In Search Of Pockets Of Predictability

Alan Taxonera Morera

Master´s Thesis
Swiss Federal Institute of Technology of Zurich (ETH)
Escola Tècnica Superior d´Enginyeria Industrial de Barcelona (ETSEIB)

Supervisors:
Prof. Dr. Didier Sornette
Dr. Ryan Woodard

Zurich,
August 24, 2008

> "*Men are moved by two levers only: fear and self-interest*"
>
> Napoleon Bonaparte (1769–1821)

The little corporal might have oversimplified the incredibly complex set of emotions that govern and control human conduct. His statement, however, gives us an idea of the underlying forces that are at work on a normal trading day and offers us a hint on why we so often act irrationally. Throughout his career as a general, consul and later emperor, Napoleon himself was very aware of this reality. His superlative skills combined with a deep understanding of human nature might have been decisive factors for his success and grandeur.

Abstract

Spidyn is a technical indicator aimed at discovering pockets of predictability of the order of a few days in stock markets, based on the detection of positive or negative unsustainable price accelerations. Coupled with an appropriate execution strategy, the Spidyn indicator could be used for algorithmic trading purposes.

The goal of this Master´s Thesis is threefold. First, try to verify in a robust way the predictive power of Spidyn by capturing a signal that confirms an added value to the indicator. Second, by incorporating this knowledge into our trading strategies, conduct portfolio tests and compare the results obtained with previous ones. Last but not least, and almost equally meaningful, is to get a feel for finance and complex systems and try to understand some of the underlying mechanisms of the stock market. I wrote this paper in a linear and structured fashion, with the idea in mind that it would encourage and aid another master´s student or PhD in this pursuit.

Contents

List of Figures **v**

List of Tables **vi**

1 Introduction **1**
1.1 The Stock Market 1
1.2 The S&P500 3
1.3 Market Efficiency 4
1.4 Investing Principles 4
1.5 Technical Versus Fundamental Analysis 6
1.5.1 Comparison 6
1.5.2 Indicators 6
1.5.3 Where We Stand 6

2 The Spidyn **8**
2.1 Overview 8
2.1.1 A Word of Caution 8
2.2 Description of the Laboratory 9
2.3 Data 9
2.3.1 Time Series 9
2.3.2 Returns Versus Log-returns 9
2.3.3 Delta-hedged Prices 10
2.4 The Black Box 11
2.4.1 Purpose 11
2.4.2 Parameter Space 12
2.5 Strategy 13
2.5.1 Purpose 13
2.5.2 The Contrarian Strategy 13
2.5.3 Parameter Space 14
2.6 Previous Knowledge 15
2.7 Motivation 15

3 Distribution Tests **17**
3.1 Goals 17
3.2 Justification 17
3.2.1 Differences in Input Time Series 18
3.3 Normality Plots 18
3.4 The Anderson-Darling Test 25

3.5 Visual Inspection 25
3.6 Conclusions 27

4 Correlation Tests **28**
4.1 Linear Correlation 28
4.2 Negative Products 32
4.3 Subsets 32
4.4 Conclusion 33

5 The Signal **34**
5.1 Objective 34
5.2 Binning 34
5.2.1 Linear Binning 34
5.2.2 Quantile Binning 35
5.3 Probability of Success 36
5.4 Testing Against Randomness 36
5.4.1 Purpose 36
5.4.2 Procedure for Reshuffling Time Series 36
5.5 Results 37
5.6 Conclusions 40
5.6.1 Bullish Regime 40
5.6.2 Bearish Regime 41

6 Portfolio Tests **42**
6.1 Notions 42
6.1.1 Definitions 42
6.1.2 Markowitz 43
6.2 Quantile-based Thresholds 43
6.3 Description of Statistics 45
6.4 Conclusions 46
6.4.1 Bullish Regime 46
6.4.2 Bearish Regime 47
6.5 Where To Move Next 49

A Portfolio Test Results **51**

B How the Greatest Investors Tell Us To Invest **62**

Acknowledgments **65**

Bibliography **66**

List of Figures

1.1 Response of the market to the 1987 crash 2
1.2 The S&P500 acceleration over the last 50 years 3

2.1 Comparison between normal and delta-hedged prices 11
2.2 A time series of Spidyn indicators 13

3.1 Differences regarding the type of input data 19
3.2 Normality plot for T=10 21
3.3 Normality plot for T=30 22
3.4 Normality plot for T=60 23
3.5 Normality plot for T=150 24
3.6 Histogram of indicators of the *Microsoft* stock 26
3.7 Histogram of sampled distribution 26

4.1 Linear correlation for 500 companies (I) 29
4.2 Linear correlation for 500 companies (II) 30
4.3 Linear correlation for 500 companies (III) 31
4.4 Cross correlation of the *Google* stock 33

5.1 Probability of success using linear binning 35
5.2 Probability of success measured for a bullish market 38
5.3 Probability of success measured for a bearish market 39

6.1 Histogram of thresholds for different quantile values 44

List of Tables

3.1 Results of the Anderson-Darling test 25

A.1 Bullish market using quantile = 1% 52
A.2 Bullish market using quantile = 2% 53
A.3 Bullish market using quantile = 5% 54
A.4 Bullish market using quantile = 15% 55
A.5 Bullish market using tIn = tOut = -0.5 56
A.6 Bearish market using quantile = 1% 57
A.7 Bearish market using quantile = 2% 58
A.8 Bearish market using quantile = 5% 59
A.9 Bearish market using quantile = 15% 60
A.10 Bearish market using tIn = tOut = -0.5 61

Chapter 1

Introduction

1.1 The Stock Market

Due to a variety of reasons, the stock market has received ample attention throughout its relatively short history. On one level, the stock market is an unbounded object of fascination to many academics and traders alike who try to diagnose its dramatic crashes and anticipate its course. However, on a completely different level, its potential for lucrative gains attracts hordes of traders seeking greater wealth. Ultimately, the function of any stock market is to bring buyers and sellers together to enable the trading of company stock and derivatives at an agreed price. Throughout the literature and folklore, we frequently hear of stories of investors who made fortunes overnight or of others who lost millions of dollars and ended rather tragically. These anecdotes trigger our imagination and resonate well with the gambler in all of us.

The stock market is in many respects a self-organizing system constituted by millions of agents trading actively worldwide. In the framework of complex systems, financial markets are envisioned as systems whose large number of interacting constituents self-organize their internal structure and their dynamics with sometimes macroscopic or emergent properties. This perspective challenges the previous analytical approach, consisting of decomposing a system in smaller components and attempting to understand the whole as the sum of its parts [1].

The crash of 1987 caused multi-million dollar losses and sent shock waves across the financial world; from the NIKKEI in Tokyo to the FTSE in London; from the DJIA in New York to the DAX in Frankfurt. Figure 1.1 illustrates the impact it had on the S&P500 index and how the general public responded in times of panic and despair. Surprisingly enough, the market seemed to reach a consensus. The mechanism, with which the investing community agreed on a price, is reminiscent of a damped oscillator in physics [2]. This and other similar phenomena are an accurate mirror of the rich patterns of the stock market and of the sophisticated interplay between its interacting agents.

The market is extremely volatile. As a result from this, you can never be sure of what it will do tomorrow. Even J.P. Morgan when he was asked how the market would react, he coldly replied, “It will fluctuate”. Markets manifest whimsical symptoms as they shift through different regimes and moods in relatively short periods of time. One day the market will be unusually quiet,

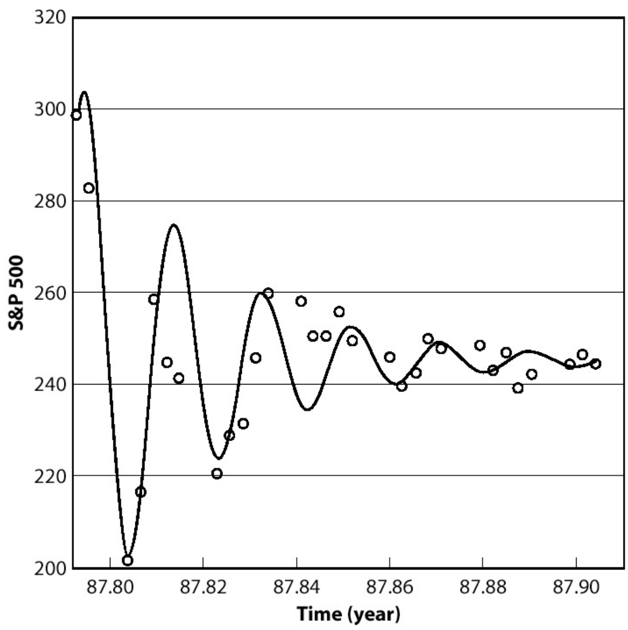

Figure 1.1: Response of the market to the 1987 crash

while others it will be incredibly turbulent with huge ups and downs. Sometimes it will exhibit a bullish attitude and suddenly for no apparent reason, switch over to a bearish regime. In today's world financial markets, every single trade and transaction can be stored and later analyzed. The scientific exploration of these records of social human interactions is a challenge where many fundamental questions still remain open and unresolved.

Stock market prices change at all time scales. If there is no characteristic scale in stock market price fluctuations, crashes are nothing but small drops that did not stop [3]. "According to this view, since crashes and bubbles belong to the same family as the rest of returns we observe on normal days, they should be unpredictable because their nucleation is not different from that of the multitude of small losses which cannot be predicted at all" [1]. Ever since Bachelier investigated the properties of prices at stock markets, it is known that price changes follow, to a first approximation, a random walk (or Brownian motion) behavior [4]. As an important consequence, you cannot simply predict what the future price of an asset will be through a statistical analysis of past prices. Nevertheless, as Bachelier remarks himself, in lack of deterministic predictability you can still rely on statistical predictability in financial markets. The most important prediction of the random walk model is that the square root of the amplitude of fluctuations increases in proportion to time [1].

1.2 The S&P500

On many occasions we will refer to the S&P500. The S&P500 is a stock market index created in 1957 containing the stocks of 500 large-cap corporations, most of which are American. A few of the companies that comprise the index are *Amazon, Coca Cola, Hewlett Packard, Ebay, Ford, Goldman Sachs, Intel, Microsoft, Nike, Wal-Mart*, etc. Widely regarded as the best single gauge of the U.S. equities market, the index can be used as building blocks in portfolio construction. All of the stocks of the S&P500 are traded on the two largest U.S. markets, the New York Stock Exchange (NYSE) and Nasdaq. It is ranked right after the Dow Jones Industrial Average (DJIA), in terms of most highly monitored indexes. Since it is market-value weighted, movements in price of companies that have bigger market capitalizations have a greater effect on the index.

One of the great puzzles of finance is the *equity risk premium*. This refers to the fact that stocks have historically outperformed bonds by a substantial amount. The average annual real return (inflation-adjusted return) on the U.S. stock market for the past century has been about 7.9%. In the same period, the real return on a relatively riskless bond was around 1.0%. The equity premium is the difference between these two returns, which is 6.9%. According to some academics, the difference is too large to reflect a proper level of compensation that would occur as a result of investor risk aversion [11].

Paul Samuelson, a famous economist who was awarded the Nobel Prize in Economics in 1970, has reported that even good investors on average seem to find it hard in the long run to do better than common-stock indexes [11]. Figure 1.2 shows the evolution of the S&P500 over the last fifty years. We can see that the trend is extremely clear: it has been steadily increasing. One must note that the Y-axis is plotted in logarithmic scale, so in fact the acceleration is exponential and not linear as it appears to the eye. After taking a glance at the following graph, the buy & hold strategy seems to work remarkably well. This raises the following question: can dynamic strategies outperform the acceleration of the market? As we will see later on, this topic is central to our study of the Spidyn indicator.

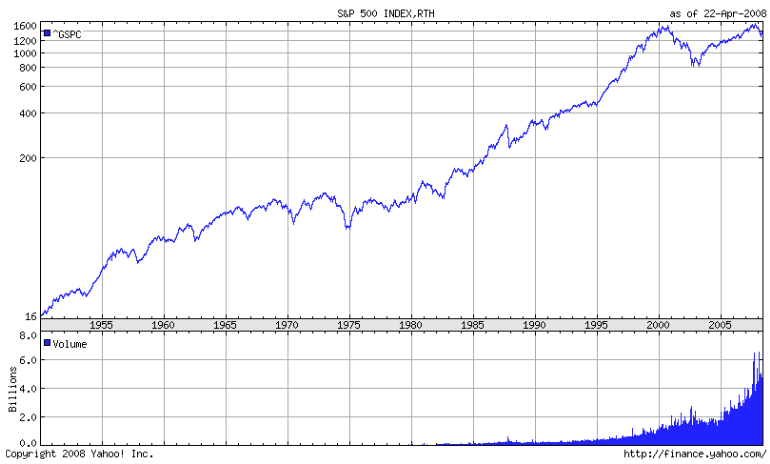

Figure 1.2: The S&P500 acceleration over the last 50 years

1.3 Market Efficiency

An issue that is the subject of intense debate among academics and financial professionals is the *Efficient Market Hypothesis*. To us, it is relevant because it conditions the existence of pockets of predictibilty. The Efficient Market Hypothesis states that at any given time, stock prices fully reflect all available information. The implications of it are truly profound. Most individuals buy and sell stocks under the assumption that the securities they are buying are worth more than the price that they are paying, while securities that they are selling are worth less than the selling price. But if markets are efficient and current prices fully reflect all information, then buying and selling securities in an attempt to outperform the market will effectively be a game of chance rather than skill. There are three forms of the efficient market hypothesis [7]:

1. *Weak-form Efficiency*- This view suggests that technical analysis techniques will not be able to consistently produce excess returns. The reason being that share prices exhibit no dependencies, which implies that future price movements are determined entirely by unexpected information and therefore are random.

2. *Semi-strong-form Efficiency*- Semi-strong-form efficiency implies that neither fundamental analysis nor technical analysis techniques will be able to produce excess returns. This occurs because share prices adjust to available new information very rapidly and in an unbiased fashion, such that no excess returns can be earned by trading on that information.

3. *Strong-form Efficiency*- This view implies that stock prices reflect all information, public and private, and no one can earn excess returns. In other words, even insider information is of no use. To test for strong-form efficiency, a market needs to exist where investors cannot consistently earn excess returns over a long period of time. However, given the hundreds of thousands of fund managers worldwide, even a normal distribution of returns should be expected to produce a few dozen excellent performers.

We assume that the market is close to being efficient and will shortly stabilize and reflect more accurately a more objective value of its assets. As we have seen, this theory asserts that prices of stocks, bonds or any other asset which is traded, reflect all known information in the market. Hence, it is impossible to consistently outperform the market by using any information that the market already knows. Paradoxically, the more intelligent and hard working the investors are, the more random the sequence of price time series generated by the market becomes. So the most efficient market, is one in which price changes are completely random and unpredictable. The market has been reported to be close to efficient by a series of scholars [15]. This point is important, as it hints there is room for pockets of predictability to exist.

1.4 Investing Principles

What is an investment really, but a bet in which one tries to guess how others will value an asset? There is an endless assortment of definitions out there of what an investment is, but maybe the only point they all share in common is that their goal is to increase our wealth. The fact of the matter is that all of us out there are trying to make a living, and parallel to that, most of us are also trying to get rich, to improve our standard of living or simply to sustain a lifestyle with which we feel comfortable. There is absolutely no question these are the top reasons why investors think in the first place of entering the stock market.

How would an investor behave if before a trading day he had complete information of the closing prices for each stock? It is reasonable to think he would invest all of his wealth in the single best performing stock of the market. In fact, if the information is reliable enough, he should. Now let us think of how he would act if he had access to only some information. Let us suppose, the direction of the change of stocks is known to him but not the magnitude. The most logical thing to do would be to split his wealth among those stocks that will increase in value, since there is no reason to invest more money on one particular stock than on another. This strategy makes sense because although it is likely it is not optimal, the investor knows he will make money out of the deal. But how would he invest his beloved wealth if he had absolutely no information about the future? What if instead he was showered everyday with turbulent news, unfounded rumors and conflicting opinions? This is, we are afraid, the situation for most investors in the stock market, however sophisticated, influential or wealthy they are.

If you are a stakeholder of a stock, it is almost certain you would wish the price of it to skyrocket. When this occurs, it is generally because of one of the following reasons. Very often a boom is explained as being a result of an overvaluation as seen from the eyes of nervous investors, leading to inflated expectations of the performance of stocks. This in turn results in a *positive feedback*, which is a mechanism of amplification found to exist in many systems. Sometimes the run-up is ignited by a more plausible reason, such as a real improvement of the company held by investors. Let us imagine *Apple* announces the release of yet another revolutionary product or *Porsche* reports generous dividends among its shareholders or *Toyota* patents a new technology for hybrid cars. These are all good reasons that provide the necessary ground for stock prices to explode. What if Mr. Warren Buffet spends millions of dollars on *Exxon* or an article on the Internet claims the *Google* run-up will last another year. These, we suspect, are not so good reasons.

Investments in the stock market are based on a straightforward rule. If we expect the price to go up, we should buy and hold the stock until the upward trend changes direction (or in other words, we should go “long”). In such a situation, the investor should close his or her position and sell the stock at a higher price, thus cashing in the difference. A person will profit from a long position in the market if the price of the stock increases. Nevertheless, if we expect the market to go down, we should stay out of it or sell by borrowing a stock only to rebuy it later for a smaller price (commonly referred to as being “short”). There are two very interesting ideas behind the concept of selling short. In the first place, it allows the common trader to deal with amounts of money that he might not necessarily own at the moment, but that at the same time are necessary to participate in a promising activity. If the operation works out, he or she will later be able to pay back the amount due, and will retain the difference in price. Secondly and contrary to what many people think, during a financial collapse when all stocks are plummeting and alarms are flying thick and fast, a great opportunity arises to make money.

In the end, it is all about timing. You can make money anytime, as long as you anticipate the market´s movements and act accordingly. Nevertheless, one must not neglect the effects of liquidity. The presence of liquidity constraints might prevent us from being able to sell a certain stock at a desired instant. For strategies that operate with very small margins, this nuisance might be relevant since it can seriously downgrade its performance. Liquidity has been gaining some attention lately and is believed by some academics to be a main driver of the prices of stocks.

1.5 Technical Versus Fundamental Analysis

1.5.1 Comparison

Two main approaches coexist in valuating stocks: fundamental analysis and technical analysis. On the one hand, fundamental analysis focuses on the value of a company, while ignoring the market. Usually it involves analyzing its management, income statement, intangible assets and competitive advantages. On the other hand, technical analysis concentrates on identifying trends by means of computer models and algorithms. Price charts help traders identify market trends, while technical indicators help them judge a trend's strength and sustainability. A distinguished British economist named Keynes argued that not only are stock prices determined by the firm´s fundamental value, but in addition, mass psychology and investor expectations play a major role [12]. He even went as far as to say that professional money managers prefer to devote their energy not to estimating fundamental values but rather, to analyzing how the crowd of investors is likely to behave in the future. Technical analysis is oblivious to the actual value of a company and alternatively tries to foresee the changes in valuation by the general public. Both fundamental and technical analysis offer us a distinctive and complementary view of companies. Therefore, both are worthy of our consideration in any study which aims to detect superior stocks. Except for the very last section in the appendix, titled "How the Masters Tell Us to Invest", this paper is written under the lens of technical analysis.

1.5.2 Indicators

A technical indicator is a series of data points that is derived by applying a formula to the price time series of a stock. In general, they serve three broad functions: to alert, to confirm and to predict. Regardless of the complexity of the formula, technical indicators can provide a unique perspective on the strength and direction of the underlying price action. To enable analysis, a series of data points over a period of time is required to create valid references. As this series is generated, buy and sell signals are issued. Ultimately, the aim of technical indicators is to identify trends in the stock market and aid a portfolio selection strategy.

1.5.3 Where We Stand

Currently, many efforts are being put forward in a variety of fields to improve our understanding of the stock market and help us devise a trading strategy which is both profitable and robust. The field of research is extremely active, and counts with a huge number of teams of researchers and companies devoting substantial amounts of energy and working around the clock in hopes of "beating the market". Due to obvious reasons, it is also one of the best funded. This climate of euphoria and exaltation has created countless indicators over the last years. Some technical analysis software programs come with dozens of indicators built in, and even allow users to create their own [8]. The majority of these indicators lack the predictive power required for a real-world implementation and eventually fall into obscurity. Others such as Commodity Channel Index (CCI), Relative Strength Index (RSI), Rate of change (ROC) or Moving Average Convergence Divergence (MACD) become popular in financial circles [9]. Among a vast ocean of indicators that try to offer a leap forward, is the Spidyn.

Chapter 2

The Spidyn

2.1 Overview

Spidyn is a technical indicator aimed at detecting windows of predictability within a short frame of time, typically of the order of a few days. The indicator reflects whether recent prices tend to accelerate in a non-usual way. We believe a large negative value of the indicator indicates a recent mini-crash, while a large positive value indicates a recent mini-bubble. It is unclear, however, as to the precise location of the transitional region between the two behaviors.

The stock market experiences very small but traceable fluctuations which have a potential for gains which cannot be overlooked. While it is possible that for sustained earnings of considerable amplitude one has to rely on long-term trends, it must be said that short-term strategies can also be quite lucrative. The indicator operates on a time scale that lies between that of intra-day trading (where trading is conducted with a precision down to tenths of a second) and long-term trading (which ranges from years to even decades).

"Independence between successive returns is remarkably well verified most of the time. However, it may be that large drops may not be independent. In other words, there may be occasionally bursts of dependence or pockets of predictability present in the market" [1]. In this context, a contrarian strategy based on the Spidyn indicator could be devised to exploit these rare moments of collective and some might add irrational behavior. One of our tasks will consist in trying to determine if such opportunities exist.

2.1.1 A Word of Caution

Throughout the entire text, one has to bear in mind that the core code of Spidyn is as of August 2008 of limited access and that neither me nor my supervisor Ryan Woodard had access to it. This piece of information has important implications that one cannot simply neglect. Sometimes it has been hard to interpret or comprehend some of the results without any more additional knowledge as to how the black box really calculates the indicator. Regardless of this fact, we have tried not to let this setback get in the way of our research and in the end we believe some interesting results have been attained.

2.2 Description of the Laboratory

The algorithm itself is provided in a black-box, the code of which was developed by Didier Sornette and Didier Darcet, and implemented by Yann Ageon in 2004. The main files used to generate the Spidyns are:

- `SpiDynTest.cpp` – contains the main function.
- `SpiDyn.h` – header file to recognize the Index function.
- `SpidDyn.a` – a static library.

One of the main difficulties of setting up the infrastructure required for each experiment is that a number of completely unrelated systems should work shoulder to shoulder to obtain results. To begin with, code written in the Python language by Ryan Woodard retrieves financial data from the Internet. Then, it interacts with the black-box compiled in C++ under Linux to produce the index. The parameters of each particular experiment are specified on a separate text file, which is linked to the Python program. Subsequently, Matlab® code developed by Gilles Daniel generates data structures with which the program can work. The `P.mat` structure contains all the information concerning the price time series, whereas the `I.mat` has recorded all the data regarding the indicator. Finally, the Matlab® code computes a series of statistics, graphs and reports to evaluate the performance of our portfolio trading strategy.

2.3 Data

2.3.1 Time Series

The Spidyn indicator is built on the daily closing price of stocks. We work with historical financial time series retrieved from the *Yahoo! Finance* website. Throughout our study, we have focused on the S&P500, although it could have been any other market. We have conducted our experiments mainly in two periods: 2000-2002 and 2003-2006, a bearish and bullish period respectively. Prior to 2001, we find there are important gaps in our data, which is why we have tried to avoid those years whenever it was possible. Had it not have been the case, we would have conducted experiments on a broader period of time. For future reference, we recommend the use of another database. Given the recent spike of volatility of stock markets due to the liquidity crunch that originated in the U.S. subprime market, it would also be beneficial to run our analysis on the most recent data.

2.3.2 Returns Versus Log-returns

For sake of simplicity, we will sometimes refer to log-returns simply as returns. Both are the two most widely used definitions of returns, although there seems to be some confusion when it comes to their usage. Consider the expression of returns from time $t-1$ to t:

$$R_{dt}(t) = \frac{P(t) - P(t-dt)}{P(t-dt)} \tag{2.1}$$

By definition, a log-return is:

$$LR_{dt}(t) = log\left(\frac{P(t)}{P(t-dt)}\right) \tag{2.2}$$

where dt is the time scale over which the return is calculated and can be for example one day, one week, one month, one minute or any other value. For a small x, we can write the following first order approximation:

$$log(1+x) \simeq x \tag{2.3}$$

In the same way, for small price variations we have:

$$log\left(\frac{P(t)}{P(t-dt)}\right) = log\left(1+\frac{P(t)}{P(t-dt)}-1\right) \simeq \frac{P(t)}{P(t-dt)} - 1 = R_{dt}(t) \tag{2.4}$$

Equation 2.4 shows that returns are a first order approximation of log-returns. Choosing returns or log-returns is not innocent. Opting for returns means that the money generated is not compounded. In other words, at $(t-dt)$, our wealth is $P(t-dt)$ and it remains so until time t. Choosing log-returns implies that at each dt, the return expressed in equation 2.1 is reinvested and accrues interest. Therefore, from a portfolio analysis view point, choosing log-returns means that the investor views his portfolio wealth with a continuous time view, for instance, he can withdraw his money out any time he pleases. Adopting returns suggests that the investor has fixed (say monthly or quarterly) times in which he evaluates his returns. Log-returns will always be numerically smaller than returns expressed as a percentage.

2.3.3 Delta-hedged Prices

Input time series may consist of regular prices or of what is known as delta-hedged prices. Following a suggestion from two collaborators of Gilles Daniel, he started testing his strategies not directly against stock prices, but against delta-hedged prices. Before any analysis, he rebuilt the time series of stock prices to hedge them against the S&P500 index. This is equivalent to shorting the index for $1 every time we open a long position of $1 in our trading strategy. The rationale behind working with a delta-hedged or delta-neutral dataset is that we want to spot the price time series that accelerates in an unsustainable way in comparison to the rest of the market. For instance, if the market as a whole keeps going up, then what we call "unsustainable acceleration" for a given stock should be revised [13]. Figure 2.1 compares the evolution of genuine and delta-hedged prices, for the *Ebay* stock in the period between 2003 and 2006. The exact procedure we adopted to build delta-hedged prices is the following [14]:

1. Compute the log-returns of asset i.

$$R_i(t) = log\left(\frac{P_i(t)}{P_i(t-1)}\right) \tag{2.5}$$

2. Compute the log-returns of the index j.

$$R_j(t) = log\left(\frac{P_j(t)}{P_j(t-1)}\right) \tag{2.6}$$

3. Reconstruct the log-returns of market-adjusted stock.

$$R_k(t) = R_i(t) - R_j(t) \tag{2.7}$$

4. Use these log-returns to build the delta-hedged stock prices.

$$P_k(t) = P_k(0)e^{\sum_{k=1}^{t} R_k(t)} \tag{2.8}$$

5. Run the Spidyn DLL on P_k.

6. At each entry signal, buy the delta-hedged stock P_k, which is equivalent to buying P_i and selling short P_j. At each exit signal, sell P_k.

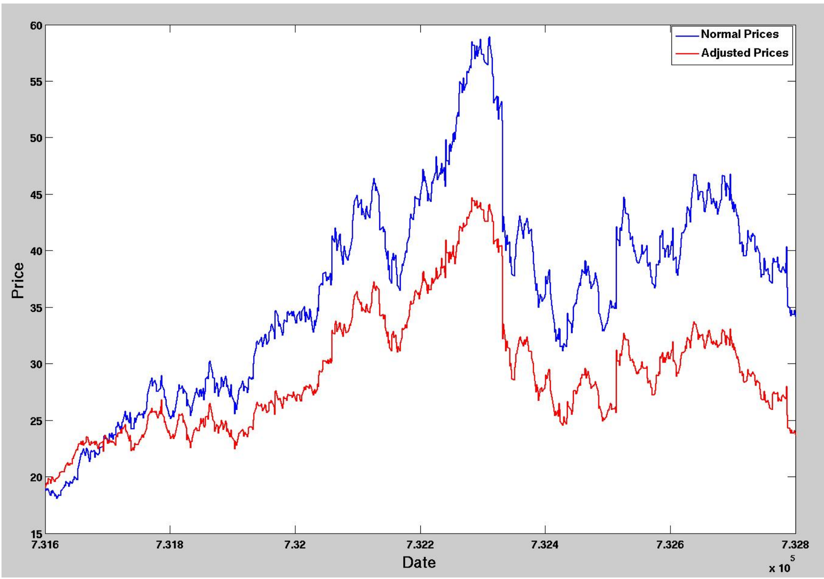

Figure 2.1: Comparison between normal and delta-hedged prices for the *Ebay* stock from 2003-2006.

2.4 The Black Box

2.4.1 Purpose

We were able to generate the indicators by means of the so-called black box. The black box takes in a number of input parameters and then produces a single number usually of the order of ± 1. The function `Index` is at the heart of the trading platform, as it is responsible for calculating the Spidyn indicators. Its header is as follows:

```
   double Index(double pdData[ ], long lDataNumber, int iBeginDegree,
int iEndDegree, double pdWeight[ ], bool& bSuccess)
```

2.4.2 Parameter Space

At the time of the writing of this paper, the source code of the `Index` function is kept secret. Nonetheless, we have access to some vital information, in particular to its input parameters and to a brief description of their meaning.

- `lDataNumber` – is equivalent to T, which is the size of the sliding window to compute the indicator on.
- `pdData` – is the price time series, the length of which has to be equal to T.
- `iBeginDegree` – is the minimum coefficient of the polynomial to fit. In the code, *degree* is equal to the number of degrees of freedom the polynomial has and not the highest exponent. For example, a polynomial of degree 3 would be $P(x) = a_o + a_1x + a_2x^2$ and not $P(x) = a_o + a_1x + a_2x^2 + a_3x^3$. It remained equal to 2 throughout the entire study.
- `iEndDegree` – is the maximum coefficient of the polynomial to fit. In the code, *degree* is equal to the number of degrees of freedom the polynomial has and not the highest exponent. It remained equal to 5 throughout the entire study.
- `pdWeight` – is a vector which gives the weight of each polynomial. Currently it allows to calibrate the relative weights given to the velocity, acceleration and jolt of a price time series. They are set to $\frac{1}{3}$, $\frac{1}{3}$ and $\frac{1}{3}$ respectively.
- `bSuccess` – is a non-specified boolean.

A time series of indicators is nothing but the result of several iterations of this process, each time with a moving window that slides along the price time series. If we take N as the length of the price series and T the window size on which the indicator is computed on, the number of iterations to perform is exactly *(N-T+1)*. Strictly speaking, the first *(T-1)* values of the Spidyn time series are composed of zeros, for it cannot operate with anything less than the first T prices. Given the fact that on average indicators are slightly positive, this suggests that times where the market is accelerating dominate over the times it is crashing. On the other hand, the occurrence that usually the most negative indicator is more extreme than its positive counterpart, might indicate that crashes are easier to identify than bubbles. To our knowledge, the Spidyn algorithm involves to some degree the fitting of polynomials and high order derivatives. Qualitatively, the index is a measure of the projected return outside the window divided by the daily variability observed within it [5].

Figure 2.2 portrays how the index describes an erratic movement around the central value of zero. As can be seen, at times the indicator surpasses the [-1,1] range. These values do not hold any other special meaning besides being a reference. Note also the initial flat section in which the indicator is not computed until we reach a price time series of at least length T.

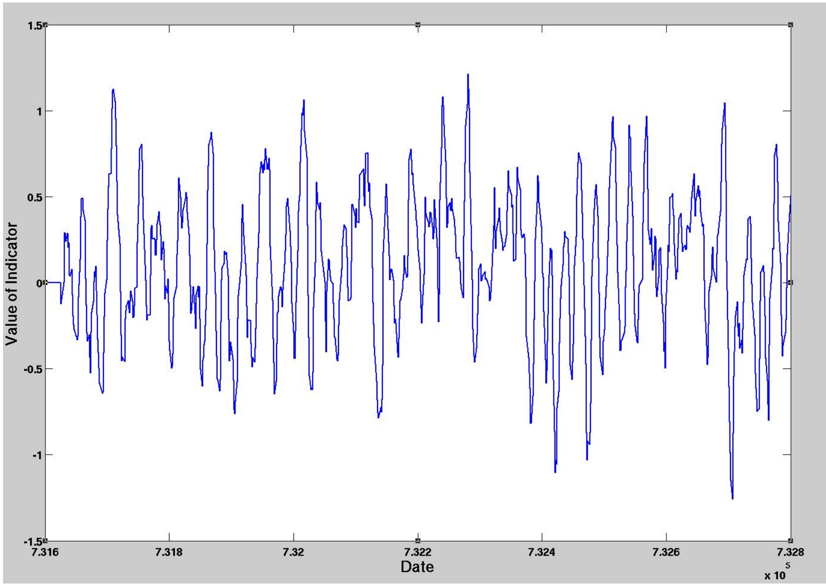

Figure 2.2: A time series of Spidyn indicators of the *Amazon* stock from 2003-2006 using normal prices.

2.5 Strategy

2.5.1 Purpose

The role of the strategy is almost as important as the indicator itself. After all, Spidyn is just a number. If we want to manage our portfolio in a real world scenario, we need an automated system that will interpret the indicator and act according to an organized set of instructions. In portfolio management, there are a number of decisions that have to be made on a daily basis. When do we open a position? How much do we invest on a specific asset? When do we close that position? Generally speaking, the more sophisticated our strategy is, the more parameters we have to calibrate.

2.5.2 The Contrarian Strategy

A *trend reversal* is defined as a sudden change in the price direction of a stock, index, commodity, or derivative security. A reversal can be a positive or negative change against the prevailing trend. Technical analysts watch for these patterns because they can indicate the need for a different trading strategy on the same security. For example, if a technical analyst holds a stock and notices a reversal pattern, he may want to consider closing his existing long position and assuming a short position to profit from the downward movement of the stock's price.

Why do such sudden changes occur? Many believe that something definable as a "market psychology" exists, and that sufficiently large herd effects can cause bubbles and crashes. Some traders and financial writers even see the market itself as possessing its own moods and personality, sometimes describing the market as "bullish" or "bearish". Moreover, it may very well be that nonlinear and non sustainable regimes occur on shorter time scales. Are large and well-defined market regimes nothing but small ones that did not stop or do they belong to another class?

By means of technical analysis, such non sustainable market regimes could be tracked and then exploited according to a mathematical methodology commonly used for stock investing known as *mean reversion*. Mean reversion is a theory suggesting that prices and returns eventually move back towards the mean or average [10]. When the current market price is less than the average price, the stock is considered attractive for purchase, with the expectation that the price will rise. When the current market price is above the average price, the market price is expected to fall. In other words, deviations from the average price are expected to revert to the average. This mean or average can be the historical average of the price or any other relevant average such as the growth in the economy or the average return of an industry. This theory has led to many investing strategies involving the purchase or sale of stocks whose recent performance has greatly differed from their historical averages. The Spidyn indicator is based on such a mean reversal strategy for stock trading.

Spidyn detects increases or decreases in the price of stocks that are unsustainable, reflecting an overestimation/underestimation due to an optimistic/pessimistic atmosphere. As its name suggests, the contrarian strategy is based on swimming upstream. A large positive amplitude indicates the occurrence of a mini-bubble. Instead of buying like the majority of investors, we do the opposite and sell. The reasoning behind our actions is that we expect prices are soon going to drop because they are unsustainable. In our study we have not allowed short selling, therefore, we can only sell a stock that we have already acquired. Similarly, large negative amplitudes suggest a mini-crash. Rather than joining the crowd of investors by selling, we go out and buy even more of that stock. The logic is that we expect prices are soon going to rise because again, they are unsustainable.

2.5.3 Parameter Space

It must be stressed that the landscape of parameters we are dealing with is extraordinary. In fact, this is one of the greatest challenges in our research. If we do not internalize this fact, we run the risk of not seeing the forest for the trees. We have already covered the variables closest to the black box. It is now about time we define and explain the parameter space that encompasses our trading strategy. The most important variables include:

- `InvestFraction`- the fraction of money invested when a new stock is incorporated in the portfolio. The percentage is drawn from our total wealth, which may exist in the form of cash or money invested in the stock market.

- `tIn`- we enter a position when $i(t) < tIn < 0$, where i(t) is the value of the indicator. It has to be surpassed in a downward direction. As a result, we enter the market before an expected rebound in the price occurs.

- `tOut`- we exit a position when $i(t) > tOut \geq tIn$. It has to be surpassed in an upward direction. Setting `tIn` and `tOut` is critical to the performance of our strategy and therefore a topic of major concern in our research.

- `MaxLeverage`- the maximum leverage allowed. It is calculated as the money invested on the stock market over out total wealth.
- `Stop-Loss orders`- These were introduced to protect ourselves from important *drawdowns* by exiting the market. A drawdown is defined as a persistent decrease in the price over consecutive days. In our strategies, when our current wealth has lost X percent in Y days we automatically cut all our positions and wait for another Z days before entering the market.

2.6 Previous Knowledge

The preceding researcher in charge of the study of Spidyn, Gilles Daniel, did a splendid job of building a Matlab® environment for algorithmic trading revolving around the indicator. He also implemented various trading strategies based on the Spidyn indicator and back-tested them extensively on past and surrogate data. To put it in a nutshell, his analysis lead him to the following conclusions [13]:

1. When larger than a given threshold, the Spidyn indicator demonstrates some predictive power on stock price returns.
2. It is possible to design simple trading strategies based on this indicator which robustly exhibit Sharpe ratios slightly larger than one.
3. The signal-to-noise ratio identified by the Spidyn indicator does not appear high enough yet to move directly to a real-world implementation of algorithmic trading strategies.

Another share of knowledge of the index was acquired through a series of meetings with Didier Sornette. One of the notions he transmitted to us was that most of the time Spidyn is not revealing relevant information. The signal, if any, we assume to be very weak. As a consequence, for the signal to emerge we must apply some kind of statistical averaging. He also provided some insight and theoretical background as to the way the indicators are generated and the meaning of its parameters. More importantly, his feedback on results and view on the subject helped optimize our efforts as we moved forward.

2.7 Motivation

Now that the stage for Spidyn and all its components has been set up, it is about time we discussed our aspirations and scope of this project. Like in virtually any study related to the field of complex systems, it is not easy to achieve a major step forward in a conclusive manner. We strive to uncover and contribute with something new and useful, but at the same time we have to be realistic with our own inherent limitations of time and knowledge in the field. In the early stages of our research, we were confronted with the following question, “In which direction should we set out to pursue our goal?”

Upon reading the reports Gilles Daniel had generated, one thing struck us as odd. The thresholds that delimited buying and selling orders, (`tIn` and `tOut` respectively) were equal to a constant. They were not dependent on other factors which could play a key role such as time, company, market mood, etc. How could possibly a fixed value, effectively bound a buying and selling opportunity for all companies throughout the shifting market regimes? It suddenly came to mind that we could attempt to convert a static parameter into a dynamic parameter, one that could possibly capture that the market was alive and breathing. By dynamic, we understand that the thresholds that define the trading strategy are time and company dependent. The first conception we had of these "smart thresholds" differs slightly from that of the final implementation, which makes use of a supposedly strength of the indicator and incorporates the idea of quantiles to make companies comparable. In later chapters, we will explain in detail all of these concepts and the sequence of actions that lead to their development.

Chapter 3

Distribution Tests

3.1 Goals

We initiated on the study of the distribution of the indicator and its statistical properties as a solid starting point in our research. Some of the topics addressed are:

- Do the indicators follow a normal distribution?
- Can we identify outliers within indicators belonging to one stock?
- Are indicators drawn from original and delta-hedged time series identical?
- Are distributions from different companies statistically equivalent or do they differ?
- Can we improve our understanding of the indicator after exploring a fraction of its parameter space?

3.2 Justification

The assumption of normality is extremely significant because in many statistical procedures we take for granted that variables come from a normal distribution. One of the traditional ways of testing whether a sample is drawn from a normal distribution or not consists in using normal probability plots. Besides normality plots, another powerful tool for verifying if a distribution approximates that of a Gaussian, is the Anderson-Darling test.

Notwithstanding the fact that Spidyn belongs to a normal distribution or any other, studying its distribution is relevant in identifying anomalous data. This is a key issue in validating a mathematical model. A problem often encountered in statistics is that you are never 100% sure if strange observations suggest the model does not fit well or if it is due to randomness. To illustrate this, let us look at an example. Our best friend calls to tell us that he has found a very special nickel on the floor. After having a strong urge to flip the coin ten times in a row, he always seems to get heads. Well, there is nothing wrong in flipping a coin ten times and getting ten heads. It is unlikely, but stranger things have happened. Similarly, the question here would be, are the ten heads a result of using a biased coin or can they be entirely attributed to luck? We should probably start off by either submitting the coin to examination or repeating the experiment a number of times (with us on the scene this time. . .).

Another issue of great importance is confirming whether indicators belonging to different companies are drawn from the same distribution or not. In other words, are distributions of indicators statistically identical? If that is true, with what confidence level can this be assured? Prior to this verification, a common entering and exiting threshold has been used for all companies. This has important implications in our strategy, for it treats all stocks equally. Is it likely that, given a fixed threshold, we obtain the same proportions of data beyond it, regardless of the stock we pick?

Without extensive knowledge as to how the indicator is generated, it is extremely difficult to foresee how tweaking some of the parameters will affect Spidyn. We feel less inclined to adjust the degrees and weights of the polynomials, as doing so would require a substantial level of understanding of how the Spidyn functions internally. Although one can always speculate beforehand, how will the indicator respond if the window size is increased? Does choosing between normal and delta-hedged prices have an impact distribution-wise? Under what conditions does Spidyn approximate best to a normal distribution?

3.2.1 Differences in Input Time Series

We have seen a comparison between adjusted prices and original ones in figure 2.1. Apparently, the shape of both time series is very alike and only a certain multiplicative factor seperates one from the other. Now, we are in particular even more committed in quantifying how much they alter the outgoing indicators. Is opting for one type of input prices or another relevant to Spidyn?

Figure 3.1 portrays the absolute value of the difference between Spidyns generated from ordinary and market-adjusted prices of *IBM* for the period 2003-2006. Bearing in mind that the Spidyn algorithm outputs a value of the order of ± 1, the differences cannot be neglected. As a consequence, choosing the type of input data is not innocent and has a profound effect on the value, but not the behavior of the prices. In 80%–90% of the experiments conducted during our research, we fed normal prices to the black-box. The reason being that we wanted to keep alterations and variables to a minimum, so that results were as robust as possible.

3.3 Normality Plots

A normality plot is a graphical technique for assessing whether or not a sample of data is normally distributed. If all the data points fall near the theoretical normal distribution line, an assumption of normality is reasonable. Otherwise, if the points curve away from it, then the assumption is not justified. The window size usually used in prior tests is of the magnitude of 30 or 60 days. A good share of understanding can be inferred from the plots obtained in the following experiments. To carry them out, we randomly picked four stocks of the S&P 500 and monitored them as we made adjustments to the parameter T. Starting from the upper-left hand corner, the stocks are: *Noble Corp., Nisource Inc., National Oilwell Varco and Novell Inc.*

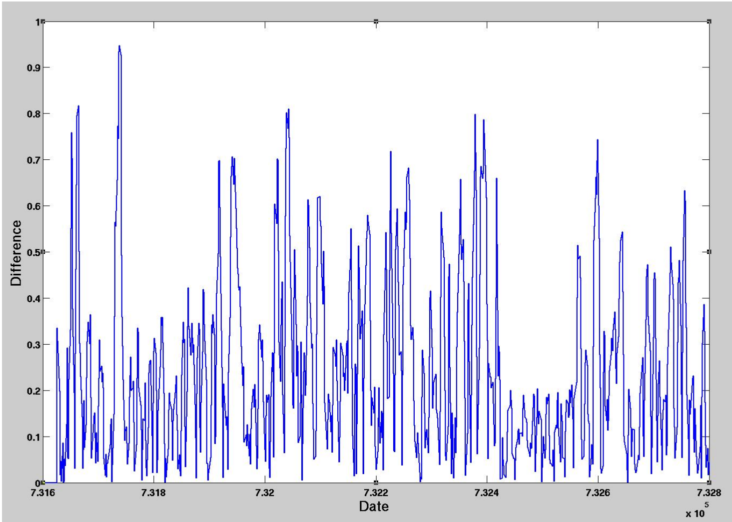

Figure 3.1: Differences in indicators regarding the usage of normal or delta-hedged prices of IBM for 2003-2006.

With the intention of making graphs comparable, all normality plots dealing with the same window size, have been scaled down to the one of least resolution. The symmetry of the graphs has been maintained for visualization purposes. Note how the amplitude of the indicators changes, as a consequence of regulating T. This effect will be discussed in a future section dealing with the conclusions of the distribution tests.

Figure 3.2 represents the normality plots for tiny values of the window size. For $T = 10$ days, we learn that Spidyn acts in some respects like noise. Under these circumstances, the indicator fits non-surprisingly quite well with a normal distribution. We should not miss the fact, that for some stocks, signs of fat tails still remain visible.

Figure 3.3 shows the normality plots for $T = 30$. For central values, the hypothesis that they are distributed according to a normal distribution seems reasonable. In the tails, however, one can appreciate in the case of some stocks a considerable deviation from normality. On the lower right corner, the plot for this company externalizes an obvious departure from normality. This already starts to point out that different companies express distinctive dynamics. We should also remark that the most negative indicator is practically always numerically more distant from zero than the most extreme positive value. Is it reasonable to suggest that Spidyn is best at predicting rebounds when the price is plummeting than bursts of bubbles?

As T increases to 60 days, this phenomenon is even more palpable. In figure 3.4, we see how *fat tails* begin to emerge. Fat tails are said to exist when extreme events occur with a much higher probability than they would if assuming a normal distribution. In finance, fat tails are generally considered undesirable because of the additional risk they imply. This is good news for investment targets and for the study of Spidyn, since it means that a few rare observations might carry hidden information that could be advantageous for identifying pockets of predictability. In fact, it would be discouraging to observe that Spidyn indicators adjust themselves perfectly with a normal distribution, as this would signal investing opportunities are improbable.

By the time T reaches 150, Spidyn´s distribution is deformed beyond recognition as seen in figure 3.5. We identify peaks building up around zero, suggesting a significant amount of indicators lie in the vicinity of this region. When we integrate an infinitesimal value beyond zero, we drastically increase the cumulative probability function. If we push it to the limit, the Gaussian model cannot adequately describe the distribution that Spidyn seems to follow.

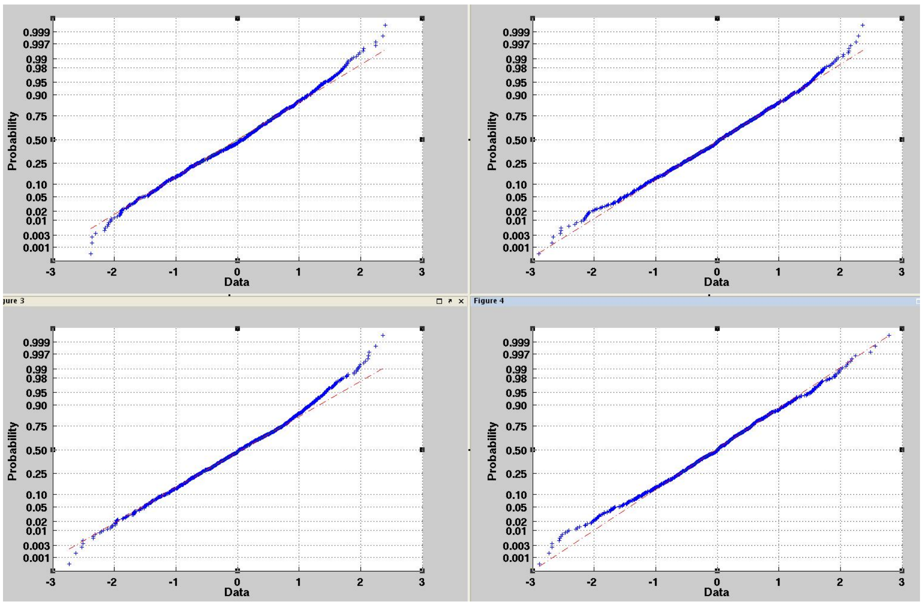

Figure 3.2: Normality plot for T=10 days during 2003-20006 using normal prices. Starting from the upper-left hand corner, the stocks are: *Noble Corp., Nisource Inc., National Oilwell Varco and Novell Inc.*

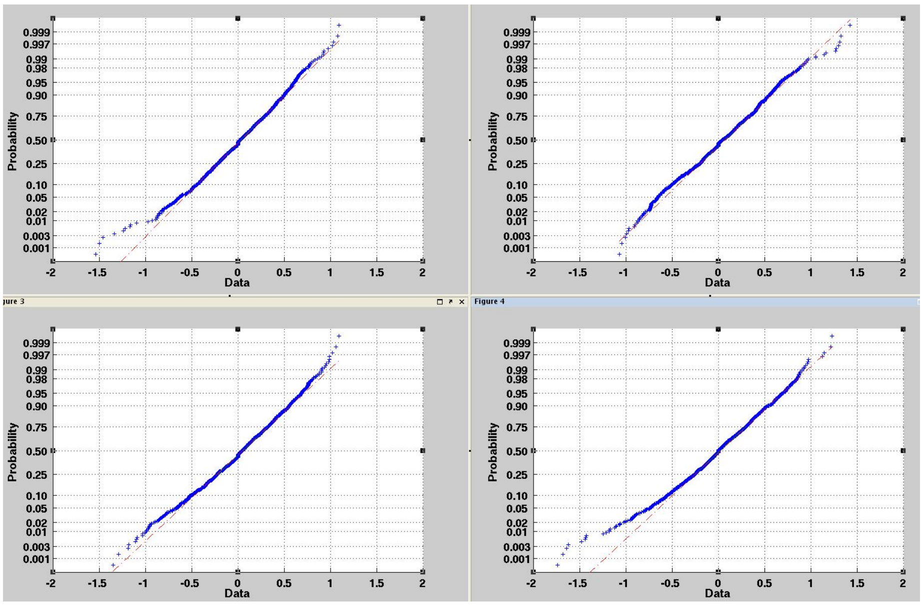

Figure 3.3: Normality plot for T=30 days during 2003-20006 using normal prices. Starting from the upper-left hand corner, the stocks are: *Noble Corp., Nisource Inc., National Oilwell Varco and Novell Inc.*

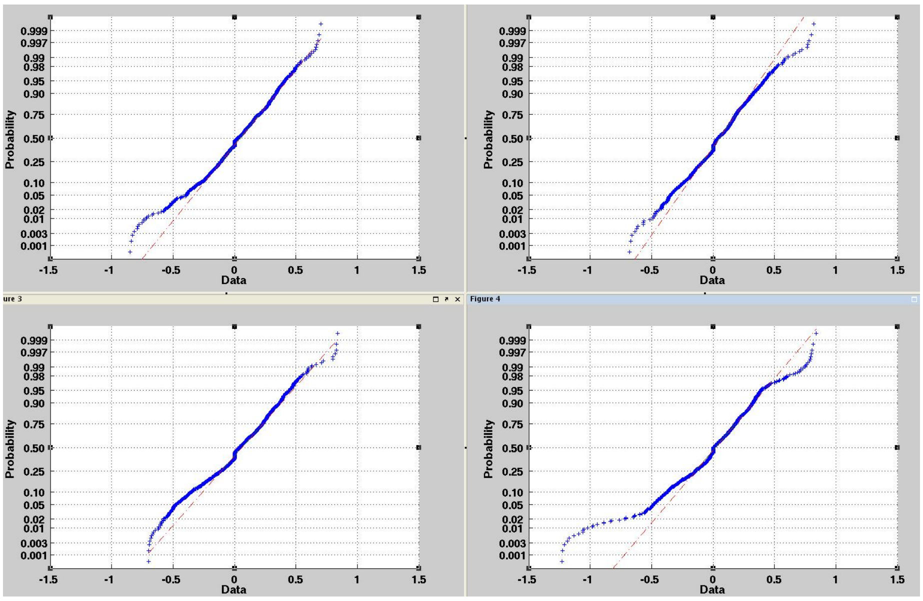

Figure 3.4: Normality plot for T=60 days during 2003-20006 using normal prices. Starting from the upper-left hand corner, the stocks are: *Noble Corp., Nisource Inc., National Oilwell Varco and Novell Inc.*

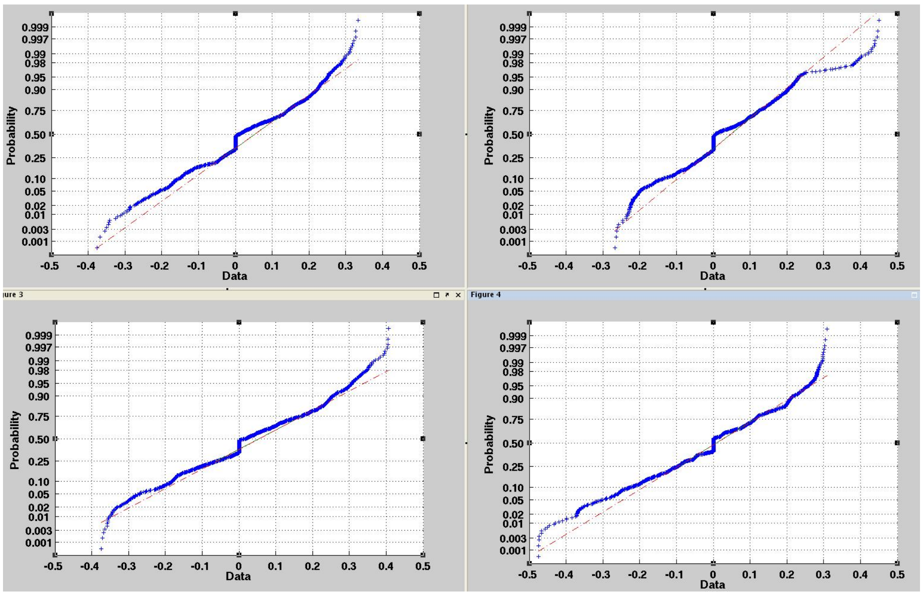

Figure 3.5: Normality plot for T=150 days during 2003-20006 using normal prices. Starting from the upper-left hand corner, the stocks are: *Noble Corp., Nisource Inc., National Oilwell Varco and Novell Inc.*

3.4 The Anderson-Darling Test

To ensure the reliability of the results obtained from the normality plots, we conduct an extra hypothesis test known as the *Anderson-Darling* test. The Anderson-Darling test is a powerful procedure to detect departures from normality. It makes use of the specific distribution by calculating critical values that are employed in assessing normality of a sample. This means Anderson-Darling is not distribution free, in the sense that the critical values calculated depend on the distribution being tested. One of its distinctive features is that it is highly sensible to small deviations in the tails of the distribution, meaning that a low number of unusual observations is sufficient to reject the null hypothesis which states that Spidyn is normally distributed. The opposite occurs with other tests, such as the *Kolmogorov-Smirnov*, which in contrast, place more weight on divergences found in the central area of the Gaussian bell. We analyzed for what percentage of stocks the normality assumption is justified. For a given significance of 95%, we obtained the following results:

Window size	Probability Associated
10	71%
30	18%
60	0.6%
150	∼0%

Table 3.1: Results of the Anderson-Darling test

3.5 Visual Inspection

Until now, we have conveyed the idea that Spidyns do not fit well to a Gaussian distribution. To satisfy our curiosity, we proceed to another test. First, we will first obtain both the mean and standard deviation from a distribution of indicators. Next, we will draw a normal distribution with the same μ, σ and number of observations. Finally, we will compare both plots and judge for ourselves how strongly they resemble each other.

Fig 3.6 displays the histogram of empirical values of Spidyn for *Microsoft* stock and $T = 30$, whereas Fig 3.7 the histogram of the artificially sampled distribution. Even though both graphs have some similarities, we cannot dismiss them so easily for being drawn from the same distribution. The authentic distribution of indicators exhibits a slimmer waist and fatter tails, especially the negative one. Moreover, the proportion of values found close to zero is very high, as opposed to that of the resampled distribution, which manifests smaller tails and a heavy middle section.

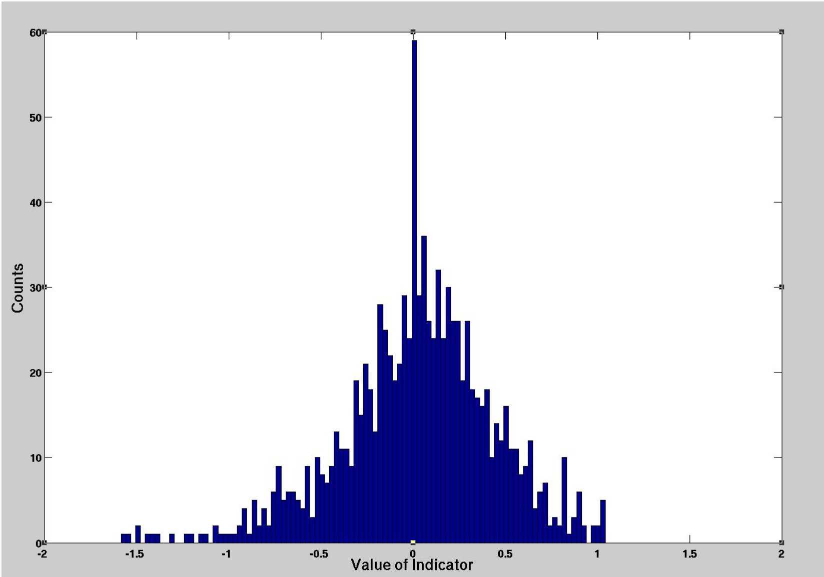

Figure 3.6: Histogram of indicators of the *Microsoft* stock for T=30 days during 2003-2006.

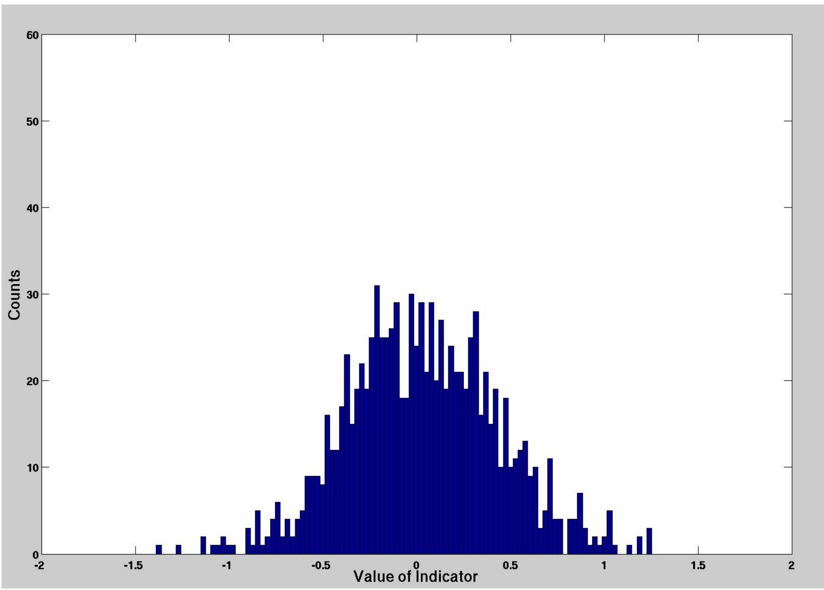

Figure 3.7: Histogram of sampled distribution with the same μ, σ and number of observations as the distribution in figure 3.6.

3.6 Conclusions

A close relationship exists between the value of the window size of Spidyn and the range of the distribution of indicators. By range of a distribution, we mean the maximum value observed minus the minimum value. We have found that increasing the window size results in decreasing the range of the index. Looking back retrospectively, the effect of the variable T has on the distribution of the index is quite logical. Let us not forget that we are trying to fit polynomials on a price time series. The longer the time series is, the flatter the polynomial will tend to be; thus, averaging out the movements of the market. However, when the same polynomial is fit on a shorter window, it will be prone to the short range variability of the market. It goes without saying, but we cannot shrink the window endlessly. After all, attempting to fit a fifth degree polynomial on a time series of only five points is a joke! Ultimately, the most visible outcome of using narrow windows to calculate the Spidyn is that indicators have a larger absolute value.

Spidyn indicators are likely not to follow a normal probability function. We have seen this from three different angles: normality plots, Anderson Darling tests and a visual comparison. Using any of the three methods listed above, we have discovered that strong deviations at the edges prevent us from validating the null hypothesis which states that indicators follow a normal distribution. Furthermore, we have learned that not all companies exhibit the same distribution, which means that `tIn` and `tOut` should incorporate some modification to take it into account. Presently, both thresholds do not contemplate this particularization among stocks. This is likely to be the single most important result to come out of the distribution tests.

Based on past experience, there is evidence to suggest that Spidyn performs better on average, when the price is quickly plummeting and then rebounding rather than the other way around [5]. Negative Spidyns are more reliable because from a descriptive point of view, crashes occur in a dramatic and abrupt way. Bubbles, on the other hand, build up gradually throughout a wider time frame and their nucleation is much more subtle and perplexing. Therefore, bubbles are harder to pinpoint before they burst and become crashes.

We have witnessed the presence of fat tails for the majority of variable configurations. This inevitably raises another issue. Are these abnormal values of Spidyn indicators linked to abnormal values of returns in the future? Is there evidence to suggest that there is additional information hidden in the strength of the Spidyn indicator? As we will see later, quantile-based thresholds derived from these questions.

To a lesser extent, we have grasped the importance of paying close attention to our input data. We have detected considerable differences in the values of the indicators generated by the black box, when using one type of input time series or another. Unfortunately, this does not throw any light as to which time series is more convenient for executing our strategy based on the indicator.

Chapter 4

Correlation Tests

4.1 Linear Correlation

Dependence leads to predictability, it is an undisputed fact. If through some means we could establish a link between indicators and lagged returns, it could be a move in the right direction for successfully detecting pockets of predictability. We know correlations within returns are extremely small because any significant correlation would lead to an *arbitrage* opportunity that would be rapidly exploited and swept away. In economics, arbitrage is defined as the practice of taking advantage of a price differential between two or more markets.

Are Spidyns and future returns correlated? The goal of this test is to detect such a correlation. We are especially anxious to discover a negative correlation between a vector of Spidyns and another one of returns projected into the future. The reason is that we would like to observe that negative Spidyns lead to positive returns and that positive Spidyns lead to negative returns. We make use of the Matlab function `corrcoef`, which returns a $N \times N$ correlation matrix where N is equal to the number of variables. In our study, $N = 2$. Each element $R_{i,j}$ is calculated by the following equation:

$$R_{i,j} = \frac{Cov(i,j)}{\sqrt{Var(i) \times Var(j)}} \tag{4.1}$$

$R_{1,2}$ or $R_{2,1}$ is a statistical measure of the strength with which the indicator is linked with returns delayed a certain factor of τ days. In other words, it quantifies how future returns can be predicted from the knowledge of a single measure of the indicator in the past. A correlation function that is zero for all nonzero time lags, would imply that lagged returns and indicators hold no linear relationship whatsoever. This is indeed disastrous news for Spidyn enthusiasts.

Consider Figures 4.1 and 4.2. Both represent the accumulated lagged linear correlation for 500 stocks. Put into simple terms, each line represents a company and each point of the line the $R_{2,1}$ value for a certain delay of τ days. The only difference between the two graphs is the criteria used on how to color them. In figure 4.1, we use blue for all of those companies which at $\tau = 1$ days express a negative linear correlation. However, in figure 4.2 we color those companies which display a negative correlation after $\tau = 9$ days. The difference is truly remarkable. For one day lags, more than half of the stocks clearly manifest a negative linear correlation. A few days later, we observe that approximately half of the stocks exhibit a negative correlation, suggesting the sign of the linear correlation is random and therefore that the link is no longer present.

Figure 4.3 is very similar to the previous two experiments except for one difference. This time we plot the correlation between indicators and a random distribution of returns centered at zero, meaning that approximately half of the values are negative and the other half positive. As we expected, we obtain a graph where the linear correlation behaves like noise for all values of τ.

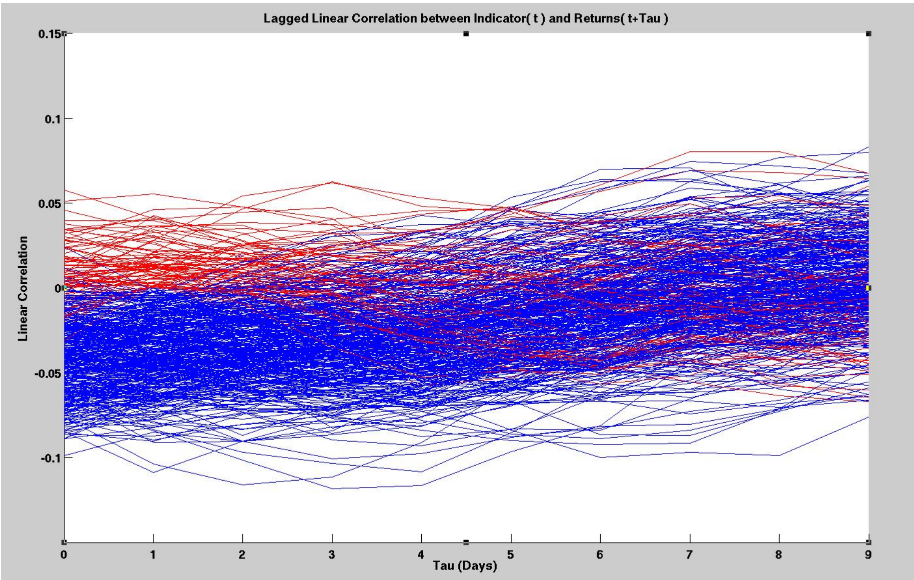

Figure 4.1: Linear correlation for 500 companies for the period 2003-2006 and T=30 days. Blue stocks have a negative linear correlation between Spidyn indicators and returns after τ=1 days.

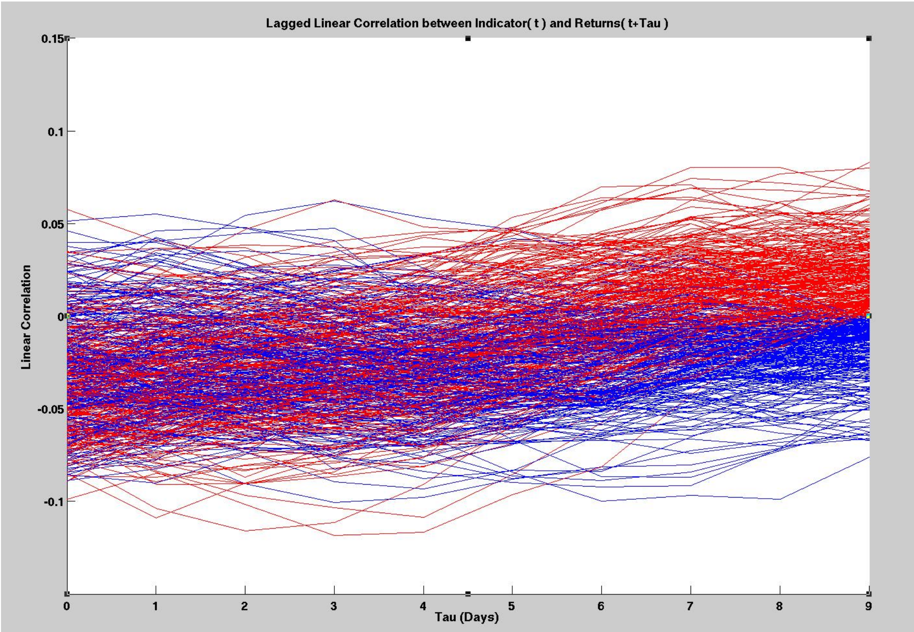

Figure 4.2: Linear correlation for 500 companies for the period 2003-2006 and T=30 days. Blue stocks have a negative linear correlation between Spidyn indicators and returns after τ=9 days.

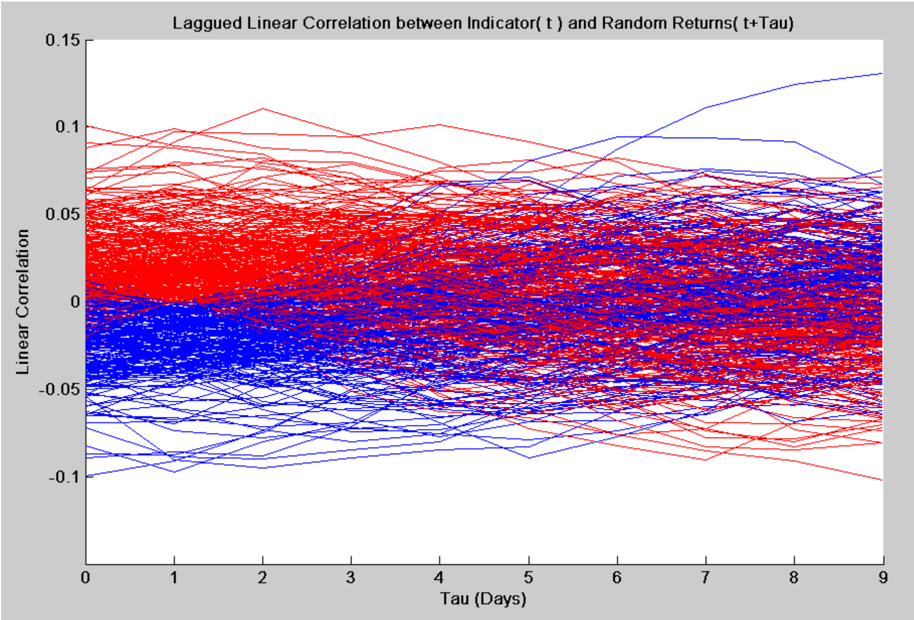

Figure 4.3: Linear correlation between indicators and a random distribution of returns centered at zero. The period is 2003-2006 and T=30 days. Blue stocks have a negative linear correlation between Spidyn indicators and random returns after τ=1 days.

4.2 Negative Products

In this experiment, we have randomly picked a company out of the S&P500. For each day, we calculate the following product:

$$W(j) = indicator(i, j) \times return(i, j + \tau) \tag{4.2}$$

where i refers to a company and j to a day. $W(j)$ is nothing but one term in the sum entering the definition of the $Cov(i, j)$ in the numerator of equation 4.1. The idea is that we are searching for a shift towards the negative values in $W(j)$. Positive values of $W(j)$ indicate that a positive indicator follows a positive return or that a negative indicator follows a negative return. Neither of them are of our interest.

Figure 4.4 exhibits how vector $W(j)$ oscillates and pulsates around zero. This delicate balance between randomness and order perceived is reminiscent of the one found in return time series. Aesthetic considerations aside, the graph is very noisy to say the least. It is extremely difficult to say that there is a shift towards negative values, however small it may be, without entering the terrain of speculation. Figures 4.1-4.3 are quite smooth because the curves $R_{1,2}$ are plotted as a function of the lag τ, which changes relatively little of the correlation coefficient. In contrast, figure 4.4 is very noisy, since we plot only one term of the numerator of equation 4.1 as a function of time (and not time lag).

Note that we are calculating two very different things in figures 4.1-4.3 and in figure 4.4. This can be inferred by comparing the roughness of the various plots. Figures 4.1, 4.2 and 4.3 are surprisingly smooth because one value represents the correlation between two vectors of more than a thousand observations, which tends to average out this coefficient. On the other hand, figure 4.4 is a very particular case: a product between two almost random signals for one company and one time lag. It is likely that if we plotted $W(j)$ for other time lags, we would notice a varying presence in the negative values. And if we calculated the average movement among companies, it should be consistent with the linear correlation tests. This is important to bear in mind, in practice however, it is extremely hard to verify using this methodology.

4.3 Subsets

A subset is a set contained within a set. Much work has been done by academics, in finding linear correlations among subsets of features in high-dimensional data. We studied a simple case by comparing a subset composed of those positive indicators that after τ time steps ended up in a negative return, to absolutely all positive indicators. Similarly, we followed the same procedure and tested negative indicators that lead to lagged positive returns, against all negative indicators. Could the distribution of successful indicators be characterized and set apart from the multitude of indicators? A modification of the Anderson Darling test was utilized in these subset tests. This new function is a generalization of the previous test and allows us to verify if k independently sampled populations come from the same distribution. Using a 95% confidence, we found that there was virtually no difference whatsoever in the distributions. We are not surprised

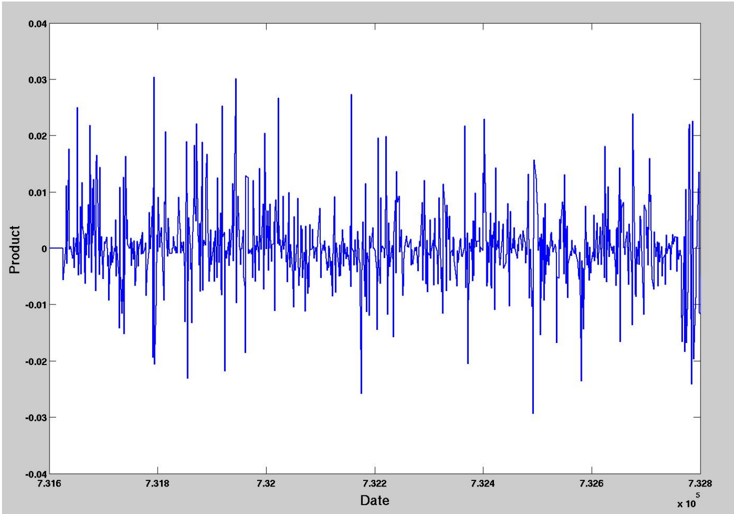

Figure 4.4: Cross correlation of the *Google* stock for 2003-2006 and $\tau = 1$ day.

by the failure of the subset test. Had these correlations been so conspicuous, such relations would have been already identified previously.

4.4 Conclusion

Figures 4.1 and 4.2 have delivered to us the first piece of evidence that supports the belief that Spidyn might have strength. These correlation tests have signaled to a faint but visible link between Spidyns and lagged returns. We can argue one way or another, whether the slight negative correlation observed is a consequence of the link between negative Spidyns and positive returns or because of the link between positive Spidyns and negative returns. However, the information is mixed as both correlations add on to the same direction. At the current stage, it is impossible to detangle one from the other. What is almost certain though, is that there is some relationship between the two and that this relationship seems to fade away shortly. After just a few days, future price returns lose all relationship with the value of Spidyn and cannot be predicted by linear extrapolations of the past. Non-linear correlations may better capture the connection. Nonetheless, such dependencies are much harder to detect and carry out in practice.

With little success, we have attempted to distinguish successful indicators from the rest. There is no reason to believe these indicators originate from another kind of distribution than their other fellow indicators. Furthermore, cross-correlation tests were found to be extremely noisy. In spite of our efforts, it is still premature to announce that Spidyn indicators manifest strength. We are in need of an indication that cannot be so easily refuted. In the next chapter, we will discover how we captured a signal that to a large extent does precisely that.

Chapter 5

The Signal

5.1 Objective

So far, we have sensed a mounting pile of evidence that links the sign of the indicator with the sign of lagged returns. We are still eager to go beyond that. It would be even more valuable to us if we could identify a bond between the value of the indicator and the sign of future returns. At any rate, our goal is to capture a signal that confirms the strength of the indicator, thus revealing some degree of predictability of Spidyn. If proven to be the case, this information could be employed to adjust more accurately the entering and exiting thresholds.

5.2 Binning

In order to understand the importance of the upcoming graphs and follow the discussion, one has to realize how they were developed. While the complexity of the procedure used to create the graphs is not enormous, the reader must be at least familiar with some concepts before we proceed to its implications.

5.2.1 Linear Binning

Initially, a naive approach was used to conduct the binning process. Each distribution of indicators was linearly divided according to its most extreme negative and positive values, meaning each binning had essentially two degrees of freedom. As a result of using this strategy, all bins for one company had the same length but contained varying amounts of data. Now that the distribution of Spidyn has been studied to some extent, one can easily understand that many of the points conglomerated around the central bins and that few indicators, if any, were found in the bins at the edges. Of course, the two bins at the edges had to include at least one indicator by definition. This method suffered from the fact that same bins belonging to different companies were not directly comparable as each binning varied significantly from stock to stock. For example, some companies with *outliers* present in their distributions lead to large sizes of bins, several of them empty. Many scientists and academics in the field use the terms outliers, kings or black swans to describe an observation that is numerically distant from the rest of the data. Moreover, even within the same company, bins could not be compared because of what we have mentioned in the previous section. After running a few experiments, the graphs produced with the linear binning procedure were found to be extremely noisy and inconclusive compared to those using quantile binning. Calculations resulted in messy averages, especially in those areas with few indicators per bin, that is, the edges. Figure 5.1. pays tribute to the phrase, “A picture is worth more than one thousand words”.

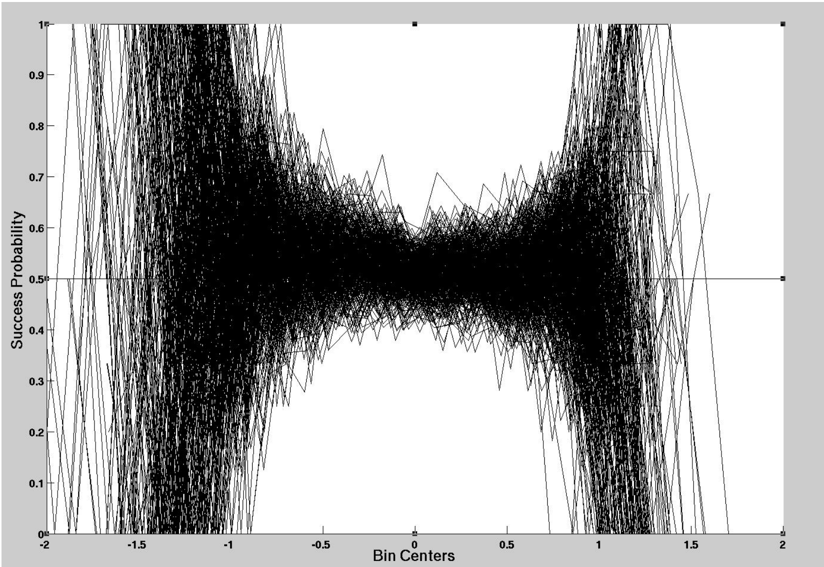

Figure 5.1: Probability of success using linear binning for the period 2003-2006. Each line represents a stock of the S&P500. The edges are extremely noisy due to a lack of indicators on which to compute the probability of observing a positive return after $\tau = 1$ day.

5.2.2 Quantile Binning

The X-axis represents Spidyn´s space split in a number of different bins. Following a suggestion from Professor Didier Sornette, the concept of quantiles was employed to guarantee that each bin which was going to be analyzed contained the same number of indicators. Quantiles refer to a fraction of points below a certain value. For example, a 30% quantile is the point at which 30% percent of the data fall below and 70% fall above that value. Hence, if we work with 20 bins, each bin contains 5% of all the data belonging to one company. The logic behind associating the same number of indicators to each bin is that statistics computed from each bin have the same value and are equally weighted. Of course, a 100% probability of success with 100 points is more significant than the same probability of success with 2 points. In respect to the number of bins we choose to work with, the more we rely on (fewer points per bin), the nosier the graphs will turn out. On the other hand, the less bins we employ (more points per bin), the larger the portion of variability that is averaged out.

5.3 Probability of Success

The Y-axis in figures 5.1,5.2 and 5.3 quantify the probability of success of an indicator belonging to a given bin, conditioned on the fact that a certain lag of τ days links Spidyn indicators with positive returns. In other words, we calculated the conditional probability of observing a positive return after τ days, for an indicator corresponding to a specific bin. For each company and bin, we compute a probability between zero and one. What we see plotted in figures 5.2 and 5.3, however, is the superposition of all lines, each representing a particular stock. This is achieved by averaging all 500 probabilities for each bin of Spidyn´s space. Note that the average probability for an indicator to lead to a drop in prices is the complementary of the probability observed on the plots. Each of the four lines plotted on a graph corresponds to a specific lag or delay. The values of the lag that were found appropriate to carry out the simulations are 1, 3 , 10 and 30. The first two numbers try to identify a short-term response of the returns, whereas the values of 10 and 30 are aimed at visualizing the behavior of the indicator once the link is presumably broken. To sum up, each point plotted describes the average probability of success for an indicator belonging to a certain range in terms of quantiles, averaged over the 500 companies which comprise the market, for a defined lag of τ days.

5.4 Testing Against Randomness

5.4.1 Purpose

Spidyn looks for dependencies. How does it react if these dependencies are washed away? We stress the importance of testing against randomness, whenever we believe to have discovered the latest technique in prediction. Surprisingly, many researchers are oblivious to this step. We would be surprised by the amount of times a discovery we believe of utmost significance is matched or even surpassed by the effects of a completely random process. For this reason, we compare the probability of success using original price time series, against randomized ones, in hopes they are distinguishable.

5.4.2 Procedure for Reshuffling Time Series

We explain the methodology used to generate a random price time series, starting off with the original one. It is advisable not to reshuffle directly the prices since we will destroy both the correlation structure and the distribution of returns, resulting in meaningless noise. Instead, a better strategy consists in:

1. Calculate the returns, R.
2. Shuffle the returns, R_k.
3. Fix the first price of the new time series to be the same as the original one, $P(0)$.
4. Reconstruct a new price time series from the returns obtained in step 2 with the following equation:

$$P'(t) = P(0)e^{\sum_{k=1}^{t} R_k} \tag{5.1}$$

where R_k are the shuffled returns.

In doing so, we preserve the distribution of returns but destroy the correlation structure in the market. Indicators drawn from non-reshuffled returns manifest more order through the correlation structures of returns than Spidyns associated to random price time series. We conducted one hundred experiments, every time repeating the steps mentioned earlier to generate new price time series. Afterwards, we obtain for each bin one hundred independent measures of the probability of success. Finally, we plot the 80% and 90% intervals of confidence for the distribution of these probabilities.

Considering intervals were generated from random prices, they shouldn't be exactly horizontal but close. After removing the underlying structure of the market, there is no particular reason to believe some Spidyns hold a higher degree of predictibility than others. By the way, it should be highlighted that the lag used for the bandwidths in figure 5.2 and 5.3 is $\tau = 1$ day. Had we chosen a larger value, these intervals of confidence would be slightly wider. Because a 30 day lag is made of returns equal to the sum of 30 uncorrelated daily returns and variance is additive, confidence levels are likely not to be the same. Let´s recall the fact that the standard deviation of a return over thirty days is $\sqrt{30} \times \sigma$, where σ is the standard deviation of a return over one day. Thus, the error bandwidths grow with the square root of the time lag τ, which is 30 in our example.

5.5 Results

Figure 5.2 and 5.3 are possibly our largest contribution to the study of the Spidyn indicator. They are significant for a variety of reasons which we will discuss shortly. The first graph corresponds to the period from 2003 to 2006, a time when the market was exhibiting a notable bullish behavior. On the other hand, the second plot belongs to the years from 2000 to 2002, a period where many will agree the market was going through a bearish regime.

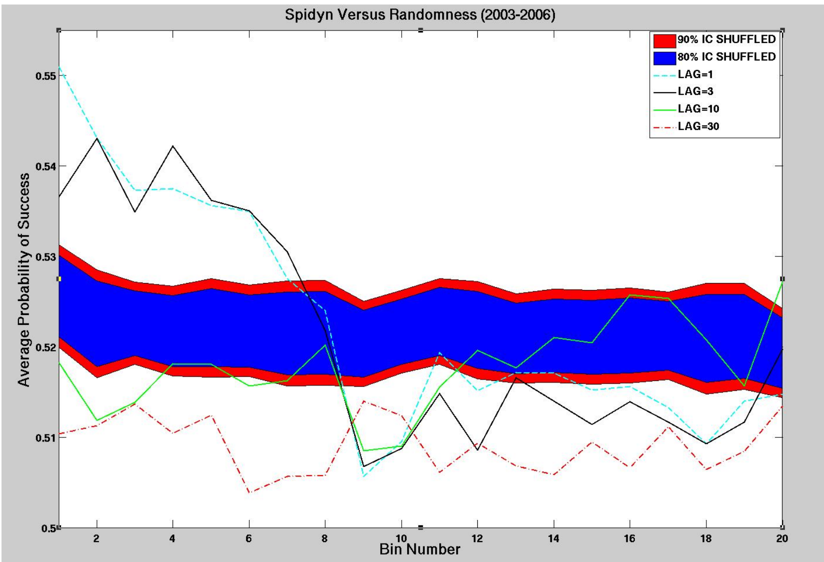

Figure 5.2: Probability of success measured for a bullish market (2003-2006) using quantile binning. The intervals of confidence were calculated with $\tau = 1$ day.

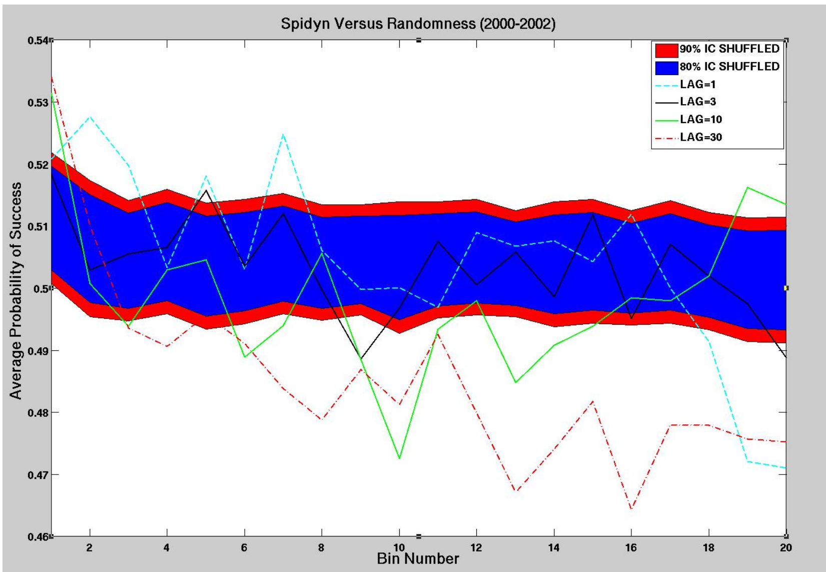

Figure 5.3: Probability of success measured for a bearish market (2000-2002) using quantile binning. The intervals of confidence were calculated with $\tau = 1$ day.

5.6 Conclusions

5.6.1 Bullish Regime

Figure 5.2 captures adequately the signal we had long been searching for, and is relevant for at least three reasons. Firstly, it lays evidence on the connection between the strength of the Spidyn indicator and the sign of lagged returns. Secondly, it also reveals that this link fades away after lags of the order of only a few days. Thirdly, it validates the added value of the Spidyn indicator when tested against randomness.

The single most outstanding feature of the 2003-2006 graph, is the characteristic leap located in the low quantiles. As we move towards the more negative end of the spectrum, a suggestive jump of 2–3% is observed for very small lags. Judging from the variability present on the graph, this leap is significantly higher than the variability within bins. According to the plot, in order to achieve the biggest chance of predicting an increase in price, we must concentrate on the 5% most negative indicators and wait one trading day. This probability on average is approximately 55%, which in finance is substantial. However, Spidyn loses its predictive edge pretty quickly and beyond the 20% quantile it cannot be set apart from noise. We would like to observe a drop in the probability of success for large quantiles, implying that big positive indicators have the virtue of predicting more accurately drops in prices. This is certainly not the case.

For lags of one and three days, both are nicely above the 80% and even 90% confidence bands of the reshuffled time series for quantiles less than 30%. This may suggest that Spidyn captures the structure of the market somewhat successfully and has indeed some added value, at least for bullish periods. Continuing on this strain of ideas, it also means that the jump previously discussed cannot be attributed exclusively to luck. To a lesser extent, it points to an underlying and subtle structure in the stock market that can be swept away by solely reshuffling the returns.

For lags of ten or more days, the chance of having positive returns following any Spidyn indicator is basically noise around a value which is close to the probability of simply spotting a positive return on any given day. During the period analyzed, this was between 51% and 52%. We are able to see how the signal is lost after the course of just a few days. Consequently, the existence of pockets of predictability is indeed very fragile and temporary.

Overall, the implications of this graph are meaningful, since it indicates that very negative Spidyns carry additional strength and offer some degree of predictability. We are more inclined to believe now that the negative correlation spotted earlier during the correlation tests, originate predominantly from the relationship between negative indicators and positive returns and not the other way around. It also reinforces the supposition Spidyn works better for crashes than bubbles, since we do not see a decline in the probability of success for the most positive bins. Last but not least, we are far from a flat line around 50%, which would indicate the indicator is incapable of finding unsustainable trends.

5.6.2 Bearish Regime

The probability of success studied in the period from 2000-2002 exhibits a discordant feature if compared to the one from 2003-2006. During this time of crashes and downfalls, large lags seem better suited in linking negative indicators to positive returns. A reasonable explanation for this phenomenon is that during this interval of time, a number of prolonged crashes were followed by sudden rebounds. In mid-crash, negative indicators were thriving, but from a correlation standpoint, small lags of the order of a couple of days were not explaining positive returns. On the contrary, larger lags were required on average for these negative indicators to be linked to positive returns, which were abundant in the rebound phase. It can be easily seen that the unconditioned probability of having a positive return on any given day is much lower than during a bullish market. In particular, we learn this from observing the intervals of confidence which are now centered around 50% or 51%. The bandwidths have also increased drastically as a result of a higher variability associated with the intervals during this time.

The signal we picked up for the bullish regime of the market is not visible anymore. Small lags are no longer outside the intervals of confidence, conveying that they are not discernable from random behavior. This discredits the added value of Spidyn in predicting increases in prices in times of a generalized pessimistic market sentiment. Furthermore, large quantiles persist in not revealing any helpful information on prices drops.

There is yet another characteristic of the graph that is noteworthy. For delays of thirty days and indicators greater than the 40% quantile, we notice that the line falls strongly outside the confidence band. This can also be seen to happen transiently for ten day lags and indicators over the 50% quantile mark. Both are indications that upward trends led to an anomalous drop subsequently in the following 10 and 30 days. This is clearly a diagnostic that there is a strong deviation from absence of dependence. It does not translate into a prediction which is successful, but there is an anomaly. The market was anomalous during this period.

Overall, we see that Spidyn is strongly time-dependent, besides being company dependent, which has been seen previously. Hence, we believe the `tIn` and `tOut` thresholds should take into consideration both dependencies. From a descriptive point of view, both graphs are significantly different. They externalize divergent dynamics of the market, to which Spidyn is not indifferent.

Chapter 6

Portfolio Tests

6.1 Notions

6.1.1 Definitions

Portfolio optimization is one of the key topics in finance. It can be characterized as a search for a satisfactory compromise between maximization of the investor´s capital and minimization of the related risk. In finance, a *portfolio* is an appropriate collection of investments held by an institution or a private individual. *Assets*, or components in a portfolio may include stocks, bonds, options, warrants, gold certificates, real estate, futures contracts, production facilities, or any other item that is expected to retain its value. *Portfolio management* involves deciding which assets to include in the portfolio, how many to purchase and when to purchase them, given the goals of the portfolio owner and changing economic conditions. In order to accomplish this some sort of performance measurement, most typically the expected return on the portfolio and the risk associated with this return are used in the selection process. In our case, the Spidyn with the help of a predefined strategy administers and controls our portfolio. If possible, we would like to analyze everyday our possibilities and reallocate our wealth as we receive more information. That is not possible without paying the price of *slippage*. Slippage is a friction cost manifested in financial markets, provoked by market orders which are usually not executed at the order price, due to limited liquidity and finite human execution time.

In portfolio optimization, there is no such thing as a best solution. Instead a collection of optimal strategies exist, each one corresponding to a level of risk aversion of the investor. *Risk aversion* is the reluctance of a person to accept a bargain with an uncertain payoff rather than another bargain with a more certain, but possibly lower, expected payoff. Let us imagine a person is given the choice between two scenarios, one certain and one not. In the uncertain scenario, the person is to make a gamble with an equal probability between receiving \$100 or nothing. The alternative scenario is to receive a specific dollar amount with a deterministic probability equal to 1. Investors have different risk attitudes. A person is:

- risk-averse if he or she would accept a certain payoff of less than \$50 (for example, \$40) rather than the gamble.

- risk neutral if he or she is indifferent between the bet and a certain $50 payment.
- risk-seeking if the certain payment must be more than $50 (for example, $60) to induce him or her to take the certain option over the gamble.

6.1.2 Markowitz

Significant progress in our understanding of the stock market was acquired by Markowitz with his mean-variance Portfolio Theory[16], the capital asset pricing model of Sharpe [17] or Black and Scholes´s option pricing and hedging theory[18]. For over half a century, Markowitz´s mean-variance (MV) approach has been the standard for efficient portfolio construction. He stated that a rational investor should either maximize his expected return for a given level of risk, or minimize his risk for a given expected return. Today, nearly all commercial portfolio optimizers for asset allocation are based on some variation of it because his theory makes investment sense and is easy to implement. Nevertheless, mean-variance optimized portfolios have been shown not to perform well[19]. The approach places excessive weights on assets with large excess returns, regardless of possible estimation errors in the input data.

6.2 Quantile-based Thresholds

Now that some evidence has been laid on the connection between the strength of the Spidyn indicator and lagged returns, (especially between negative Spidyns and positive returns and for lags of the order of a few days) it is appropriate to incorporate this piece of knowledge into our trading strategy. As a reminder, it has also been suggested in chapter 2 that not all distributions among companies are statistically identical. On a subsequent test, we look at it from another perspective. First, we fix a certain quantile for all 500 companies and plot these values. Next, we determine if the supposition that a certain quantile is represented by the same value for all stocks is believable by evaluating the dispersion of this distribution. If not proven the opposite, the idea that a constant can be applied as a threshold to all stocks in the S&P500 cannot be opposed. Figure 6.1 displays such distribution for four different quantiles. After examining the graph, the hypothesis that companies possess the same distribution of Spidyns is not defensible, since the standard deviation of the threshold is found to be very large. Small quantiles exhibit larger standard deviations because they are more susceptible to variations, as they depend on fewer data and in turn become noisier. Under the light of these results, we consider that each company must be treated independently and compared against a different threshold.

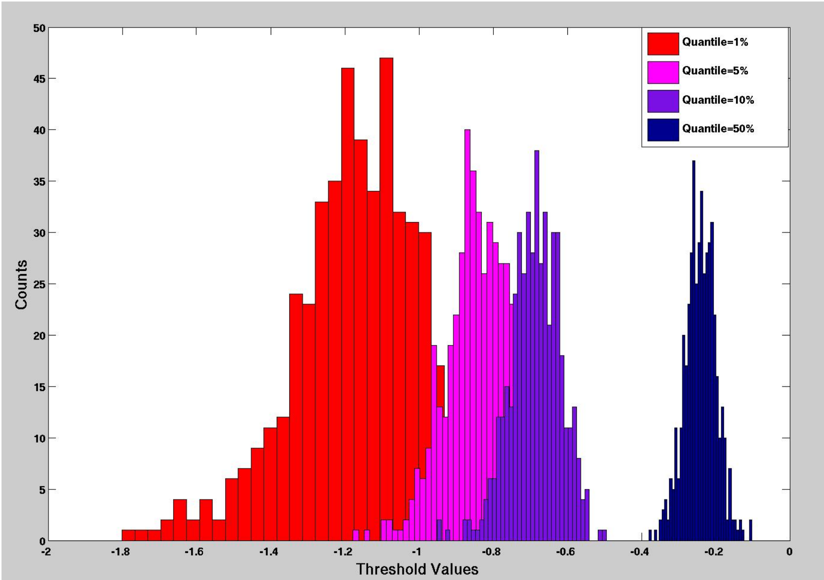

Figure 6.1: Histogram of thresholds for different quantile values

We are thus in need of a technique that enables us to scale all stocks according to a common measure and allows us to trade exclusively with the most negative segment. Quantiles have been used for a second time. Previously, we have seen how they were used in the binning process and how they were instrumental in capturing the signal. Quantiles are a unifying criteria because they treat all companies in relative terms and normalize them so that they are comparable. Parenthetically, by this we mean that an indicator of -0.5 from two different companies might not be analogous to each other in relative terms, whereas a 5% quantile is. Continuing on this strain of ideas, we will trigger buying orders conditioned on the fact the indicator is confined in a low quantile, the value of which is another parameter of itself. Depending on the definition of `tIn` and `tOut`, we will restrict or widen the range of Spidyn values where trading is accepted. The greater the quantile value on which the threshold is based, the greater the volume of trading. From a more practical point of view, a possible implementation of this idea would be to work with a vector of `tIn´`s, one per company.

It is worth mentioning that by calculating these thresholds from historical data, we are in fact cheating since in a real world scenario we would not a priori know the distribution of Spidyns. By means of back testing, we could assess the suitability of this modification and if proven to work, the distribution of Spidyns could be updated as we move forward into time and experience different market regimes. One aspect to take into consideration is that the system now has memory and depends on the starting time of the simulation. Should it be reset at some point? Nevertheless, if we imagine there is only one market regime from the present onwards to infinity, `Tin` and `Tout` will eventually stabilize since we will be adding more data with the same proportions.

The adjustment of the `tOut` parameter is likely to be more complex, although a variety of options can be tested. This inevitably raises the following question:

when is it optimal to abandon the market? If we recall the graph that showed the acceleration of the S&P500 over the last decades, we are better off holding the stock indefinitely. Some strategies which could be tested are for instance:

1. `Tout` = `Tin`.
2. `Tout` = 0.
3. Hold until the Spidyn is contained in the X% most positive percentile.
4. Hold until the price declines more than a certain value, no matter what the value of the index is.

In our portfolio tests, we used strategy number one.

6.3 Description of Statistics

The goal of the following experiments will be to determine if portfolios using "adaptable thresholds" outperform portfolios with static thresholds. By static we understand that `tIn` and `tOut` are invariant throughout time and among companies. The portfolio tests results can be found in appendix A. The metrics needed to interpret them are described below:

- `Deals-` refers to the number of effective deals that have been traded and is actually equal to or less than the number of signals we receive to open a new position. The reason why sometimes a buy order is not satisfactory in spite of getting a signal is because `Stop-Loss` orders get triggered.

- `Annual Turnover-` it quantifies how quickly a fund turns over its holdings. Annual turnover is equal to the total transaction volume of the trades in one year divided by the total portfolio size.

- `Average Profit & Loss per Deal-` the average return per deal expressed as a percentage:
$$\text{P\&L} = \frac{\text{(sell price - buy price)}}{\text{buy price}} \tag{6.1}$$

- `Sharpe per Deal-` is a measure of the return per unit of risk in an investment. In Gilles´ code, $R_f = 0$, returns are the P&L´s per deal and σ their standard deviation. The complete formula is as follows:
$$\text{Sharpe Ratio} = \frac{\text{E[R-}R_f\text{]}}{\sigma} \tag{6.2}$$

- `Portfolio Performance per year-` measures how much money our portfolio has generated as a whole with respect to the initial value of our wealth, which is $1 in cash. Gains are compounded, meaning that the money earned through trading is reinvested and not withdrawn from our investment universe.

- `Sharpe-` same concept as Sharpe per deal, although in this case it is measured up against the returns of our total wealth and its variability.

- `Success Rate`- we consider that a deal is successful if and only if sell price > buy price.
- `Maximum Drawdown`- largest consecutive drop of the value of our wealth in relation to its peak.

6.4 Conclusions

The ultimate verification of Spidyn´s predictive power is a real-world portfolio implementation. If the indicator provides some degree of predictability, this should be reflected in the following portfolio tests, which in principle are the best test of the system. It seems more sensitive to identify a Spidyn signal than the quantile prediction of “success”. The latter, however, is useful as an independent statistics of the indicator.

6.4.1 Bullish Regime

Even in the presence of a wide range of parameters, some trends become visible in our tables of portfolio results. In line with our results obtained from chapter 5, the average `P&L per deal` and `Sharpe per deal` apparently increases as we restrict trade to the indicators in the lower quantiles. This has to be seen in the light of figure 5.2 and reflects the notion that using a low quantile value to define the thresholds, we achieve more profitable trades on average. To our knowledge, this fact has never been actually proven before. However, we also see how the number of deals decreases for small quantile values and reduces this recently discovered advantage. The natural reaction is to try to compensate this effect by increasing `Investfraction`. But to account for the difference in volume traded, `InvestFraction` would have to be ridiculously high in the case of restraining activity to low quantile thresholds. Would a potential client be willing to invest 30% of his total wealth on a new opening position? Probably not.

As `InvestFraction` increases, the number of deals stays the same, since in theory the number of signals we get to open new positions has to remain unaltered. There seems to be some exceptions to this rule, especially for sufficiently large values of trading. This is achieved either when `InvestFraction` and/or the `quantile` value of the threshold go up. The reason why considerable amounts of trading trigger `Stop-Loss` orders is that the drop required is more likely to occur. Let´s consider a simple example with two different values of `Investfraction`, say, 5% and 10%. Our initial wealth is always \$1, all in the form of cash. In response to an entry signal, in one case we invest 5 cents and in the other 10 cents. Notice that our wealth is still \$1 for both cases since we have converted cash into stocks. If the following day there is an unfortunate 50% drop of the stock, our wealth is then \$0.975 and \$0.95 for 5% and 10% respectively. This translates into a 2.5% and 5% drop of our wealth. As can be seen, the larger the trading volume, the more prone we are to face a drop of the characteristics required to fire a `Stop-Loss` order. When a `Stop-Loss` order is issued, we are then forced to cut all positions and stay out of the market during the next days. In doing so, we miss some incoming trade signals, which could or could not be promising.

`Portfolio performance` has to be compared to the yearly performance of the index, which for this period was of 15% and annual `Sharpe ratio` of 1.2. Despite yielding good results with the quantile-based thresholds, the simpler strategy of setting `tIn` and `tOut` equal to -0.5 for all companies produces higher portfolio performances for the same value of `InvestFraction`. Why is a trading strategy with more profitable deals underperforming? The key point here is that the portfolio performance obtained needs to be compared on the basis of realistic transaction costs. For instance, when `InvestFraction` = 15% and we use fix thresholds, the `portfolio performance` of 150% is rapidly diminished if we apply transaction costs. Transaction costs are not taken into account directly in the algorithm, it is up to us to interpret if results are still profitable in a real world scenario. If a deal is made of buy and then sell, and the transaction cost is 0.05% per one way, this produces a 0.1% cost both ways. Hence, the total cost amounts to $7785 \times 15\% \times 0.1\% = 117\%$ which erases much of the 150%! Let´s shift our attention to quantile-based thresholds. For example, if we use `quantile` = 1% and the same `InvestFraction`, we gain a mere 32% for the same period. Yet if we consider the costs associated to this strategy, we only incur 14.6%. In other words, our costs represent approximately half of our earnings, whereas if we employ static thresholds this proportion increases to three fourths.

Logically, the volume of shares traded and the portfolio performance also increase at approximately the same rate as `InvestFraction`. This is not perceived in the `Annual Turnover` which remains nearly untouched, since the effect of increasing trade is neutralized by increasing the portfolio size. It must also be noted that for static thresholds and very high values of `InvestFraction`, we do not cap the `MaxLeverage` anymore. Under these conditions, maximum leverages fetch values of the order of 2. Last but not least, even though we raise the `InvestFraction`, the `average P&L` is not modified if no `Stop-Loss` orders are issued. One has to bear in mind that we are comparing two percentages, that is, two relative values. In absolute terms the gains resulting from the trade of shares could be completely different, provided that `InvestFraction` is also different.

6.4.2 Bearish Regime

Results during 2000-2002 are somewhat blurry. It is obvious that `portfolio perfomances` and their `Sharpe ratios` are significantly smaller than those of 2003-2006. As a reference, they can be compared up against the yearly performance of the SP&500 index, which for this period was of -15% and annual `Sharpe ratio` of -0.68. The trend of many of the remaining parameters is arguable and difficult to explain. For large volumes of trading (when `InvestFraction` and/or the `quantile` value of the threshold go up.), `Sharpes per deal` are extremely small and sometimes even negative. This can be expected from a period where crashes were dominating the market. However, if we restrict trade to very small quantiles, `average P&L´s` and `Sharpe per deal` are unusually high owing to a small number of extremely profitable trades (of the order of 100% per deal). This means that `average P&L´s` are actually bigger than for 2003-2006 if we keep the other parameters fixed. The reason may be that Spidyn detected and exploited successfuly such negative accelerations during the successive crashes and rebounds. `Success rates` clearly fall below those observed during the bullish regime. To a certain extent, this was predictable after analyzing figure 5.3. For big quantiles, `Maximum drawdowns` systematically increase with `InvestFraction`. The drawdown effect is explained in 6.4.1 with a numerical example. Overall, we perceive more disorder in the results obtained, as opposed to the 2003-2006 period.

6.5 Where To Move Next

Unfortunately, due to time constraints we were unable to conduct all the tests we had in mind. It would be very interesting to see some kind of return-risk representation of our strategies as a function of the quantile parameter. Can we identify a region of optimal values much in line with the efficient frontier of Markowitz? Below are a set of ideas that can either stand alone or be combined with quantile-based thresholds:

Data Resampling We conceive a price time series as only one of the many possible outcomes. By adding some noise to the time series, we obtain an ensemble of which we try to identify the average behavior. Countless ways of adding noise to a time series exist, we would like to discuss only two of them. First, a possibility that appears reasonable is to multiply one by one the original time series by a vector of random numbers of the same length, each of which has a mean equal to one and a tiny standard deviation. Another more sophisticated option is to use the high, low and opening time series instead of adding meaningless artificial noise. It must be made clear to the reader, that throughout our study we have worked solely with closing prices and that high, low and opening time series might provide additional insight. The idea of data resampling is being used for the LPPL (log-periodic power law) in the Entrepreneurial Risks group, so an implementation of it is not far-fetched. The newly adjusted Spidyn indicator would be an average of the various Spidyn indicators generated from the set of time series. It would be interesting to analyze how the Sharpe value for the portfolio reacts as we increase the value of sigma of the noise. Does it decrease? Does it stabilize before reaching zero? If so, what is this value?

Repeating Signals It consists of a very straight-forward approach. We only undertake an action if we get at least N consecutive non-conflicting signals, while the window size T remains a constant. For example, we would sell a given stock if we received 3 sell signals in a row. The higher N is, the less transaction costs we incur (since there is less trading). Nevertheless, we might also be missing out on profiting opportunities and running the risk of holding in our portfolio not-so-good securities. Despite this suggestion goes against the very spirit of Spidyn and the concept of promptly spotting pockets of predictability, it could be evaluated.

Different Time Scales The goal is to obtain indicators belonging to different time scales, that is, with varying window sizes. We compute indicators projecting back into the past with different time frames and draw a Spidyn for each one. If all signals are congruent with each other, we continue on with the trading algorithm. As with the other suggestion, we are adding restraints to our trading scheme. This proposition suggests a possible hierarchal structure is present in the market that could be seized by the indicator. It might occur that the market behaves similarly when observed with different levels of magnification.

Parameter Space Exploration Throughout our research, we have dealt with a portion of Spidyn´s parameter space. It would be interesting to test future experiments with variations of the set of parameters. In particular, we emphasize the exploration of the degrees of the polynomials and the time periods on which tests were conducted. We have worked mainly with 2000-2002 and 2003-2006, but how does Spidyn perform during other periods of time?

Dynamic Invest Fraction We have regulated the range of indicators where trading is accepted by using quantiles. This has a direct influence over the number of positions we open. However, this method could be sophisticated by adjusting dynamically the amount invested when we open such position. Would it depend on the value of the indicator or maybe some other measure?

Appendix A

Portfolio Test Results

The following tables summarize the portfolio tests conducted on a wide range of parameters during the periods of 2000-2002 and 2003-2006. Due to the large size of tables and their hard disposition, we have sacrificed visual appearance in favor of readibility and clarity.

InvestFraction	# Deals	Annual Turnover	Average P&L	Sharpe per Deal	Portfolio Perf.	Sharpe	Success	Max. Drawdown
1%	1002	16360%	0.75%	0.26	2%	1.5	60%	-0.75%
2%	1002	16350%	0.75%	0.26	3.9%	1.5	60%	-1.5%
5%	1002	16350%	0.75%	0.26	10%	1.5	60%	-3.7%
10%	1002	16360%	0.75%	0.26	21%	1.5	60%	-7.3%
15%	976	16910%	0.78%	0.26	32%	1.6	60%	-11%

Table A.1: Bullish market using quantile = 1%

InvestFraction	# Deals	Annual Turnover	Average P&L	Sharpe per Deal	Portfolio Perf.	Sharpe	Success	Max. Drawdown
1%	1688	14230%	0.63%	0.22	2.8%	1.4	56%	-1.1%
2%	1688	14230%	0.63%	0.22	5.6%	1.4	56%	-2.1%
5%	1688	14260%	0.63%	0.22	14%	1.4	56%	-5.1%
10%	1655	14620%	0.65%	0.22	30%	1.5	56%	-9.9%
15%	1587	14950%	0.69%	0.23	48%	1.6	58%	-12%

Table A.2: Bullish market using quantile = 2%

InvestFraction	# Deals	Annual Turnover	Average P&L	Sharpe per Deal	Portfolio Perf.	Sharpe	Success	Max. Drawdown
1%	3705	11410%	0.55%	0.18	5.3%	1.4	58%	-1.9%
2%	3705	11410%	0.55%	0.18	11%	1.4	58%	-3.8%
5%	3662	11610%	0.55%	0.19	28%	1.4	58%	-8.9%
10%	3467	12050%	0.50%	0.17	52%	1.6	57%	-14%
15%	3189	13180%	0.40%	0.14	58%	1.4	57%	-17%

Table A.3: Bullish market using quantile = 5%

InvestFraction	# Deals	Annual Turnover	Average P&L	Sharpe per Deal	Portfolio Perf.	Sharpe	Success	Max. Drawdown
1%	9151	8988%	0.54%	0.16	13%	1.8	57%	-3.4%
2%	9151	8932%	0.54%	0.16	28%	1.8	57%	-6.7%
5%	8602	9931%	0.45%	0.15	59%	1.8	57%	-14%
10%	7396	10870%	0.36%	0.12	85%	1.6	55%	-37%
15%	6797	13260%	0.32%	0.11	110%	1.4	54%	-40%

Table A.4: Bullish market using quantile = 15%

InvestFraction	# Deals	Annual Turnover	Average P&L	Sharpe per Deal	Portfolio Perf.	Sharpe	Success	Max. Drawdown
1%	11572	8327%	0.49%	0.14	15%	1.7	56%	-5.3%
2%	11572	8388%	0.46%	0.13	30%	1.7	56%	-10%
5%	10562	9569%	0.34%	0.10	53%	1.5	55%	-27%
10%	8976	11250%	0.31%	0.10	86%	1.3	54%	-38%
15%	7785	12160%	0.36%	0.12	150%	1.6	55%	-45%

Table A.5: Bullish market using tIn = tOut = -0.5

InvestFraction	# Deals	Annual Turnover	Average P&L	Sharpe per Deal	Portfolio Perf.	Sharpe	Success	Max. Drawdown
1%	181	32120%	1.3%	0.25	0.84%	1.1	57%	-0.57%
2%	181	32140%	1.3%	0.25	1.7%	1.1	57%	-1.1%
5%	181	32630%	1.4%	0.26	4.2%	1.1	58%	-2.8%
10%	178	32730%	1.4%	0.26	8.4%	1.1	58%	-5.6%
15%	142	31920%	1.1%	0.21	7.8%	0.77	56%	-8.4%

Table A.6: Bearish market using quantile = 1%

InvestFraction	# Deals	Annual Turnover	Average P&L	Sharpe per Deal	Portfolio Perf.	Sharpe	Success	Max. Drawdown
1%	238	22300%	3%	0.37	2.5%	1.3	69%	-0.58%
2%	230	22650%	3%	0.36	4.8%	1.2	67%	-1.1%
5%	228	22890%	3%	0.36	12%	1.2	67%	-2.9%
10%	228	22790%	3%	0.36	24%	1.3	67%	-5.6%
15%	227	22770%	2.8%	0.36	34%	1.2	67%	-8.4%

Table A.7: Bearish market using quantile = 2%

InvestFraction	# Deals	Annual Turnover	Average P&L	Sharpe per Deal	Portfolio Perf.	Sharpe	Success	Max. Drawdown
1%	598	15530%	1.9%	0.14	4.0%	1.4	53%	-3.2%
2%	588	15760%	1.9%	0.13	7.8%	1.4	53%	-6.4%
5%	588	17500%	1.5%	0.10	15%	1.2	50%	-16%
10%	564	16510%	1.0%	0.11	20%	1.1	50%	-21%
15%	539	16890%	1.1%	0.12	31%	1.1	50%	-31%

Table A.8: Bearish market using quantile = 5%

InvestFraction	# Deals	Annual Turnover	Average P&L	Sharpe per Deal	Portfolio Perf.	Sharpe	Success	Max. Drawdown
1%	1484	11090%	0.07%	-0.021	0.30%	0.08	47%	-8.5%
2%	1467	11420%	0.03%	-0.027	-0.02%	-0.003	46%	-18%
5%	1461	11880%	0.07%	-0.019	0.74%	0.05	47%	-30%
10%	1325	13220%	0.25%	0.012	7.8%	0.27	48%	-41%
15%	1255	14460%	0.24%	0.014	8.6%	0.20	47%	-52%

Table A.9: Bearish market using quantile = 15%

InvestFraction	# Deals	Annual Turnover	Average P&L	Sharpe per Deal	Portfolio Perf.	Sharpe	Success	Max. Drawdown
1%	3905	7814%	0.19%	-0.003	2.4%	0.35	50%	-12%
2%	3805	8206%	0.23%	0.004	5.5%	0.42	51%	-21%
5%	3790	8508%	0.20%	0.006	11%	0.38	50%	-40%
10%	3121	11570%	0.20%	0.007	15%	0.34	51%	-47%
15%	2702	14970%	0.37%	0.042	50%	0.79	52%	-35%

Table A.10: Bearish market using tIn = tOut = -0.5

Appendix B

How the Greatest Investors Tell Us To Invest

As a bonus, we have included an annex which attempts to collect advice from some of the most remarkable investors of the century. While going through the extensive literature, we have tried to avoid those one hit wonders, which have been overexposed in the media and offer no real value to the apprentice. Instead, we have devoted more attention to those methodical investors which have consistently beat the market as a result of a superior strategy and solid investing techniques. The following four investors have been selected based on their achievements and impact on the investing community: Benjamin Graham, Philip Fisher, Warren Buffett and Peter Lynch.

Benjamin Graham considered the first proponent of *value investing*, an investment approach that generally involves buying securities whose shares appear underpriced by some form of fundamental analysis. He is the author of the book *The Intelligent Investor*, one of the most widely recognized investment books in the world. Graham emphasizes that nobody ever knows what the market will do. That includes analysts, bankers, casual investors and your grandmother. Another point he stresses is that we are all part of the general public, and the general public is usually wrong. To counteract our inherent emotional weakness, Graham suggests to automate parts of our investment strategy and to rely on hard data. Although ignoring bold headlines and great stories from our friends is not always an easy thing to do, it is crucial for sound investment. Graham's favorite allegory is that of Mr. Market, a very obliging fellow who turns up every day at the stock holder's door offering to buy or sell his shares at a different price. Frequently, the price quoted by Mr. Market seems plausible, but often it is ridiculous. The investor is free to either agree with his quoted price and trade with him, or to ignore him completely. Mr. Market doesn't mind this, and will be back the following day to quote another price. The point is that the investor should not regard the whims of Mr. Market as determining the value of the shares that the investor owns. He should profit from flaws from the market rather than participate in them. Another contribution by Benjamin Graham was the concept of margin of safety, which put in simple terms, is the difference between a company´s *business valuation* and its *market valuation.* Business valuation refers to how much the company should be worth if it was liquidated today, whereas market valuation is the price the market has placed on the stock. Every investor should know for each stock in his or her portfolio its respective margin of safety, so that he or she can decide how much a stock´s price can drop and still be a good investment[21].

Philip Fisher regarded as a pioneer of the field of *growth investing*, a style of investment which contrasts with that of value investing. Growth investing is based on investing in companies that exhibit signs of above-average growth, even if the share price appears excessively high in terms of metrics such as price-to-earning (P/E) or price-to-book ratios. Fisher firmly believed that investors should acquire businesses with the ability to grow sales and profits over the years at rates greater that their industry average. In his classic book, *Common Stocks and Uncommon Profits*, Fisher declares that the best time to sell a stock is "almost never"[22]. He advices to look out for capable management, which according to him is one that is willing to sacrifice immediate profits for long-term gains and at the same time maintains integrity and honesty with shareholders. To succeed at investing, Fisher thought people should concentrate on their circle of competence, or in other words, in those areas the investor is already familiar with. Within that circle of competence, investors should conduct thorough and unconventional research to understand the superiority of a company over its competitors. In opposing extensive diversification, he rarely placed more than ten companies in a portfolio, and even at that, most of the money was usually concentrated in three or four stocks.

Warren Buffett For many, one of the world's greatest stock market investors. Amassing a fortune worth more than US $62 billion, he was ranked by Forbes as the richest person in the world in 2008[20]. He has repeatedly criticized the financial industry for what he considers to be a proliferation of advisors who add no value but are compensated based on the volume of business transactions which they facilitate. Moreover, he has pointed to the growing volume of stock trades as evidence that an ever-greater proportion of investors' gains are going to brokers and other middlemen. Buffett condemns the academic position that the market is efficient and that beating the S&P 500 is pure chance. Just as a business puts more money into its most successful ventures, he said, you should similarly invest more money in those stocks that are performing well. One of the world´s greatest investor does not check stock prices because they are unreliable indicators of how much a company is worth. Mr. Buffett emphasizes buying quality companies rather than speculating about the direction of the price. The reason being that good companies are still good when times are bad. He advocates his definition of a quality company as being one that often reports high net profit margins and is able to direct the revenues it generates wisely. Furthermore, he pays close attention to whether or not the company´s earnings grew enough to justify the cost of reinvesting the previous year´s earnings. This is the primary reason why he is more concerned on *return on equity* instead of plain earnings. Return on equity is a measure of a corporation's profitability that reveals how much revenues a company generates with the money shareholders have invested. Mr. Buffett determines the value of a company by projecting its future cash flows and discounting them back to the present with the rate of long-term U.S. government bonds. Like Fisher, he believes in a focused portfolio and standing by them through thick and thin. Furthermore, he discourages the common practice among investors of selling their top performers. He manages his portfolio the same way he manages a business. Would you sell a division of your business that is consistently showing a profit? Probably you would not sell it off. In fact, it is highly likely that you would invest even more money. Above all else, Warren Buffett promotes the idea of examining business, not stock prices.

Peter Lynch His most famous investment principle is simply, "Invest in what you know", popularizing the economic concept of local knowledge. This simple principle resonates well with average non-professional investors who don't have time to learn complicated quantitative stock measures or read lengthy financial reports. Mr. Lynch has also stated that the individual investor is more capable of making money from stocks than a fund manager, because as a consumer you are conducting research all the time. He has outlined many of the investments he found when he was out with his family, driving around or making a purchase at the mall, not in the office. For mainly two reasons, Mr. Lynch is known for constantly searching for companies which are buying back shares. First, when fewer shares are circulating among the general public, earnings per share increase. Second, it shows faith in the company and entices its management to work harder and be more committed with their jobs. Before buying any stock Peter Lynch always forces himself to understand why he buys, by giving a two-minute monologue that covers the reasons why that particular stock is interesting to add in his portfolio. He explains that the real benefit of this is that it makes you know your companies, which is handy no matter which way the stock price moves because you can make your buy and sell decisions from information from the company instead of information from the market.

Acknowledgments

I would like to thank my supervisor Didier Sornette, for going through many problems from a theoretical perspective and for communicating strong convictions on how scientific work should be pursued. I also would like to express my gratitude to my other supervisor Ryan Woodard, for inspiring discussions on many issues, for helping me out with technical problems, and especially for his generous attitude of never imposing his own ideas, but instead allowing me to develop my own lines of thought. On a personal level, I am greatly indebted to both my parents, Angel Taxonera de la Cruz and Joana Morera Riera, for having granted me this priceless opportunity to study abroad and for being a perpetual wheel of support. Finally, I wish to thank the fellow student from my home university, who turned down the position to come to Zürich and enabled me to be where I am now. I will never know who it was.

Bibliography

[1] Didier Sornette. *Why Stock Markets Crash?* Princeton University Press (2004).

[2] Entrepreneurial Risks Notes. Bubbles and Crashes in Entrepreneurial Dynamics. *http://www.er.ethz.ch/teaching/Lect_6_Bubbles_crashes_08.pdf*

[3] Bak, P. *How Nature Works: The Science of Self-Organized Criticality* (1996).

[4] L. Bachelier. *Théorie de la spéculation, Annales scientifiques de l'.N.S.* 3e serie (1900).

[5] Didier Sornette, personal communication.

[6] Rajnish Mehra. *The Equity Premium: Why Is It a Puzzle?*(2003).

[7] Types of Market Efficiency. *http://en.wikipedia.org/wiki/Efficient_market_hypothesis*

[8] Technical indicators *http://www.ashkon.com/settechnicalindicators.htm*

[9] Investopedia Definitions *http://www.investopedia.com*

[10] Mean Reversion *http://en.wikipedia.org/wiki/Mean_reversion*

[11] Paul Samuelson and William D. Nordhaus. *Economics*, Mc Graw-Hill (2004).

[12] Keynes, J.M. *The General Theory of Employment, Interest and Money* (1936).

[13] Gilles Daniel. *Spidyn Report* (August 2007).

[14] Gilles Daniel. *Spidyn Report* (April 2007).

[15] Matteo Marsili. *Dissecting Financial Markets: Sectors and States* (2008).

[16] Harry Markowitz. *Portfolio Selection*, The Journal of Finance (1952).

[17] Sharpe. *Capital Asset Prices: A theory of market equilibrium under conditions of risk*, The Journal of Finance (1964).

[18] Black, F. and Scholes. *The pricing of options and corporate liabilities*, The Journal of Political Economy (1973).

[19] Michael J. Best and Robert R. Grauer. *On the sensitivity of mean-variance-efficient portfolios to changes in asset means: Some analytical and computational results* (1991).

[20] Forbes Website. *http://www.forbes.com/2008/03/05/richest-people-billionaires-billionaires08-cx_lk_0305billie_land.html*

[21] Benjamin Graham. *The Intelligent Investor*, Harper Business Essentials (2003).

[22] Philip Fisher. *Common Stocks and Uncommon Profits*, PSR Publications (1984).