both academia and industry. The academic benefits include gaining pattern statistics and confidence intervals. Aggregating any bias between professional analyst recommendation and the output of the classifier is another benefit. Not having to replicate the rigorous methodology for bandwidth selection and bypassing the need to rely on a trained technical analyst to engage in empirical research. We can upload the weights and have a ready to go classifier that will run on any program with SVM and image processing.

For industry, the benefits include program trading. This may be useful in using hybrid patterns to optimize profit over out of sample data and developing a trading strategy.

The potential contribution of such methods includes aggregating the output between professional analyst recommendations, speed and accuracy implications, gaining pattern statistics (confidence intervals). These implications are a contribution to both academia and industry via a better understanding of nonlinear visual chart patterns, and program trading.

The potential advantage of using image processing methods is many. Normal time series methods would not be able to uncover these patterns. We can aggregate any bias between the output and professional recommendation, we can gain confidence intervals for the patterns, we can forecast with them, we can also gain hybrid patterns based on probability thresholds. We can also uncover the patterns on images where the raw data is not available. This has implications for academia and industry including market efficiency research and program trading.

The paper we have provided shows that the digital image classifier can successfully uncover patterns that were previously used and identified by kernel estimation. We show with a high degree of confidence that it can uncover the patterns. We also show that the bandwidth can be set objectively and rigorously in smoothing data which is commonly done by visual inspection. We have provided a framework that can feed into both academia and industry. Via pattern classification, pattern statistics, aggregating bias between the output and professional analyst recommendation, and program trading.

Ever since computers were invented, we have wondered whether they might be made to learn. If we could understand how to program them to learn – to improve automatically with experience – the impact would be dramatic. Imagine computers learning from medical records which treatments are most effective for new diseases, houses learning from experience to optimize energy costs based on the particular usage patterns of their occupants, or personal software assistants learning the evolving interests of their users in order to highlight especially relevant stories from the online morning newspaper. A successful understanding of how to make computers learn would open up many new uses of computers and new levels of competence and customization. And a detailed understanding of information processing algorithms for machine learning might lead to a better understanding of human learning abilities (and disabilities) as well.

We do not year know how to make computers learn nearly as well as people learn. However, algorithms have been invented that are effective for certain types of learning tasks, and a theoretical understanding of learning is beginning to emerge. Many practical computer programs have been developed to exhibit useful types of learning, and significant commercial applications have begun to appear. For problems such as speech recognition, algorithms based on machine learning outperform all other approaches that have been attempted to date. In the field known as data mining, machine learning algorithms are being used routinely to discover valuable knowledge from large commercial databases containing equipment maintenance records, loan applications, financial transactions, medical records, and the like. As our understanding of computers continues to mature, it seems inevitable that machine learning will play an increasingly central role in computer science and computer technology.

The patterns LMW (2000) study are rule-based and non-linear, but common in technical trading circles. Standard time series analysis would fail to detect the classic patterns. Technicians employ the use of charts to visually detect non-linearities. The aim of this research is to train a computer to visually detect non-linearities common in technical analysis and mimic the technician's ocular facilities. With current digital image technology, it is possible to show pictures to a computer and have it process data as a human would see it. A digital classifier that can visually recognize common patterns in stock price data is not only novel but would have commercial and academic applications.

A properly trained classifier can recognize patterns in digital images which would be unrecognized by untrained technical analysts and ordinary time series methods. It would be more robust to the pattern definitions in chapter 1 when aggregated across professional analyst recommendations. It would also allow us to optimize profitability over hybrid patterns in out of sample testing. For accuracy, we would optimize the threshold for pattern identification.

The advantages of digital image processing over analog image processing are that analog image processing is left up to the interpreter. Digital image processing takes a nonparametric problem (analog image) and makes it parametric (digital image) by turning it into 1's and 0's. This allows for a variety of statistical methods to uncover the information and condenses a digital image in binary information. There are many techniques of turning an image into a vector of information (discussed in chapter 2). The purpose of this is to couple image processing with machine learning (which is also known as computer vision) and use it to mimic the way humans make decisions on financial data. The advantage of this to the current methods are one) aggregating bias between professional analyst recommendation and the classifier output. Thus making it a more robust pattern database. Two) gaining pattern statistics such as confidence intervals and hybrid patterns. Three) benchmarking for potential speed and accuracy. Four) gaining pattern profitability. Five) forecasting future returns based on the occurrence of chart patterns. Traditional time series methods (AR1, etc.) would not otherwise detect the patterns.

2. Literature

Support Vector Machines recognize a positive class from a negative class (Platt, 1999). This may be mapped through a sigmoid function to obtain the probability that a specific example belongs to a given class (Platt, 1999).

Some studies take a sliding window approach where they slide a trained classifier across the larger window (Dilal Triggs (2005)). Other studies locate high probability regions to test a trained classifier using convolutional neural networks (Redmon, (2016)).

Some extreme examples of image processing and convolutional neural networks are shown by Mnih, et al. (2013) convolutional neural networks can learn how to play Atari games from watching videos and adjusting the penalty/loss function. These videos of games are just digital images shown at a fast frame rate. The AI can break the information into pixel values and process it with a high degree of accuracy.

Coupling advanced statistical learning theory with digital image processing is deemed computer vision. Current works in this area determine to find faces(Ai, Liang, Xu. (2001), Jones, Viola (2003)), pedestrian detection (Dalal and Triggs (2005)) object detection (Redmon, et al. (2016)), and self-driving vehicles. Other topics include learning to play video games from video feed (Minh et al. (2013)). A natural extension to this is automating technical analysis.

Methods by Dalal and Triggs (2005) uses support vector machines for digital image processing. Support Vector Machines (SVM), are a flexible implementation of classification (James, Witten, Hastie, Tibshirani (2013)). SVM is discussed in depth in Platt (1999), with applications in C language Joachims (1999). It was originally developed and discussed by Cortes, Vapnik (1995).

3. Statistical Learning Theory

Statistical learning theory discusses learning an outcome from input data. For quantitative data, it is called making predictions. For qualitative data, it is called classification. Classification may be applied to data such as digital images, speech, and textual analysis.

4. Computer Vision

Computer vision is the intersection between computer science, machine learning, probability, and statistical learning theory. Books on this topic include Bishop (2006); Mitchell (1997); MacKay (2003); Hastie, Tibshirani, and Fiedman (2005); Witten and Frank (2005); V. Vapnik (2013). Books on probability include Ross (2014). Books on digital image processing include Gonzalez and Woods (2002) and Burger and Burge (2008), Shapiro and Stockman (2000).

5. Classification Methodology

We take the 400 perms that aren't a part of the current DJIA holdings. We use them for training. We preprocess the data by getting extreme values from a digital image. We

We split our training data into two sections. One using 60% training and 40% testing, and another using 10% training and 90% testing. Both achieve similar accuracy in the high 80%'s for training and testing and is robust to additional machine learning methodologies. We use both support vector machines and logistic regressions. We use SVM for their flexible implementation and logistic regression for their fast training time.

We use the best available model from Mathematica (which ends up being a Logistic Regression) benchmarked on speed and accuracy. We compare the result to a Support Vector Machine, given that models particular flexibility and praise in the literature.

The purpose of this section is to train a digital image classifier to make predictions in data based on observations it hasn't seen before. We pre-process all of our pattern images to condense to the extrema from pixel values. This is used for training and testing. We show with a high degree of accuracy that new data can be classified to the correct (true) pattern and with a high degree of accuracy that training data can be classified to the correct (true) pattern. This has implications for both academia and industry. Namely, pattern statistics, and program trading.

We split the data into training and testing. For training, we use patterns from stocks, not on the DJIA. These come from CRSP permno's 10001-10713 and about 150 random draws. We then use the (held out) DJIA 30 for testing.

Some permno's don't have available patterns for testing. These are normally stocks without enough price history. 40% of our total patterns come from our 30 DJIA stocks. The majority of these are 'survivors.' The 60% remaining come from the 407 stocks, not in our sample for testing.

We 6721 patterns from 407 different premno's that aren't on the DJIA for training and have 4607 patterns that are on the 30 DJIA for testing. We use logistic regressions for training since the results are not highly varied and they are faster. Training takes 11.3 seconds for 6721 examples. The accuracy for training is 89.5%. If we use a Support Vector Machine for training it takes 1 minute and 27 seconds with 88.3% accuracy. Patterns found for training data are shown below. Head and Shoulders are the most observed.

Digital Image classification for 8 technical indicators detected among non DJIA stocks from 1925-2019. Each stock's price history is scanned for the occurrence of the following 8 technical indicators: head-and-shoulders (HS), inverted head-and-shoulders (HIS), broadening top (BTOP), broadneing bottom (BBOT), triangle top (TTOP), triangle bottom (TBOT), rectangle top (RTOP), rectangle bottom (RBOT).							
BBOT	BTOP	HS	HIS	RBOT	RTOP	TBOT	TTOP
391	956	1803	790	317	753	529	1182

Table 11. Training Data

There are 6702 patterns in total. The most often pattern in training data is the head-and-shoulders. The second most often is the triangle top. Followed by broadening top, inverse head-and-shoulders, rectangle top, triangle bottom, broadening bottom, rectangle bottom. Again the data is consistent that there are more top patterns than bottom patterns.

The classes used in training correspond with the pattern types. They are shown with probabilities in the table below. For a logistic regression and Support Vector Machine Respectively:

Training accuracy for 8 technical indicators detected among non DJIA stocks from 1925-2019. Each pattern is analyzed for accuracy using two methods of image classification, support vector machines and logistic regression. The following patterns are analyzed: head-and-shoulders (HS), inverted head-and-shoulders (HIS), broadening top (BTOP), broadneing bottom (BBOT), triangle top (TTOP), triangle bottom (TBOT), rectangle top (RTOP), rectangle bottom (RBOT).								
Method	BBOT	BTOP	HS	HIS	RBOT	RTOP	TBOT	TTOP
Logistic Regression	.92	.94	.93	.93	.65	.67	.92	.91
SVM	.92	.93	.93	.93	.64	.66	.92	.92

Table 12. Training Results

The classifier misses rectangle bottom, and rectangle tops. The two classifiers are close in accuracy. Logistic regressions have higher overall accuracy than SVM so they are used in analyzing which patterns are misclassified and where. The truth tables for a logistic regression and SVM are omitted but are similar to that for the testing data. The following figure shows the misclassification of the 8 technical patterns.

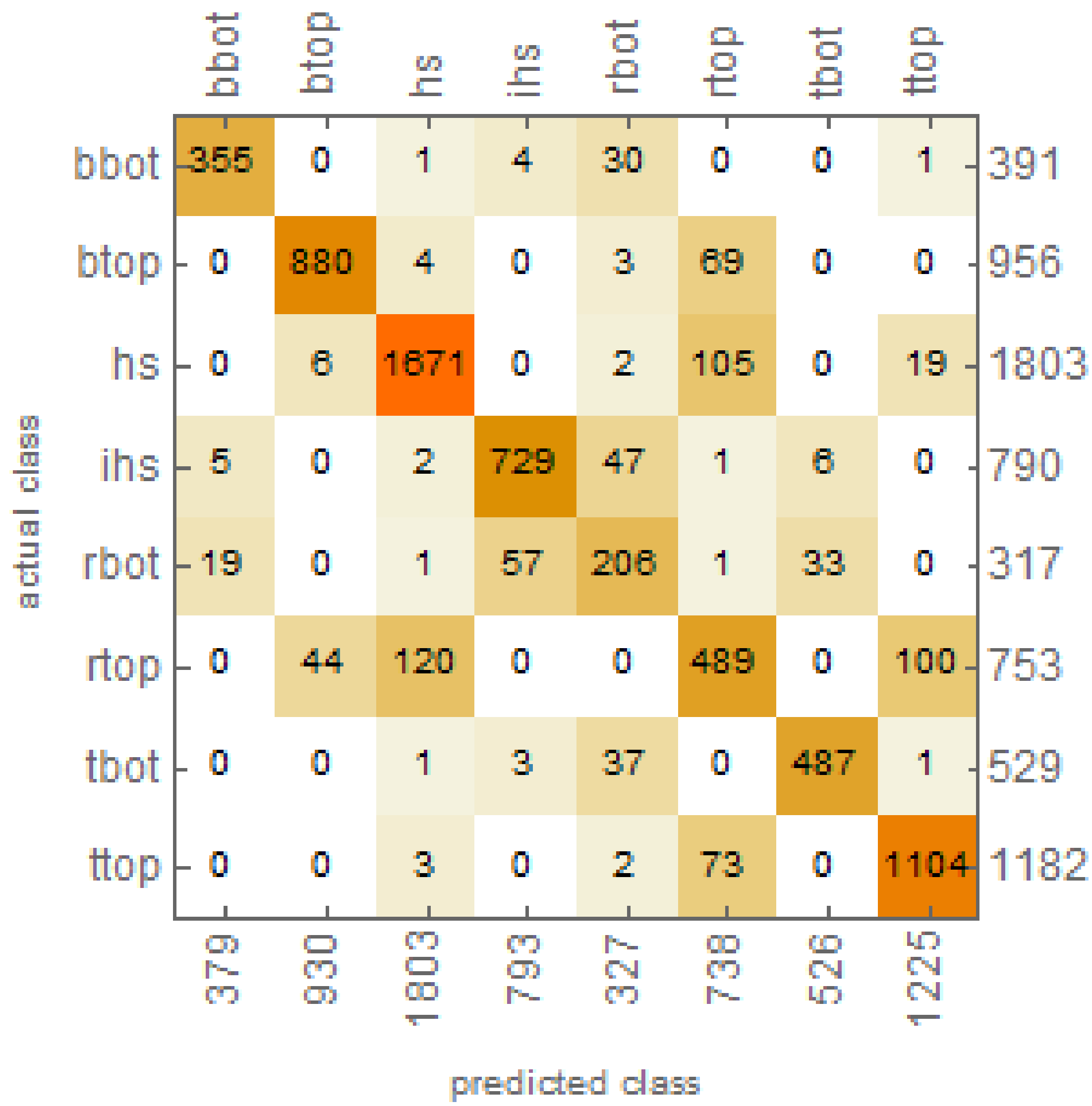


Figure 28. Truth Table Training Data using Logistic Regressions

On the testing data we don't have 100% accuracy because a number of the rectangle top patterns are being confused with a broadening top, head-and-shoulders, or triangle top.

We extend the classifiers to include

Test data from the DJIA components are considered. The number of patterns are 4620. The same pattern classes are used from training. They are shown with probabilities below. These are the same patterns from the above section. The classifier hasn't seen these patterns before and it's task is to define which pattern it is based on five consecutive extrema.

Testing accuracy for 8 technical indicators detected among DJIA stocks from 1925-2019. Each pattern is analyzed for accuracy using two methods of image classification, support vector machines and logistic regression. The following patterns are analyzed: head-and-shoulders (HS), inverted head-and-shoulders (HIS), broadening top (BTOP), broadneing bottom (BBOT), triangle top (TTOP), triangle bottom (TBOT), rectangle top (RTOP), rectangle bottom (RBOT).								
Method	BBOT	BTOP	HS	HIS	RBOT	RTOP	TBOT	TTOP
Logistic Regression	.91	.92	.93	.93	.67	.69	.91	.90
SVM	.91	.91	.93	.93	.67	.67	.91	.90

Table 13. Testing Data

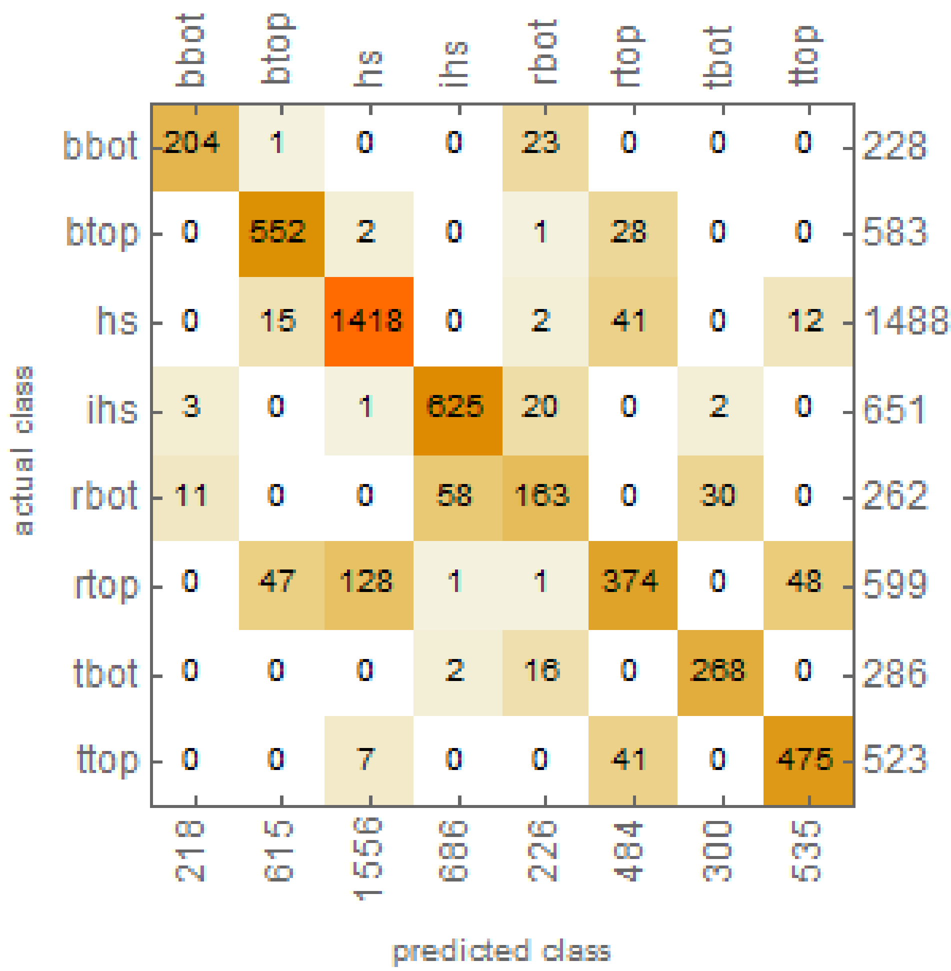


Figure 29. Truth Table for Tested Data using Logistic Regressions

Truth tables for a logistic regression and SVM are shown above. The implications are similar to the training data and are focused around the misclassification of a head and shoulders for a rectangle top and vice-versa. This seems to be a problem and is also discussed in literature. Noting from Edwards and Magee (2007) that a head and shoulders can slide in to a rectangle formation.

Training and testing aren't much different. The most misclassified patterns are rectangles. The most accurate patterns out of sample are the head and shoulders (inverse head and shoulders). In sample isn't as important because we can never test on in sample data.

The implications on the returns are shown below. We use the highest predicted class at the time of pattern completion for entering the market. Missclassification only plays a role when the pattern for a top is confused with a pattern for a bottom. Most of the misclassified rectangle tops are misclassified as a head and shoulders which has the same outcome on decision making. We still would enter a short position. The other misclassified rectangle tops are for a broadening top, and a triangle top. Again there is no information loss. There is one that is misclassified for an inverse head and shoulders which would be a problem, so long as the resulting trade were in the wrong direction. (i.e. maybe it's trying to tell us something?)On the other way around, there are 47 triangle tops that are misclassified as rectangle tops.

The implications for trading and returns are only affected to the extent that the patterns change sign. As long as the reversal is of the same type and the pattern is recognized in the same place it won't affect the outcome on returns.

Referring back to the literature from chapter one. It is well justified by Edwards and Magee (1966,2007) that these patterns can be confused. A lot of the visual properties are similar for the patterns.

The above two graphs show in sample and out of sample data respectively. The patterns found are similar in accuracy and the errors are in the same places.

We proceed with conditioning the returns on each of the patterns as part of the output of our original classifier. The summary statistics are shown in the table below:

Enacted returns for 8 technical indicators detected among DJIA stocks from 1925-2019. Each pattern is enacted from the conditional returns and probabilities from classification. The following patterns are analyzed: head-and-shoulders (HS), inverted head-and-shoulders (HIS), broadening top (BTOP), broadneing bottom (BBOT), triangle top (TTOP), triangle bottom (TBOT), rectangle top (RTOP), rectangle bottom (RBOT).		
Moment	Classification Conditional Returns	Raw Conditional Returns
Mean	0.0697	0.0684
Standard Deviation	1.56	1.57
Median	0	0
Skewness	21.39	22.03
Kurtosis	4.830	4.814

Table 14. Classifier Summary Statistics

This shows that the two distributions, true returns and out of sample output from the classifier are no different. The main difference here is that instead of being 100% conditioned on the return, the conditional distribution is broken among all of the classes.

When we look at the effect on individual patterns, we see that some of the patterns have returns outside of their true class. This isn't possible without classification. When we restrict the returns to just the patterns found as truth. We get results similar to the true results without classification.

Enacted returns for 8 technical indicators detected among DJIA stocks from 1925-2019. Each pattern is enacted from the conditional returns and probabilities from classification. The following patterns are analyzed: head-and-shoulders (HS), inverted head-and-shoulders (HIS), broadening top (BTOP), broadneing bottom (BBOT), triangle top (TTOP), triangle bottom (TBOT), rectangle top (RTOP), rectangle bottom (RBOT).								
Moment	BBOT	BTOP	HS	HIS	RBOT	RTOP	TBOT	TTOP
Mean	0.0020	-0.0009	0.0013	-0.0007	0.0009	0.0009	0.0012	0.0011
Standard Deviation	0.0176	0.0172	0.0155	0.0152	0.0123	0.0139	0.0142	0.0180

Table 15. Classifier Conditional Returns

These show that the conditional returns and standard deviations are close to the result without digital image classification. The potential advantage of this (why use it?) is that we get the confidence intervals for each pattern. When one true pattern is confused for another the statistic changes as long as the reversal type is different.

6. Robustness Checks

For an exercise we reduce the amount of training data to 1000 observations and place the remaining 5702 unused training patterns on the testing data. We show that the classifier has similar accuracy.

Results from holding out additional training data to reduce the number of examples to 1000 (from 6721) doesn't change the results on a logistic regression. We add the unused training data to our test sample and have similar accuracy of 87.67%. Using SVM the classifier has 84.9% accuracy in sample and 84.28% out of sample.

This shows the ability of a digital image classifier to learn the pattern rules from association without them being explicitly defined. This is true of the early examples as well but this has a lower split with the data. 1,000 training, 10,341 testing.

Testing accuracy for 8 technical indicators detected among DJIA stocks from 1925-2019, and non-DJIA Stocks from 1925-2019 from 1000 training observations. Each pattern is analyzed for accuracy using two methods of image classification, support vector machines and logistic regression. The following patterns are analyzed: head-and-shoulders (HS), inverted head-and-shoulders (HIS), broadening top (BTOP), broadneing bottom (BBOT), triangle top (TTOP), triangle bottom (TBOT), rectangle top (RTOP), rectangle bottom (RBOT).								
Method	BBOT	BTOP	HS	IHS	RBOT	RTOP	TBOT	TTOP
Logistic Regression	.9146	.9186	.9304	.9200	.6296	.6545	.9203	.9082
SVM	.8796	.9142	.8961	.9023	.5943	.5773	.8545	.8680

Table 16. Robustness Check 1, Data Split 10% Training 90% Testing

The truth tables are similar for both classifiers. The main advantage for improving this can be made in correctly classifying the rectangle bottom and rectangle top patterns.

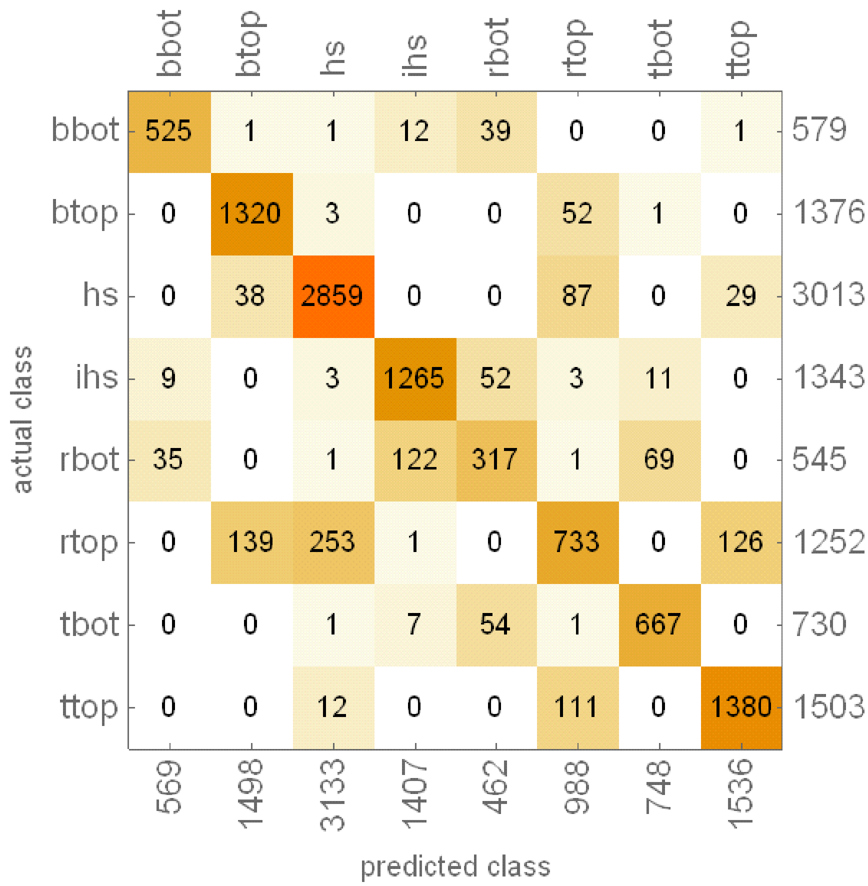


Figure 30. Truth Table on Extended Data with Logistic Regressions

The return implications for out of sample data are compared with the logistic regression classifier results to the truth (raw data). Given the high degree of accuracy we want to know if it affects the outcome for trading purposes. The returns from the out of sample data from classification are deemed ‘Classify’ and the true returns from the out of sample data are deemed ‘Raw.’ We enact the returns across the probabilities for each pattern. We also look at the true returns conditioned only on long returns and short returns.

The last problem we look at is splitting the data the other way with 90% training and 10% testing. To do this we omit 1000 patterns from the non-DJIA stocks and use the remaining patterns along with those that are on the DJIA for training. The purpose here is to see if we can improve our accuracy. We test Support Vector Machines, Logistic Regressions, and Gradient Boosted Trees.

In Sample accuracy, and out of sample accuracy are recorded. All three examples are 81% or better off of the first 1500 training observations. They reach a plateau which seems to be because of the rectangle patterns.

In the table below we check additional models including Random Forest (4.22s Training Time, 86.3% in sample accuracy), Gradient Boosted Trees (19.7s training time, 87.3% in sample accuracy), Logistic Regression (6.35s training time, 88.4% in sample accuracy), Support Vector Machine (1 min 44s training time, 88.0% in sample accuracy Neural Networks (1 min 52 second training time, 82.7% accuracy), Naïve Bayes (2.21s training time, 75.0% accuracy), Decision Tree (2.76s training time, 84.6% accuracy, Markov (5.57s training time, 69.0% accuracy), Nearest Neighbors (2.22s training time, 85.1% accuracy), Thus confirming our results of the ability of a digital image classifier to learn the pattern rules from labeled examples.

Model training time and sample accuracy for predicting 1000 observations of the following chart patterns: head-and-shoulders (HS), inverted head-and-shoulders (HIS), broadening top (BTOP), broadneing bottom (BBOT), triangle top (TTOP), triangle bottom (TBOT), rectangle top (RTOP), rectangle bottom (RBOT).		
Method	Training Time	Model Accuracy
Logistic Regression	6.53s	88.40%
Random Forest	4.22s	86.30%
Gradient Boosted Trees	19.7s	87.30%
Support Vector Machines	1 min 44s	88.00%
Neural Networks	1 min 52s	82.70%
Naïve Bayes	2.21s	75.00%
Decision Tree	2.76s	84.60%
Markov	5.57s	69.00%
Nearest Neighbors	2.22s	85.10%

Table 17. Model Training Time and Accuracy

This table shows that the fastest model is a Nearest Neighbor with 85.10% accuracy on known examples. The highest accurate model on known examples is a Decision Tree with 84.60%. SVM, Gradient Boosted Trees, Random Forest and Logistic Regressions have the highest Model Accuracy. We compare the results of these methods in predicting our patterns in the table below.

Testing accuracy for 8 technical indicators detected among 1000 non-DJIA stocks from 1925-2019, from 10341 training observations. Each pattern is analyzed for accuracy using two methods of image classification, support vector machines and logistic regression. The following patterns are analyzed: head-and-shoulders (HS), inverted head-and-shoulders (HIS), broadening top (BTOP), broadneing bottom (BBOT), triangle top (TTOP), triangle bottom (TBOT), rectangle top (RTOP), rectangle bottom (RBOT).									
Method	BBOT	BTOP	HS	IHS	RBOT	RTOP	TBOT	TTOP	Model Accuracy
Logistic Regression	95.0000%	93.2039%	92.4731%	93.2642%	61.9718%	64.6288%	93.4911%	93.0946%	88.70%
Random Forest	95.1220%	95.2978%	92.7536%	95.4315%	65.6250%	66.0714%	94.1176%	94.3878%	90.10%
Gradient Boosted Trees	89.7436%	92.9936%	89.2794%	91.0891%	53.1250%	58.7156%	90.0585%	93.2292%	86.40%
Support Vector Machines	93.8272%	93.2907%	90.1304%	94.2408%	61.3333%	61.1111%	91.5663%	92.4675%	87.00%

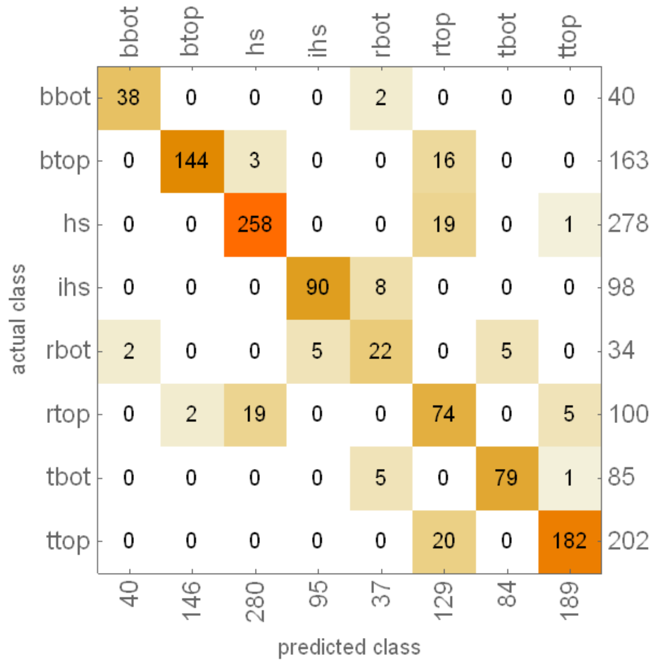
Table 18 Robustness Check 2. 90% Training 10% Testing

This is where SVM really shows its flexibility. The classifiers all perform well but SVM is more accurate across all pattern classes.

Gradient Boosted Trees are more accurate than Logistic Regressions for the Triangle Top Pattern but is similar in accuracy to both other methods.

SVM Performs the best with 90.10% overall model accuracy. All 3 classifiers perform similarly on the rectangle patterns. We can see from their truth tables that the patterns are being misclassified for head and shoulders. Part of the misclassification with rectangle bottoms might be that there are so few patterns found.

Logistic Regression (a)



Random Forest (b)

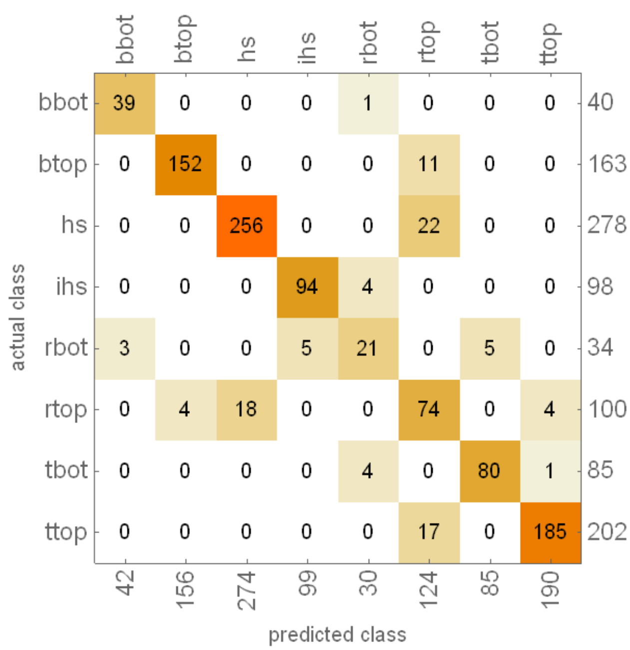


Figure 31. Truth Tables for 90% Training, 10% Testing Data Split

Gradient
Boosted
Trees (c)

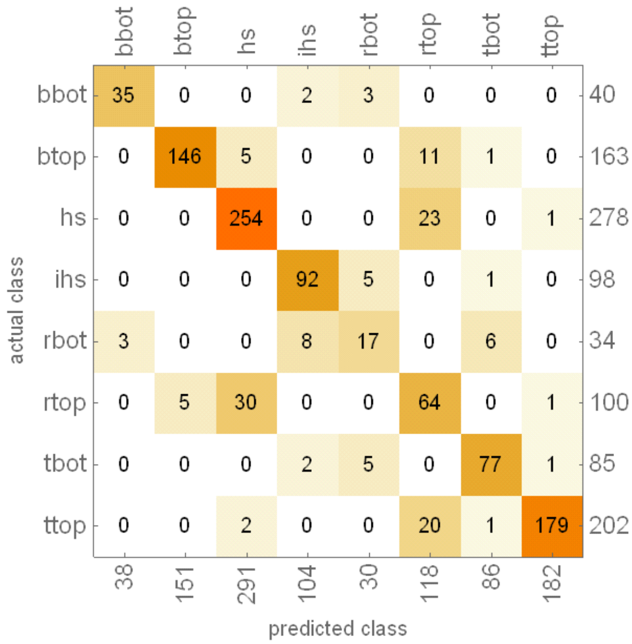


Figure 31. Truth Tables for 90% Training, 10% Testing Data Split

Support
Vector
Machines
(d)

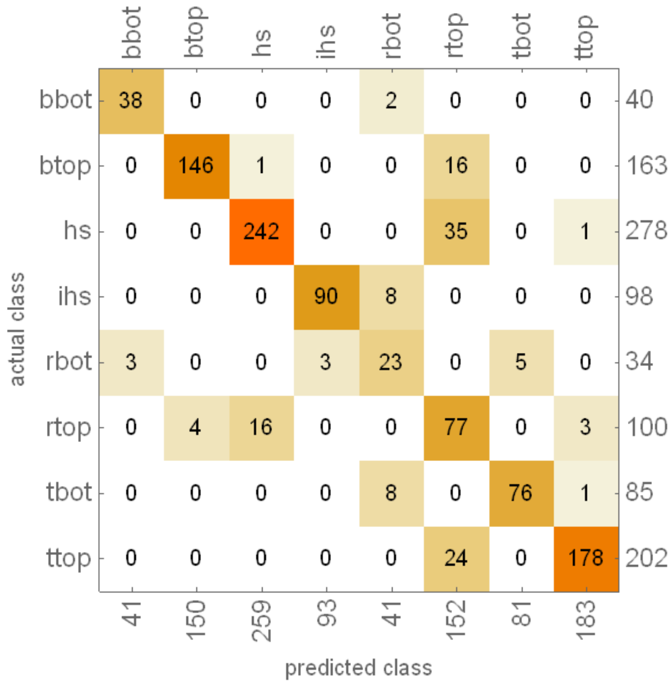


Figure 31. Truth Tables for 90% Training, 10% Testing Data Split

This shows the misclassification is similar to all models but SVM performs better with more training data. This seems to be an advantage for the model and is likely the model that would be used on a larger dataset with more pattern observations.

Future work would likely employ the use of more patterns for training so that the 30 DJIA components result in 10% testing data. Or, classifying the entire CRSP database which would require the use of a super computer.

The above analyses show the viability of digital image classification in predicting visual nonlinear chart patterns. The training data we used is from the chart patterns identified from image data in part 2. We use non Dow Jones Industrial Average stocks to predict patterns from the Dow Jones Industrial Average stocks. The resultant classifier is highly accurate, robust to additional machine leraning models, and data splits. We show that the classifier can predict either training or testing data out of sample and is not time variant. Final conclusions for all three chapters are in the section below.

7. Pattern Prediction

In this section we take pattern windows with enough extrema to complete a pattern but that do not meet any of the rules defined in section one. We use this as testing data for our machine learning classifier. Our goal is to see how well the model can predict patterns and hybrid patterns of the nonlinear technical chart patterns previously identified.

We take all of the nonpattern trading data from our methodology in chapter two. Namely plotting the prices and smoothed prices. We find the extrema values and keep the windows with enough extrema to complete a pattern. We discard the pattern windows that have completed patterns as we have already analyzed those.

The contribution here is that we can get pattern statistics from extrema that wouldn't otherwise meet the pattern rules identified by Lo, Mamaysky and Wang 2000.

For conditional returns we can take the patterns that are a top and condition on when they are a higher probability than the patterns that are a bottom. We can do the same for the bottom patterns and condition on them when they are a greater probability than the patterns that are a top. We can call these PTOP and PBOT. They will be our 'hybrid' patterns.

The contribution is that we can test if these 'hybrid' patterns have any predictive ability. We can again compare the conditional distribution of one day returns to the unconditional distribution (every other day). We can also gather summary statistics. We can also test the goodness of fit statistic and the Kolmogorov-Smirnov for these patterns.

8. Conclusion

We address two calls for future work in this paper. One for advances in the rigorous methodology for finding patterns including using local polynomial regressions. We also address finding new ways to uncover the patterns using digital images. A third contribution is that we use an additional measure of uncovering the patterns which is through adding machine learning to predict the pattern trading rules. We find that a digital image classifier shows both high in sample accuracy and out of sample accuracy.

In part one we show a high degree of non-randomness in uncovering the nonlinear visual patterns. Over 9500 daily price observations on a single security, we find that at most 68 patterns are uncovered (Head and Shoulders).

In part two we show that image processing can uncover the extrema points represented by a single pixel value (per extremum) and the resultant patterns are the same as in kernel smoothing methods which are represented by price and time observations. Prior work shows that the patterns are difficult to uncover on random geometric Brownian motion.

We match the results to LMW, 2000 by extending the data to 98 years. We use local polynomial regressions to uncover the patterns (previous work by LMW (2000)) used kernel smoothing. Then we use digital image processing to uncover the extreme values from very thin plots of the values. Finding extrema from these values allows us to use them in the pattern rules as defined previously. All returns are found out of sample since patterns are completed 3 days before the returns are taken.

Every pattern has at least 4 deciles with 10% or more matches when tabulating the returns. This shows that at 40%-50% of the deciles they contain informational value.

In part 3 we show is that digital image processing can be useful in uncovering the patterns. We show that image classification can uncover pattern trading rules without them being explicitly defined via supervised learning. We provide 5 consecutive extrema, and the true pattern type. We use state of the art machine learning tools such as Support Vector Machines, Logistic Regressions, and Gradient Boosted Trees to uncover the pattern rules with high success.

We show that using classifier out to condition on the pattern probabilities provides a reduced market exposure when patterns are showing conflicting sign. The resultant effect is a slightly higher average return and slightly lower standard deviation. We show that the number of patterns found on the DJIA is similar to that of image processing without classification which proves the validity of a digital image classifier to be useful in predicting stock price patterns. If we had a different number of patterns, or different/lower returns we would suggest that classification cannot uncover the patterns. This further shows the randomness of the patterns.

The added feature of using the extrema in a digital image classifier give us pattern confidence intervals. They allow us to make inferences on new data that hasn't been observed without explicitly defined rules. This feeds in to future implication for academia and industry. Namely, program trading and market efficiency research. Additional patterns may be tested using the framework outlined in Edwards and Magee (2018) and bias from the classifier may be benchmarked against professional analyst recommendations.

The advantages of digital image processing over the current methods for automatic pattern recognition are that it more closely resembles the human learning process, doesn't require raw data (once a image classifier is trained), can be aggregated between the bias of the output and a professional analyst recommendation, can easily be adjusted to multiple time frames and additional patterns.

We show that splitting the data between 60% training, 40% testing, 10% training, 90% testing, 90% training, and 10% testing all yield similar results. SVM is a better performer with more training data (90% training, 10% testing) while Logistic Regressions perform better on fewer pattern observations. From a practical standpoint, more pattern observations would likely be used for training purposes and Support Vector Machines show promising results.

References

1. Ai, Haizhou, Luhong Liang, and Guangyou Xu. "Face detection based on template matching and support vector machines." *Image Processing, 2001. Proceedings. 2001 International Conference on*. Vol. 1. IEEE, 2001.
2. Akaike, Htrotugu. "Maximum likelihood identification of Gaussian autoregressive moving average models." Biometrika 60.2 (1973): 255-265.
3. Avramov, Doron, Guy Kaplanski, and Avanidhar Subrahmanyam. "Stock Return Predictability: New Evidence from Moving Averages of Prices and Firm Fundamentals." *Available at SSRN 3111334* (2018).
4. Bishop, Christopher M. Pattern recognition and machine learning. springer, 2006.
5. Blume, Lawrence, David Easley, and Maureen O'hara. "Market statistics and technical analysis: The role of volume." The Journal of Finance 49.1 (1994): 153-181.
6. Brock, William, Josef Lakonishok, and Blake LeBaron. "Simple technical trading rules and the stochastic properties of stock returns." *The Journal of finance* 47.5 (1992): 1731-1764.
7. Brown, David P., and Robert H. Jennings. "On technical analysis." The Review of Financial Studies 2.4 (1989): 527-551.
8. Burger, W., and M. J. Burge. "Digital image processing. Texts in computer science." (2008).
9. Campbell, J. Y., Campbell, J. W., Lo, A. W. C., MacKinlay, A. C., Champbell, J. J., LO, A. A., & Lo, A. W. (1997). *The econometrics of financial markets*. princeton University press.
10. Chang, PH Kevin, and Carol Osler. *Evaluating chart-based technical analysis: the head-and-shoulder pattern in foreign exchange*. No. 9414. 1994.
11. Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine learning*, *20*(3), 273-297.
12. Covel, Michael. Trend following (updated edition): Learn to make millions in up or down markets. FT Press, 2009.
13. Craven, Peter, and Grace Wahba. "Estimating the correct degree of smoothing by the method of generalized cross-validation." Numerische Mathematik 31 (1979): 377-403.

14. Dalal, N., & Triggs, B. (2005, June). Histograms of oriented gradients for human detection. In *international Conference on computer vision & Pattern Recognition (CVPR'05)* (Vol. 1, pp. 886-893). IEEE Computer Society.
15. Dunis, Christian L., et al. "Forecasting IBEX-35 moves using support vector machines." Neural Computing and Applications23.1 (2013): 229-236.
16. Du Plessis, Jeremy. "The definitive guide to point and figure." Harriman House, Petersfield UK (2005).
17. Edwards, Robert D., W. H. C. Bassetti, and John Magee. Technical analysis of stock trends. CRC press, 2007.
18. Edwards, Robert D. "Technical Analysis of Stock Trends, John Magee." Boston, Mass (1966).
19. Edwards, Robert D. "& John Magee." Technical Analysis of Stock Trends (1948).
20. Elder, Alexander. Come into my trading room: A complete guide to trading. Vol. 146. John Wiley & Sons, 2002.
21. Elder, Alexander. Trading for a living: psychology, trading tactics, money management. Vol. 31. John Wiley & Sons, 1993.
22. Fama, E. (1970). Efficient capital markets: A review of theory and empirical work. Journal of Finance, 25, (2), 383–417.
23. Fama, E. (1998). Market efficiency, long-term returns, and behavioral finance. Journal of Financial Economics, 49, (3), 283-306
24. Fama, E. F., M. F. Blume. 1966. Filter rules and stock market trading. Journal of Business 39226–241.
25. Gasser, Theo, and Hans-Georg Müller. "Kernel estimation of regression functions." Smoothing techniques for curve estimation. Springer, Berlin, Heidelberg, 1979. 23-68.
26. Gonzalez, Rafael C., and E. Richard. "Woods, digital image processing." ed: Prentice Hall Press, ISBN 0-201-18075 8 (2002).
27. Goyal, Amit, and Ivo Welch. "A Comprehensive Look at the Empirical Performance of Equity." (2008).
28. Grundy, Bruce D., and Maureen McNichols. "Trade and the revelation of information through prices and direct disclosure." *The Review of Financial Studies* 2.4 (1989): 495-526.

29. Han, Yufeng, Ke Yang, and Guofu Zhou. "A new anomaly: The cross-sectional profitability of technical analysis." Journal of Financial and Quantitative Analysis 48.05 (2013): 1433-1461.
30. Härdle, W. (1990). *Applied nonparametric regression* (No. 19). Cambridge university press.
31. Härdle, Wolfgang, and Oliver Linton. "Applied nonparametric methods." Handbook of econometrics 4 (1994): 2295-2339.
32. Hastie, Trevor, et al. "The elements of statistical learning: data mining, inference and prediction." The Mathematical Intelligencer 27.2 (2005): 83-85.
33. Haug, M. & Hirschey, M. (2006). The January effect. Financial Analysts Journal, 62, (5), 78-88.
34. Hurvich, Simonoff, Chih-Ling Tsai, "Smoothing parameter selection in nonparametric regression using an improved Akaike information criterion"J.R. Statlist. Soc. B (1998) 60, Part 2, pp. 271-293
35. Hsieh, David A. "Testing for nonlinear dependence in daily foreign exchange rates." Journal of Business 62.3 (1989).
36. Jasemi, Milad, and Ali M. Kimiagari. "An investigation of model selection criteria for technical analysis of moving average." Journal of Industrial Engineering International 8.1 (2012): 1-9.
37. James, Gareth, Witten, Daniela, Hastie, Trevor, Tibshirani, Robert "An Introduction to Statistical Learning with Applications in R." Springer Texts in Statistics. (2013)
38. Joachims, T. (1999). Svmlight: Support vector machine. *SVM-Light Support Vector Machine http://svmlight. joachims. org/, University of Dortmund*, *19*(4).
39. Jones, Michael, and Paul Viola. "Fast multi-view face detection." Mitsubishi Electric Research Lab TR-20003-96 3.14 (2003): 2.
40. kedd: Kernel estimator and bandwidth selection for density and its derivatives. R package version 1.0. 3 (2015)
41. Kulkarni, Sanjeev, and Gilbert Harman. An elementary introduction to statistical learning theory. Vol. 853. John Wiley & Sons, 2011.
42. Linton, D. (2010). Cloud Charts. *Updata*, *London*.

43. Lo, Andrew W., Harry Mamaysky, and Jiang Wang. "Foundations of technical analysis: Computational algorithms, statistical inference, and empirical implementation." The journal of finance 55.4 (2000): 1705-1765.
44. Lo, Andrew W., and A. Craig MacKinlay, 1988, Stock market prices do not follow random walks: Evidence from a simple specification test, Review of Financial Studies 1, 41–66.
45. Lo, Andrew W., and A. Craig MacKinlay, 1997, Maximizing predictability in the stock and bond markets, Macroeconomic Dynamics 1, 102–134.
46. Lo, Andrew W., and A. Craig MacKinlay, 1999, A Non-Random Walk down Wall Street (Princeton University Press, Princeton, N.J)
47. MacKay, David JC, and David JC Mac Kay. Information theory, inference and learning algorithms. Cambridge university press, 2003.
48. Malkiel, Burton Gordon. A random walk down Wall Street: including a life-cycle guide to personal investing. WW Norton & Company, 1999.
49. Mitchell, Tom M. "Machine learning. 1997." Burr Ridge, IL: McGraw Hill 45.37 (1997): 870-877.
50. Mnih, Volodymyr, et al. "Playing atari with deep reinforcement learning." arXiv preprint arXiv:1312.5602 (2013).
51. Neely, Christopher J., et al. "Forecasting the equity risk premium: the role of technical indicators." *Management science* 60.7 (2014): 1772-1791.
52. Osler, C. L., & Chang, P. H. (1995). Head and shoulders: Not just a flaky pattern. *FRB of New York staff report*, (4).
53. Platt, John C. "12 fast training of support vector machines using sequential minimal optimization." *Advances in kernel methods* (1999): 185-208.
54. Pring, Martin J. *Technical analysis explained: The successful investor's guide to spotting investment trends and turning points*. McGraw-Hill Professional, 2002.
55. Rapach, David E., Jack K. Strauss, and Guofu Zhou. "International stock return predictability: what is the role of the United States?." *The Journal of Finance* 68.4 (2013): 1633-1662.
56. Redmon, Joseph, et al. "You only look once: Unified, real-time object detection." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.

57. Ross, Sheldon M. Introduction to probability models. Academic press, 2014.
58. Scott, David W. "Scott's rule." Wiley Interdisciplinary Reviews: Computational Statistics 2.4 (2010): 497-502.
59. Scott, David W., and Stephan R. Sain. "Multidimensional density estimation." Handbook of statistics 24 (2005): 229-261.
60. Simon, H. A. (1955). A behavioral model of rational choice. The Quarterly Journal of Economics, 69, (1), 99-118
61. Shapiro, Linda, and George Stockman. "Texture." Computer vision. Prentice Hall, 2000. 235-248.
62. Smirnov, Nikolai Vasil'evich. "On the probabilities of large deviations." Mat. Sb 40 (1933): 443-454.
63. Sweeney, Richard J. "Some new filter rule tests: Methods and results." Journal of Financial and Quantitative Analysis 23.03 (1988): 285-300.
64. Trevor, Hastie, Tibshirani Robert, and Friedman JH. "The elements of statistical learning: data mining, inference, and prediction." (2009).
65. Treynor, Jack L., and Robert Ferguson. "In defense of technical analysis." The Journal of Finance 40.3 (1985): 757-773.
66. Turlach, B. A. (1993, January). Bandwidth selection in kernel density estimation: A review. In *CORE and Institut de Statistique*.
67. Van Dijk, M. (2011). Is size dead? A review of the size effect in equity returns, Journal of Banking and Finance, 35, (12), 3263-3274.
68. Vapnik, V. (2013). *The nature of statistical learning theory*. Springer science & business media.
69. Witten, Ian H., and Eibe Frank. "Data Mining: Practical machine learning tools and techniques 2nd edition." Morgan Kaufmann, San Francisco (2005).
70. Zhu, Yingzi, and Guofu Zhou. "Technical analysis: An asset allocation perspective on the use of moving averages." Journal of Financial Economics92.3 (2009): 519-544.

Vita

Matt Lutey is from Upper Michigan and completed his B.S. (pre-med) in finance and M.B.A. at Northern Michigan University in 2011 and 2013 respectively. He completed his M.S. in financial economics at the University of New Orleans in 2017 and his Ph.D. in 2019. He accepted a tenure track position in finance at Indiana University Northwest starting August 2019.