ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

SCHOOL OF ECONOMIC SCIENCES

DEPARTMENT OF ECONOMICS

EFFICIENT MARKET HYPOTHESIS
AND THE STOCKHOLM EXCHANGE MARKET

THEOFANIS ELEFTHERIOS
GKILIS

Dissertation submitted
in partial fulfilment of the necessary prerequisites
for the acquisition of the MSc Degree

Supervisor: Tzavalis Elias, Athens University of Economic and Business
Examiner 1: Pagratis Spyros, Athens University of Economic and Business
Examiner 2: Arvanitis Stylianos, Athens University of Economic and Business

Athens
[January, 2018]

ABSTRACT

A plethora of papers have been published on the subject of Efficient Market Hypothesis (EMH) since it was developed in the seventies. Some of them provide empirical and theoretical background that support the EMH, while others disprove the theory. In this thesis, we are not only revisiting the very definition of EMH but also, we are testing the theory against a new set of data.

To begin with, we are going to look into the foundation that was set by Samuelson (1965), Fama (1965) and Jensen (1968). Then, we will discuss the mathematical framework on which the theory was build and we will analyze the forms on EMH as well as the tools that one can use to test against the null hypothesis of the theory. After having established these basic concepts, we will review the history of the EMH and we will look into the anomalies that are proven to be the main weaknesses of the theory.

Besides the theoretical approach, we will also take a look into the EMH tests in depth. Then we are going to review our test subject – the Stockholm Stock Exchange. Lastly, we will attempt to apply those tests to recent data provided by the MSCI data base about the Swedish market. We will analyze the methodology behind the tests used and we will discuss our finding and how these results support of disagree with the EMH.

Contents

1. Introduction

The Efficient Market Hypothesis (EMH) is one of the most controversial aspects of financial theory. Its importance is great since it provides the basis for the development of various financial models. Thus, the subject of its validity is of equal importance.

According to the EMH, asset prices fully reflect all available information of the market under equilibrium. Such information may include past stock prices and current value of fundamental variables in addition to inside information of managers about dividend policy. Essentially, EMH states that stocks trade at fair value. As a result, it is impossible for an investor to earn abnormal returns consistently on a risk-adjusted basis.

The main force that drives this theory under equilibrium is the intense competition among investors and the ability to trade freely. When a new piece of information arrives in the market that offers added value to an asset, investors immediately adjust their actions by buying that asset which will push the assets price upwards. Thus, all available information is fully reflected on the assets price, disallowing excess profits from the trading of the asset under equilibrium.

It is information that EMH is built around, but information itself can include a plethora of things. For that reason, EMH has been divided into three forms, each one reflecting a different type of information. The stronger the form, the more difficult for EMH to hold. Despite the great contribution of the EMH in the literature, there is plenty of research that disproves the theory, especially in short term scenarios. This is a result of various anomalies that occur in the market, and these anomalies are the main reason of the theory not being accepted by various financial analysts.

The purpose of this thesis is developed into two parts. Firstly, it is to familiarize the reader with the EMH by presenting the theoretical background behind the theory. Secondly, it aims to present a case of empirical work which will provide an overview of how the Weak form of EMH is tested against the Swedish Stock Market.

Starting from Chapter 2, we will discuss the concept of EMH by reviewing the most important papers that established the theory. An in-depth look will be presented in the mathematical framework of the theory in an attempt to give reliability to the theory. Subsequently, the forms on EMH will be presented. We will look on what basis they are developed, what kind of information they can reflect and we will also mention some of the tests that the literature provides for each one of those forms. In Chapter 3, we will elaborate on the weak form of EMH as well as the three most popular methods of testing this form. Chapter 4 will provide an introduction to the Swedish Market Exchange while Chapter 5 includes the methodology which was used to perform the tests. Chapter 6 presents an accurate description of the data on which the tests were performed on and Chapter 7 will display the results of the tests. Lastly, on Chapter 8 the findings are discussed.

2. THE EFFICIENT MARKET HYPOTHESIS

2.1 Introduction

According to Fama (1970), generally the Efficient Market Hypothesis (EMH) is concerned with whether the price of markets “fully reflect” all available information at any point. It is assumed that all available information is commonly available to all investors. Thus, all investors can take advantage of the knowledge of information and adjust their trading patterns in order to achieve abnormal profits. As a result, any profitable opportunities will be perceived immediately and the market will be led to equilibrium.

The main condition for the EMH is the existence of a perfect capital market. In other words, there are no transaction costs in trading securities, all information is available with no additional costs to all market participants and all agree on the implications of current information to the assets price. In such a market, the price of an asset will fully reflect all available information.

But a frictionless market in which all information is freely available and investors always agree on the implications of information to the prices of the assets is not descriptive in practice. Fortunately, these conditions are sufficient but not necessary. The market may be efficient if sufficient number of investors have access to available information. Also, any disagreement among investors about the implication of information on assets prices does not itself imply market inefficiency unless there are investors who can consistently make better evaluations of available information that leads to consistent abnormal returns.

Even though transaction costs, availability of information and disagreement on implications among investors are not necessarily sources of market inefficiency, they certainly are potential sources of it. All three exists in the real world trading markets and thus measuring the effects of those factors in the process of price formation is the major goal of empirical work in this area.

The following chapter will present the theoretical and mathematical framework for the theory.

2.2 The Concept

Economists have always argued that prices are always right. This implies that asset prices already fully reflect all available information. More specifically, in the stock market prices are formed by interactions between supply and demand at every point in time. Essentially, economists believed that prices in the stock market are usually an unbiased result of fundamental factors. This idea was firstly formalized by Samuelson (1965) in his published paper “Proof that properly anticipated prices fluctuate randomly”. In this paper, Samuelson specified a model that defined the price behavior and by generalizing the martingale theorem he came to the conclusion that there can be a market that does not allow profit by extrapolation of post changes in futures price.

Working independently Fama (1970) further crystalized the idea. Eugene F. Fama had extensively researched the behavior of stock-market prices before he published his paper with title “Efficient Capital Markets: A Review of Theory and Empirical Work” where he proposed the EMH and broke down market efficiency into three forms based on the type of information related. Namely, the weak form in which the information subset of interest is past price histories, the semi strong form where the concern is the speed of price adjustment to other obviously publicly available information and finally the strong form in which the concern is whether investors or managers have monopolistic access to any information relevant to the price formation.

The proposition of EMH is that stock prices fully reflect all available information of the market about the value of firms or their investment risk. As a result, investors are unable to yield excess profits consistently.

A good description of market efficiency and the underlying mechanics that it implies is the one provided by Cootner (1964):

“If any substantial group of buyers thought prices were too low, their buying would force up the prices. The reverse would be true for sellers. Except for appreciation due to earnings retention, the conditional expectation of tomorrow’s price, given today’s price, is today’s price. In such a world, the only price changes that would occur are those that result from new information. Since there is no reason to expect that information to be non-random in appearance, the period-to-period price changes of a stock should be random movements, statistically independent of one another.”

2.3 Mathematical Framework

2.3.1 Expected Return or "Fair Game" Models

According to EMH, in an efficient market prices fully reflect all available information. This statement is general since it has no empirical implications. In order to make the hypothesis testable, we need to elaborate on how prices form and define what the term "fully reflects" implies in detail.

One option would be to assume that the prices are generated mainly by a two-parameter model as used by Sharpe (1964) and Lintner (1965). In his paper, Sharpe constructs a market equilibrium theory of asset prices under the condition of risk and provides evidence consistent with the implications and results of theories that the literature provided up until then. Lintner, using also a two-parameter model develops rigorously some of the fundamental implications of uncertainty for which there was a lot of confusion in the literature up until then. Nevertheless, the theoretical models and empirical tests for market efficiency have not been this specific. Up until now, such models have been based on the assumption that the conditions of the market can somehow be expressed in in terms of expected returns. Such models would imply that conditional on some information set, the equilibrium expected return would be a function of its risk, in the same way as in the two-parameter model. All models of the expected return theories can however be described as follows:

$$E\left(\tilde{p}_{j,t+1} \middle| \Phi_t\right) = \left[1 + E\left(\tilde{r}_{j,t+1} \middle| \Phi_t\right)\right] p_{jt} \qquad (1)$$

Where E is the expected value operator, p_{jt} is the price of asset j at time t and $p_{j,t+1}$ is its price at time t+1 when also taking into consideration the reinvestment of intermediate cash income from the asset. In addition, $r_{j,t+1}$ is the one period percentage return defined as follows:

$$r_{j,t+1} = \frac{p_{j,t+1} - p_{j,t}}{p_{j,t}}$$

Furthermore, Φ is the general symbol for the information set that we assume to be fully reflected in the price at time t, and the tildes indicate that $p_{j,t+1}$ and $t_{j,t+1}$ are random variables at time t.

The value of the expected return $E(\tilde{p}_{j,t+1}|\Phi_t)$ is determined from the particular expected return theory assumed to apply. Nevertheless, the expression (1) is still important as it implies that whatever the expected return theory at hand, the information set Φ_t is fully utilized in the determination of the expected returns under equilibrium. Thus, information Φ_t is fully reflected in the price p_{jt} of the asset.

However, we must take into consideration the fact that the expected value used is just one of many possible summary measures of a distribution of returns. The general notion that prices fully reflect available information does not provide any importance to the expected value. Thus, the result of the tests that we will discuss -that are based on this assumption- depend to some extent on the validity of this model. Of course, in order to give empirical content to the market efficiency theory, some assumptions are necessary. The assumption (1), i.e. market equilibrium can be stated in terms of expected returns as well as equilibrium price is formed on the basis of the information set Φ_t have a major empirical implication. According to this assumption, market based only on information Φ_t cannot materialize profits or returns excess of equilibrium expected profits or returns. Let:

$$x_{j,t+1} = p_{j,t+1} - E\left(p_{j,t+1}\middle|\Phi_t\right) \qquad (2)$$

then

$$E\left(\tilde{x}_{j,t+1}\middle|\Phi_t\right) = 0 \qquad (3)$$

Which by definition states that the sequence of $\{x_{jt}\}$ is a fair game with respect to the information sequence $\{\Phi_t\}$. Equivalently, let

$$z_{j,t+1} = r_{j,t+1} - E\left(\tilde{r}_{j,t+1}\middle|\Phi_t\right) \qquad (4)$$

then

$E(\tilde{z}_{j,t+1}|\Phi_t) = 0$ (5)

so that the sequence $\{z_{jt}\}$, similar to the sequence $\{x_{jt}\}$, is also a fair game with respect to the information sequence $\{\Phi\}$.

In economic terms, $x_{j,t+1}$ is the excess market value of security j at time t+1. In essence, it is the difference between the observed price and the expected value of the price based on the information set Φ_t. Similarly, $z_{j,t+1}$ is the excess market return at time t. Let:

$$a(\Phi_t) = [a_1(\Phi_t), a_2(\Phi_t), \dots, a_n(\Phi_t)]$$

be any trading system based on the information set Φ_t which tells the investors the amounts $a_j(\Phi_t)$ of funds available at time t that are to be invested in each of the n available securities. The total excess market value at t+1 that will be generated by such a system is

$V_{t+1} = \sum_{j=1}^{n} a_j\,(\Phi_t)[r_{j,t+1} - E(\tilde{r}_{j,t+1}|\Phi_t)]$,

Which has the following expectation from the fair game property of (5):

$$E(\tilde{V}_{t+1}|\Phi_t) = \sum_{j=1}^{n} a_j\,(\Phi_t) E(\tilde{z}_{j,t+1}|\Phi_t) = 0$$

Subsequently, we shall discuss two special cases of the model, the sub martingale and the random walk model as they are of great importance to the empirical literature.

2.3.2 The Sub Martingale Models

Going back to the expression (1), assume that for all t and Φ_t

$E(\tilde{p}_{j,t+1}|\Phi_t) \geq p_{jt}$ or equivalently, $E(\tilde{r}_{j,t+1}|\Phi_t) \geq 0$. (6)

This statement implies that a security j has a price sequence $\{\Phi_t\}$ that follows a sub-martingale with respect to the information sequence $\{\Phi_t\}$ which essentially says that the expected value of next period's price -based on today's available information set- is equal or greater than the current price. For the price sequence to follow a martingale, expression (6) must hold as an equality.

There is one important empirical implication that derives from a sub martingale in prices. Assume a set of trading rules for a trader that has an asset and cash under his possession. This set of trading rules defines when the trader will hold the asset, sell it short or just hold cash at any time. In this case, the assumption (6) implies that this trader would not be able to achieve profits greater than those offered from a buy-and-hold strategy for the future period in question. Such rules have been tested thoroughly in the literature, providing evidence both proving or disproving the EMH.

2.3.3. The Random Walk Models

In the early literature, efficient market would often be amalgamated with the random walk model. The term "fully reflects" would imply that successive price changes are independent and identically distributed. Formally, the model says:

$$f(r_{j,t+1}|\Phi_t) = f(r_{j,t+1}) \qquad (7)$$

Expression (7) states that the conditional and marginal probability distributions of an independent random variable are identical while the density function is the same for all t. If we return to the expression (1) and compare it with (7), we will notice that expression (7) offers some additional information in comparison to the return model summarized by (1). In more detail, let the expected return on security j be constant over time, then we have:

$$E(\tilde{P}_{j,t+1}|\Phi_t) = [1 + E(\tilde{r}_{j,t+1}|\Phi_t)]p_{jt}$$

$$E(\tilde{p}_{j,t+1}|\Phi_t) = p_{jt} + E(\tilde{r}_{j,t+1}|\Phi_t)P_{jt}$$

$$\frac{E(\tilde{p}_{j,t+1}|\Phi_t) - p_{jt}}{p_{jt}} = E(\tilde{r}_{j,t+1}|\Phi_t)$$

$$E(\tilde{r}_{j,t+1}) = E(\tilde{r}_{j,t+1}|\Phi_t) \qquad (8)$$

In other words, expression (8) implies that $r_{j,t+1}$ is independent of the information available at t, Φ_t. Expression (7) implies that the whole distribution of $r_{j,t+1}$ is independent of Φ_t. Fama (1970) argues that it is best to regard the random walk model as an extension of the general expected return or fair game efficient market model. Since a fair game model only states that the equilibrium in the market can be expressed in terms of expected returns, little is said about the detailed stochastic process that generates the returns. A random walk model appears as a sub-category of the fair game model when aspects such as preferences of traders in conjunction with the flow of new information lead to an equilibrium where the distributions of returns repeat themselves through time.

As a result, empirical tests on random walk models that are in fact tests of a fair game are more strongly in favor of the EMH.

2.4 Forms

2.4.1 Weak form of EMH

As we mentioned earlier, there are three forms of EMH: weak, semi-strong and strong. The weak form of the EMH incorporate the cases where asset prices already reflect all past publicly available information. This information is common for all investors. Under the Weak form of EMH, the stock prices change only with the arrival of new information in the market. As a result, it would be impossible for traders to earn higher profits from an investment than those imposed by the Expected Returns Model.

On the other hand, many financial analysts resort to past stock prices and trading volume data to forecast future prices and achieve above average profits. Of course, this comes to a clear opposition with the EMH theory, which essentially states that consistent successful prediction of stock prices would be impossible.

As far as the weak form of EMH is concerned, the empirical evidence is undeniable since a plethora of methods have been developed and applied to test it. Some of the empirical tests we will discuss in a later chapter.

2.4.2 Semi-Strong Form of EMH

The semi-strong form of the EMH suggests that the stock prices fully reflect all available information, including both past history of stock prices and information about the market or the company at hand at time t. As Fama (1970) states, semi strong form is concerned with information obviously publicly available. Such information is considered to be stock splits, announcements of financial reports by firms, new security issues, earnings and dividends announcements, macroeconomic factors, announced merger plans, etc. In fact, “obviously publicly available information” may include information not strictly of financial nature. For example, information about the current state of research relevant with the stocks at hand would also be included and taken into consideration.

As a result, each test only brings supporting evidence for the model. Gathering such information has been proven to be a difficult task, since it is costly and time consuming. It also demands the cooperation of specialists who can interpret the effects and gravity of such information and are able to transform it into usable data. Therefore, it may not be sufficient to simply gather information, but an extensive research has to be conducted on the implications of the information as well.

Thus, testing the semi strong form of EMH is a challenging task, especially when compared to the weak form of EMH. In addition, in comparison to the weak form of EMH, the semi strong form assumes that apart from the risk investment, the stock prices should include the effort put by the specialists and analysts which is also adding to the cost. Of course, this cost would be minimal and common for all stocks because of the large number of companies analyzing these events systematically and uniformly. Thus, once again the stock prices would reflect all available information included in the semi strong form of EMH, and prices would only change with the arrival of new information that correspond so such events as mentioned before.

In essence, the semi strong form still implies that traders should not be able to realize abnormal profit using information such as past history of prices, stock splits, research and development announcements etc. This kind of information would already be reflected in the current price and traders would not be able to take advantage of it. On the other hand, other kind of information can be used to achieve excess returns. Inside information about the future plans of a company can still be available to managers and thus they may be able to abuse this information to gain an advantage over the market. This brings us to the third form of EMH, the strong form.

2.4.3 Strong form of EMH

The last form of EMH is the strong form, and is the strictest of the three. In this form, prices reflect all information, both public and private. Public information refers to all kinds of information mentioned so far, such as past history prices and financial and non-financial events. Private information would include inside information that traders are able to gather from company managers about nonpublic details of companies that could provide an edge when invested upon. Thus, the strong form of EMH states that traders would not be able to systematically earn excess profits by taking advantage of such information. Even if arbitrage opportunities arise, they will quickly be eliminated. Once a trader has inside information and acts on it, the market recognizes the change in the trading volume and adjusts in the existence of the profitable occasion. In essence, other traders will recognize the profitable opportunity and will attempt to take advantage of it too. As a result, the price of the asset at hand will adjust quickly as the inside information enters the common information set. Thus, the price of the asset will eventually once again reflect all information, both public and private.

The main difference between the semi strong and the strong form, is that in the first case, as we mentioned, traders are able to take advantage of inside information, while in the second, they are not. In essence, in the strong form of EMH, the market anticipates in an unbiased manner all future changes in assets price. Therefore, the unexpected changes themselves expose possible inside information downgrading the opportunities for excess profit.

Nevertheless, empirical research has displayed evidence that disprove the EMH in its strong form.

2.5 Historical Background

In this section, the history of EMH will be presented. Even though the term “efficient” became popular at the beginning of the 20th century, earlier literature offered a solid background for the conception and development of the EMH.

Martin Sewell (2011) with his research note (RN/11/04) provides us with an extensive background about the roots and the evolution of the term.

The concept of a “fair game”, which is essential for the EMH, can be located back in literature of the 16th century, when the Italian mathematician Girolamo Cardano in his book “The book of games of chance” mentions:

“The most fundamental principle of all in gambling is simply equal conditions, e.g. of opponents, of bystanders, of money, of situation, of the dice box, and of the die itself. To the extent to which you depart from that equality, if it is in your opponent’s favor, you are a fool, and if in your own, you are unjust.”

Even though is it not a mathematical expression, it sets the basic idea of a fair game.

In 1863, the French stockbroker Jules Regnault noticed that the longer you hold a security the more exposed you are to its price variations. In other words, the more you hold a security, the more you can win or lose – the price deviation is directly proportional to the square root of time (Regnault, 1863). Later in the 19^{th} century Lord Rayleigh, a British physicist, was already familiar with the notion of a random walk even though his work was on sound vibrations (Rayleigh, 1880). Already a few years later, a British logician and philosopher by the name of John Venn had a clear concept of both random walk and Brownian motion (Venn, 1888). The following year, George Gibson in his book The Stock Markets of London, Paris and New York clearly mentions the concept of efficient markets, and the first steps towards the EMH are made.

In 1900, Louis Bachelier in his PhD thesis “Theorie de la Speculation” (Bachelier, 1900), he developed the mathematics and statistics of Brownian motion five years before Einstein (Einstein, 1905). In in the same paper (Bachelier, 1900), Bachelier also came to the conclusion that the mathematical expectation of the speculator is zero, 65 years before Samuelson explained in depth efficient markets in terms of a martingale (Samuelson, 1965). Bechelier’s work was way ahead of his time but unfortunately it remained in the dark until it was rediscovered by Savage in 1955 (Savage, 1955). Just five years later, the term random walk was developed and published by Karl Person (Karl Person, 1905). The same year, unaware of Bechelier’s work, Albert Einstein developed the equations for Brownian Motion (Einstein, 1905). The following year, a Polish scientist by the name of Marian Smoluchowski attempted to describe Brownian Motion (von Smoluchowski, 1906) while two years later the stochastic differential equation of Brownian motion was developed by Langevin (Langevin, 1908).

Few years later, “The Laws of Supply and Demand” was published by Diblee (Diblee, 1912) while a couple of years after that, Bachelier published “The Game, The Chance and the Hazard” (Bachelier, 1914). In 1921, Taussig published a paper titled "Is market price determinate?" (Taussig, 1921) while in 1923 an English economist by the name of John Maynard Keynes stated that investors on financial markets are rewards for risk baring instead of accurately predicting future price developments (Keynes, 1923), which is a direct consequence of the EMH. In 1925, the economist Frederick MacCauley noticed that the fluctuations of the stock market as well as the fluctuations of a chance curve that may be obtained by a dice are surprisingly similar (MacCauley, 1925). The following year, Maurice Olivier gave solid data of the leptokurtic nature of the distribution of returns in his doctoral dissertation (Olivier, 1926), while in 1927 Frederick Mills was the one to actually provide a proof for the leptokurtic nature of returns (Mills, 1927). Two years later, in 1929 the Wall Street Crash occurred, which is considered to be one of the most if not the most devastating stock market crash in the history of US.

In 1930, an American economists and businessman by the name of Alfred Cowles founded and funded the Economic Society and its journal Econometrica, while two years later he set up the Cowles Commission for Economic Research. Cowles also analyzed the performance of investment professionals and concluded that stock market forecasters cannot forecast. In 1934, Holbrook Working came to the conclusion that stock returns behave like numbers from a lottery (Working, 1934), supporting the conclusion of Cowles few years ago. In 1936, Keynes published the "General Theory of Employment, Interest, and Money" (Keynes, 1936). The following year, Slutzky showed that sums of independent random variables can be the source of cyclic processes (Slutzky, 1937).

In 1944, following the paper he published in 1933, Cowles provided more reports that showed that investment professionals do not consistently beat the market (Cowles, 1933). In 1949, Working finally showed that in an ideal futures market, any professional forecaster would be unable to predict price changes successfully consistently (Working, 1949).

In 1953, Milton Friedman showed that even in cases where the behavior or the trading strategies of investors are correlated, EMH can still apply due to arbitrage (Friedman, 1953). The same year, Kendall analyzed 22 price-series at weekly frequency and found out that they were essentially random (Kendall, 1953). In addition, Kendall was one of the first to take notice of the time dependence of the empirical variance (nonstationarity). In 1956, Bachelier's work returns on the foreground after being rediscovered by Savage, as an acknowledged forerunner in the thesis of Samuelson, in an options-like pricing model (Samuelson, 1956). In 1958, Working built an anticipatory market model while the following year Harry Roberts presented the similarities between a random walk and an actual stock series (Harry, 1959). During the same year, Osborne not only managed to show that the logarithm of common-stock prices follows Brownian motion but he also discovered evidence of the square root of time rule (Osborne, 1959).

In 1960, Larson presented his findings considering a new method of time series analysis where he notes that the distribution of price changes is "very nearly normally distributed for the central 80 percent of the data, but there is an excessive number of extreme values." In 1960, Cowles revisited the paper he published in cooperation with Jones back in 1937, correcting an error but still coming to the same conclusion of mixed temporal dependence results (Cowles, 1960). In the same year, Working showed that if use of averages is made, autocorrelation may arise even if it was not present in the original series (Working, 1960).

The following year, Houthakker used a stop-loss sell orders and found not only patterns, but also stumbled upon leptokurtosis, non-stationarity and suspected non-linearity (Houthakker, 1961). Working independently of Working, Alexander showed that spurious autocorrelation cannot be introduced by averaging if the probability of a rise is not 0.5. According to him, the random walk model has the best fit for the data, but there are also signs of leptokurtosis in the distribution of returns (Alexander, 1961). In addition, the paper was of great interest because it was the first to test for non-linear dependence. In the same year, John F. Muth introduced the rational expectations hypothesis in economics (Muth, 1961).

In 1962, Mandelbrot was the first to propose that the tails of the distribution of returns follow a power of law (Mandelbrot, 1962), while in the same year Paul H. Cootner came to the conclusion, in contrast with previous reports, that the stock market is not a random walk (Cootner, 1962). Osbourne, continuing his research on deviations of stock prices from a simple random walk, noticed that stocks tend to be traded in concentrated bursts. Also in 1962, Jack Treynor wrote an unpublished paper titled “Towards a theory of market value of risky assets, which was the first paper on the Capital Asset Pricing Model (Treynor, 1962).

In 1963, Granger and Morgenstern performed spectral analysis on market prices and found that short-run movements of the series obey the simple random walk hypothesis, in contrast to long-run movements which do not Granger & Morgenstern, 1963). In addition, the same year Mandelbrot formed and tested a new model of price behavior, which (unlike Bechelier’s) used natural logarithms of prices and also replaced the Gaussian distributions with the more general stable Paretian (Mandelbrot, 1963).

In 1964, Alexander faced the critics of the paper that he published in 1961, coming to the conclusion that the subject of his study, the S&P index does not follow a random walk. Godfrey et al. (1964) published “The random walk hypothesis of stock market behavior” while Steiger (1964) realized tests for non-randomness and concluded that stock prices do not follow a random walk. Sharpe (1964) published his Nobel prize-winning work on the CAPM.

In 1965, Fama in his paper “The Behavior of Stock-Market Prices” defined an “efficient market” and concluded that stock market prices do follow a random walk (Fama, 1965). Around the same time, Samuelson (1965) provided the first economic argument for “efficient markets”.

The following year, Fama continued his contribution to the subject. In cooperation with Blume, they published a paper that came to the conclusion that serial correlation is probably as powerful as the Alexanders filter rules (Alexander, 1961, 1964) for measuring the direction and degree of dependence in price changes (Fama & Blume, 1966). In the same year Mandelbrot (1966) proved some of the first theorems showing how, in a competitive market with rational risk-neutral investors, security values and prices follow a martingale i.e. returns are unpredictable.

Harry Roberts (1967) sided with the term Efficient Market Hypothesis and divided it into weak and strong form tests, a division which later became generally accepted and frequently used.

In 1968, Michael C. Jensen evaluated the performance of mutual funds and concluded that "on average the funds apparently were not quite successful enough in their trading activities to recoup even their brokerage expenses" (Jensen, 1968). In the same year the first ever event study was published by Ball and Brown (1968).

Fama et al. (1969), by publishing their event study in a paper titled "The Adjustment of Stock Prices to New Information", offered considerable support to the idea of efficiency in the stock market.

A milestone in the evolution of Efficient Market Hypothesis was set by the paper of Fama (1970) named "Efficient Capital Markets: A Review of Theory and Empirical Work". In this paper, he defined the efficient market as well as the three forms of the tests of the hypothesis: weak, semi-strong and strong. The same year, Granger and Morgenstern (1970) published the book "Predictability of Stock Market Prices".

In 1972, the price effects if secondary offerings were studied by Scholes (1972) where the findings supported the EMH in all but a few cases where a post-event price drift exists.

In 1973, Samuelson published the paper "Mathematics of speculative price" and LeRoy (1973) proved that under risk-aversion, there is no theoretical justification for the martingale property. The book "The Stock Market: Theories and Evidence" was also published in the same year by Lorie and Hamilton (1973) as well as the book "A Random Walk Down Wall Street" by Malkiel (1973). In addition, Samuelson (1973) included stocks that pay dividends to his earlier work (1965).

"The valuation of options for alternative stochastic processes" was published by Cox and Ross (1976). Sanford Grossman described a model which shows that "informationally efficient price systems aggregate diverse information perfectly, but in doing this the price system eliminates the private incentive for collecting the information" (Grossman, 1976). Fama (1976) also published a book under the title "Foundations of Finance".

In 1977, Osborne published a collection of lecture notes, where he discussed market-marking, random walks, statistical methods and sequential analysis of stock market data under the title "The Stock Market and Finance from a Physicist's Viewpoint" (Osborne, 1977). Beja (1977) provided evidence that a real market can never be efficient.

In 1978, Ball published a paper which provided data supporting that excess returns after public announcements of firms' earnings are common (Ball, 1978). Also in 1978, Jensen expressed his belief that "…there is no other proposition in economics which has more solid empirical evidence supporting it than the Efficient Market Hypothesis.". In addition, Jensen (1978) defines efficiency as: "A market is efficient with respect to information set θ_t if it is impossible to make economic profits by trading on the basis of information set θ_t. A theoretical model of rational agents which shows that the martingale property is not necessary to hold under risk aversion was published by Lucas (1978).

Radner (1979) by taking advantage of a theoretical model of asset trading showed that if the number of alternative states of initial information is finite, then an equilibrium of rational expectations exists that reveal to all traders all of their initial information. During the same year, Dimson (1979) revisited the problems of risk measurement when shares are subject to infrequent trading. Also, Harrison and Kreps (1979) published “Martingales and arbitrage in multiperiod securities markets” while Shiller (1979) showed that the volatility of long-term interest rates is greater than predicted by expectations models.

In 1981, it was showed by Grossman and Stiglitz (1981) that is it impossible for a market to be perfectly informationally efficient. Prices cannot perfectly reflect the available information because information is costly. If it did, investors who spent resources on obtaining and analyzing information would receive no compensation. Thus, a sensible model of market equilibrium must leave some incentive for information0gathering (security analysis).

The following year, LeRoy and Porter (1981) showed that stock markets exhibit “excess volatility” and they reject market efficiency while Stiglitz (1981) showed that even with apparently competitive and efficient markets, resource allocations may not be Pareto efficient. In addition, Shiller (1981) showed that stock prices move too much to be justified by subsequent changes in dividends, i.e. exhibit excess volatility.

Milgrom and Stokey (1982) showed that the receipt of private information cannot create any incentives to trade under certain conditions. The same year, Tirole (1982) showed that unless traders have different priors or are able to obtain insurance in the market, speculation relies on inconsistent plants, and thus is ruled out by rational expectations.

In 1984, evidence of the square root of time in earnings was found by Osborne and Murphy (1984). Also in 1984, Roll publishes a paper in which he examines US orange juice futures prices and the effect of the weather, finding excess volatility (Roll, 1984).

In 1985, De Bondt and Thaler (1985) discovered that stock prices overreact, evidencing substantial weak form market inefficiencies, setting the basis for a new financial field called Behavioral Finance which seeks to combine behavioral and cognitive psychological theory with conventional economics and finance in order to provide explanations for why people make irrational financial decisions.

Marsh and Merton (1986) analyzed the variance-bound methodology used by Shiller and came to the conclusion that this approach cannot be used to test the hypothesis of stock market rationality. They also pointed out the practical consequences of rejecting the EMH. During the same year, Fischer Black (1986) introduced the concept of “noise traders”. According to Black, noise traders are those who trade based on anything but information. He also showed that noise traders are essential to the existence of liquid markets. Furthermore, Lawrence H. Summers (1986) in his paper “Assessing Dynamic Efficiency: Theory and Evidence” argued that many statistical tests of market efficiency have very little power in discriminating against plausible forms of inefficiency. French and Roll (1986) found that during exchange trading hours, asset prices are much more volatile rather than during non-trading hours. Based on that fact they speculated that this may occur due to the fact that individuals trade on private information – the market generates its own news.

In 1987 the infamous Black Monday occurred, where stock markets around the world crashed. The crash began in Hong Kong and spread to the west, hitting Europe. The United States stock market followed after other markets had already declined by a significant margin. Dow Jones Industrial Average index recorded its most devastating drop which reached a magnitude of 22.61%.

Negative autocorrelations for stock portfolio return horizons beyond a year were found by Fama and French (1988). In the same year, Lo and Mackinlay (1988) applied the variance-ratio test on weekly stock market returns data and came to conclusion of rejecting the random walk hypothesis. Another worth mentioning paper is the one published by Poterba and Summers (1988) in which the two economists showed that stock returns are characterized by autocorrelation over short periods and negative autocorrelation over longer horizons. Conrad and Kaul (1988) in their paper “Time-Variation in Expected Returns” took an elaborate look on stochastic behavior of expected returns on common stock.

Cutler et al. (1989) found that market movement cannot adequately be explain by the flow of new information alone. Eun and Shim (1989) found national stock markets are characterized by a substantial amount of interdependence, and their results are consistent with informationally efficient international stock markets. Ball (1989) focuses on the definition of stock market efficiency. The book “A Reappraisal of the Efficiency of Financial Markets” was published by Guimaraes (1989). Shiller (1989) published “Market Volatility”, a book about the sources of volatility which challenges the EMH. Also, LeRoy (1989) on his paper “Efficient capital markets and martingales”, elaborates on the transition between the intuitive idea of market efficiency and the martingale is far from direct.

The following year, Laffont and Maskin (1990) showed that the EMH may fail if the market is characterized by imperfect competition. In addition, Lehmann (1990) found reversals in weekly returns which led to rejection of the efficient market hypothesis.

Kim et al. (1991) re-examined the empirical evidence for mean-reverting behavior in stock prices and found that mean reversion is a phenomenon that occurred entirely before World War II.

In 1992, Chopra et al. found that stocks overreact (Chopra et al., 1992). Bekaert and Hodrick (1992) characterized predictable components in excess returns on equity and foreign markets. Also, the book "Capital Ideas" was published in the same year by Peter L. Bernstein which was a critic review if the history of the ideas that shaped modern finance (Bernstein, 1992). Malkiel (1992) also published an essay on the subject of Efficient Market Hypothesis in the New Palgrave Dictionary of Money and Finance.

Jegadeesh and Titman (1993) found that trading strategies that bought past winners and sold past losers realized significant abnormal returns. The same year Richardson showed that the patterns in serial-correlation estimates and their magnitude observed in previous studies should be expected under the null hypothesis of serial independence (Richardson, 1993).

In 1994, Roll observed that it is hard to profit from even the strongest market inefficiencies (Roll, 1994). Huang and Stoll (1994) provided new evidence concerning market microstructure and stock returns predictions. Also, according to the paper published by Metcalf and Malkiel (1994), portfolios of stocks formed by experts do not consistently beat the market in terms of returns. Evidence was provided by Lakonishok et al. (1994) supporting that value strategies yield higher returns because these strategies exploit the suboptimal behavior of the typical investor and not because these strategies are fundamentally riskier.

The following year, Robert Haugen published the book "The New Finance: The Case Against Efficient Markets" in which the focuses on long-term reversals that may be a result of a short-run overreaction when the market recognizes its past error (Haugen, 1995).

The book "The Econometrics of Financial Markets" was published by Campbell et al. (1996). Furthermore, Chat et al. (1996) took a more elaborate look at momentum strategies and their results suggest a market that responds only gradually to new information.

In 1997, Andrew Lo edited two volumes that bring together the most influential articles on the EMH (Lo, 1997). Chan et at. (1997) came to the conclusion that the world equity markets are weak-form efficient. Dow and Gorton (1997) investigated the connection between stock market efficiency and economic efficiency. W. Brian Arthur et al. proposed a theory asset pricing by creating an artificial stock market with heterogeneous agents with endogenous expectations (Arthur et al., 1997).

A year later, Elroy Dimson and Massoud Mussavian provided a brief history of Market efficiency in their paper Dimson and Mussavian (1998). In his third of three reviews, Fama (1998) concluded that “market efficiency survives the challenge from the literature on the long-term return anomalies”.

In 1999, Lo and MacKinlay published the book titled “A Non-Random Walk Down Wall Street” (Lo & MacKinlay, 1999). In the same year, Haugen (1999) published the second edition of his book, which makes the case for the inefficient market, positioning the efficient market paradigm at the extreme end of a spectrum of possible states. Bernstein (1999) criticized the Efficient Market Hypothesis and claims that the marginal benefits of investors acting on information exceed the marginal costs. Zhang (1999) presented a theory of marginally efficient markets.

In 2000, Shleifer published a book under the title “Inefficient Markets: An Introduction to Behavioral Finance”, which questions the assumptions of investor rationality and perfect arbitrage. In addition, Lo (2000) published a selective survey of finance while Beechey et al. (2000) published a survey paper on the Efficient Market Hypothesis. Shiller (2000) also published the first edition of his book “Irrational Exuberance” in which he challengers the Efficient Market Hypothesis, demonstrating that markets cannot be explained historically by the movement of company earnings or dividends.

Mark Rubinstein revisited some of the most serious historical evidence against market rationality and after a more elaborated look he concluded that markets are rational (Rubinstein, 2001). Furthermore, Shafer and Vovk (2001) published “Probability and Finance: It’s Only a Game!” which shows how probability can be based on game theory and then applies the same framework to finance.

In 2002, Lewellen and Shanken published a paper titled “Learning, Asset-Pricing Tests, and Market Efficiency” in which they concluded that parameter uncertainty can be important for characterizing and testing market efficiency (Lewellen & Shanken, 2002). Chen and Yeh (2002) investigated the emergent properties of artificial stock markets and showed that the Efficient Market Hypothesis can be satisfied with some portions of the artificial times series.

In 2003, Malkiel examined the attacks on the Efficient Market Hypothesis and concludes that stock markets are far more efficient and far less predictable than some recent academic papers would suggest (Malkiel, 2003). In addition, G. William Schwert showed that when anomalies are published, practitioners implement strategies implied by the papers and the anomalies subsequently weaken or disappear. In other words, literature helps evolve the market and increases the efficiency within it (Schwert, 2003). Also, Haugen (2003) published the third edition of his book, in which he focuses on the evidence, causes and history of overreactive pricing in the stock market.

In 2004, Timmermann and Granger discussed the Efficient Market Hypothesis from the perspective of a modern forecasting approach.

Malkiel (2005) showed that professional investment managers do not outperform their index benchmarks and provides evidence that by and large market prices do seem to reflect all available information.

The following year, Blakey (2005) looked at some of the causes and consequences of random prices behavior, meanwhile Toth and Kertesz (2006) found evidence of continuously increasing efficiency in the New York Stock Market.

In 2007, Wilson and Marashdeh (2007) demonstrated that cointegrated stock prices are inconsistent with the EMH in the short run, but consistent with the EMH in the long run. The Elimination of arbitrage opportunities means that stock market inefficiency in the short run ensures stock market efficiency in the long run.

McCauley et al. (2008) showed that martingale stochastic processes generate uncorrelated, generally non-stationary increments. In addition, in the same paper he explains why martingales look Markovian at the level of both simple averages and 2-point correlations, as well as proves that arbitrary martingales are topologically inequivalent to Wiener processes. Andrew Lo wrote the "Efficient Market Hypothesis" article for the second edition of "The New Palgrave Dictionary of Economics" (Lo, 2008). Yen and Lee (2008) presented a survey article that gives a chronological account of empirical findings and conclude that the Efficient Market Hypothesis is here to stay.

Focused on the global financial crisis, Ball (2009) argued that the collapse of Lehman Brothers and other large finance institutions, is rooted on the lack of respect of the efficient market theory.

Lee et al. (2010) investigated the stationarity of real stock prices for 32 developed and 26 developing countries covering the period between January 1999 and May 2007 and conclude that conclude that stock markets are not efficient.

2.6 Anomalies

Earning excess returns without taking on more risk than the market is impossible according to the Efficient Market Hypothesis. Even though plethora of papers and a lot of testing has provided support for the EMH, it would be shortsighted to say that there is no research with results that disprove the theory.

Even though data from mutual-funds seem to be resistant with the weak form of Efficient Market Hypothesis, there are a lot of market anomalies or deviations from the expected behavior that have been observed in the financial markets and complicate the picture severely.

That does not directly state that EMH has failed, but it is important to note that EMH may not apply in all cases because not all cases meet its requirements. In many cases, the behavior of traders does not appear to be rational.

In this chapter we will discuss some of those cases and we will attempt to make clear when and why those cases occur.

2.6.1 The Calendar Effect

It has been observed that there are fluctuations in the stock prices which are directly connected to the calendar days, months or seasons. These cases are known in the literature as calendar effects. In fact, a calendar effect can me any market anomaly which appears to be related to the calendar, and such effects include the otherwise unreasonable change in behavior of traders in different days of the week, different times of the month, and different times of the year. These effects are also known as seasonal tendencies.

A lot of the empirical research focused on such seasonal tendencies have shown that, when these seasonal tendencies are taken advantage of, it is possible for traders to benefit from these predictable changes of assets price movement and materialize excess profit from them.

Some of the calendar effects that can be found in the literature are: the January effect, the weekend effect, the turn-of-the month effect and the seasonal effect. We will focus mainly on the January and weekend effect.

2.6.1.1 The January Effect

The most common anomaly which has been observed in financial markets is the January effect – a seasonal anomaly in financial markets where securities price tends to increase more at the first month of every year than in any other month. If such an event is true, then it would suggest that the market is not efficient. Traders would be able to buy stocks at the end of December and sell them during January, when the prices have increased. Thus, traders that could adopt such strategy will be able to profit consistently from it, directly opposing what market efficiency dictates.

The most common theory explaining this phenomenon is connected to the income tax. Investors -who are income tax sensitive- hold small stocks for tax reasons, sell them at the end of the year in order to claim capital loss, and then they reinvest after the first of the year. Thus, the January effect occurs. Nevertheless, the January Effect does not materialize every year. According to Siegel (1994) small stocks underperformed last stocks during some years, for example: January 1982, 1987, 1990. Similar opinion is being held by Ciccone (2013) in his article "January's Stock Temptation".

Another strategy that appears to enforce he January Effect is the so called "Window Dressing". Window dressing is a strategy used by mutual fund and other portfolio managers near the end of the year to improve the appearance of a fund's performance before presenting it to clients and shareholders. In order to "Window Dress", the managers sell stocks of small magnitude that may have high volatility, replacing them with stocks that provide stable short-term gains and thus these securities are reported as part of the fund's holdings. This way, the performance of the fund for the reporting period appears to be better than it actually is.

2.6.1.2 The Weekend Effect

Empirical researchers have noticed that it is not uncommon for stocks to have an opening price on Monday that is lower than their closing price on the preceded Friday. Fama (1980) also conducted some research on the subject using the S&P index. His findings do imply the existence of the weekend effect as he states “although the average return for the other four days of the week was positive, the average return for Monday was significantly negative during each of five five-year subperiods”.

The source that causes the weekend effect is still unclear but many theories attempted to uncover the reasoning behind this phenomenon. Some theories attribute this effect to the fact that companies tend to announce unfavorable news on Friday after the market closes resulting in depressed stock prices on Monday. Other theories attempt to explain it by taking into consideration short selling strategies that could affect stocks with high short interest positions during the weekend – when the market is closed. The weekend effect could simply be a result of trader’s fading optimism during the weekend. Research has shown that the weekend effect cannot be explained by reasonings such as differences in settlement periods for transactions occurring on different weekdays, measurement errors in recorded prices, market maker trading activity or even systematic patterns in investors buying and selling behavior. Even though the actual explanation may be elusive, the weekend effect has all but dissipated (Schwert, 2002).

2.6.2 Value, Size and Momentum Anomalies

Of course, there are many “anomalies” that are generally accepted by the literature.

2.6.2.1 The Value Effect

The value effects include the positive character of the relation between security returns and the ratio of accounting-based measures of cash flow or value to the market price of security. Investment strategies based on the value effect have a long tradition in finance (Graham and Dodd 1940). Nevertheless, some researchers argue that variables like the Earnings-to-Price (E/P) are nothing but proxies for expected returns (Ball 1978). As a result, if the CAPM is an incomplete specification of priced risk, it is reasonable to expect that E/P might explain the portion of expected return that is compensation for risk variables omitted from the tests.

One of the first to test the notion of value-related variables explaining violations of the CAPM was Base (1977). He was led to the conclusion that there was a significant positive relation between E/P ratios and average returns for U.S. which could not be explained by the CAPM. This result was verified by Reinganum (1981). Among other researchers, DeBondt and Thaler (1987) documented strong positive relation between returns and Book-to-Price (B/P). Researchers have also identified a significant relation between security returns and value ratios that use cash flow in place of earnings in the numerator of the ratio. All the forms of the value effects have been reproduced by many researchers for many different sample periods and for most major securities markets around the world.

Similar to the value ratios, dividend yield -the ratio of cash dividend to price- has also been shown to have cross-sectional return predictability. Even though a positive relation has been identified between stock returns and dividend yields, interpretation of the results has been controversial. Evidence on the subject of this relation has been provided by Litzenberger and Ramaswamy (1979) as well as Miller and Scholes (1989) and many others.

2.6.2.2 The Size Effect

Research shows that the relevant size of a firm also tends to yield abnormal behavior in terms of returns. “The Size Effect” refers to the negative relation between security returns and the market value of the common equity of a firm. According to Fama and French (2011) stocks with market capitalization in the smallest 30% of companies in the data set studied outperformed those with market capitalization in the largest 30% in the data set. Small stocks had a lead of 4.5% a year over the large stocks during the examined period. It is important to mention that even though the outperformance was impressive, it was also extremely volatile.

One of the first to record this phenomenon was Banz (1981). Banz found that the coefficient on size has more explanatory power than the coefficient on beta in describing the cross section of returns. Since then, the size effect has been reproduced for numerous sample periods and for most major securities markets around the world (Hawawini and Keim, 2000).

The literature provides a little bit more elaborate explanation for the occurrence of this anomaly. Researches have shown a high rank correlation between size and price and between the value ratios and price. Others have documented a significant cross-sectional relation between price per share and average returns. This leads to the conclusion that the separately-identified value and size effects are not independent phenomena because the security characteristics all share a common variable – price per share of the firm's common stock. Fama and French (1992) estimated an equation with multiple value and size variables included as explanatory variables. Their findings suggest that value and size provide the greatest explanatory power in describing the cross section of returns, and suggest that value and size are proxies for the influence of two additional risk factors omitted from the CAMP. In this context, the value and size variables can be viewed as capturing sensitivities to the omitted factors, and the coefficients multiplying the value and size variables are estimates of the risk premia required to compensate for that exposure. This leads us to the proposition made by Fama and French (1993) – a three factor model that describes the time series behavior of security returns:

$$R_t - r_{f,t} = \beta_0 + \beta_1\left(r_{m,t} - r_{f,t}\right) + \beta_2 SmB_t + \beta_3 HmL_t + \varepsilon_t$$

So far, we made clear that value and size effects are well defines in the literature. But what explains these effects? Both premia reflect some common element which manifests only in January that is hard to reconcile with a risk compensation story. We have discussed non-risk-based explanations of the January and Weekend effect earlier in the section.

Nevertheless, much research has been conducted that supports the characterization of the value premium as a compensation for financial distress risk. Others have argued that the size effect is actually a liquidity effect in which small-cap stocks are less liquid that large-cap stocks and therefore provide correspondingly higher returns to offset the higher transactions costs (Brennan, Chordia & Subrahmanyan, 1998). Yet, some insist that size and value effects are nothing but biases in the databases used by researchers.

2.6.2.3 The Prior Return or Momentum Effect

The momentum effect essentially states that the tendency of the price is not likely to change direction, i.e. it is probable that a rise or fall of a stock price will be followed by a rise or fall respectively. It has been shown that the history of stock returns has explanatory power in the current movement of stock returns. Stocks with prices on an upward movement during the previous 3 to 12 months have a higher expected probability of continuing this upward movement the following 3 to 12 months. The temporal pattern in prices is referred to as momentum. Jegadeesh and Titman (1993) show that a strategy that simultaneously buys past winners and sells past losers generates significant abnormal returns over holding for periods of 3 to 12 months. These profits appear to be independent of market, size or value factors and has persisted in the data for several years. With this in mind, Carhart (1995) proposes a model similar to the one of Fama and French (1993), including a momentum factor defined as the difference in returns between a portfolio of "winners" and a portfolio of "losers" in the mindset followed by Jegadeesh and Titman (1993). This coefficient appears to be not only positive but also statistically significant and it cannot be explained by the other three factors. Nevertheless, finding a risk-related explanation for the momentum effect has proven to be extremely challenging. Researchers have proposed behavioral explanations of the momentum effect that use the existence of irrational traders who underreact to new information, but such models are hostile to verification due to their nature.

2.6.3 Over/Under-Reactions

Another challenge to the efficient market hypothesis is that individuals often over- or under-react to news. Sometimes, individuals tend to be conservative and rely too much on their prior beliefs, and hence do not react as intensively as the EMH theory would suggest - in other words they under-react. On the other hand, information that is salient and prominent often captures people's attention and becomes more important in the decision-making process than objective factors would dictate. People tend to assign a heavier weight to such information in forming new beliefs, resulting in over-reaction. Prices can therefore deviate from their fair or rational market value at least temporarily.

Biased reactions such as over- and under-reactions could be consistent with the Efficient Market Hypothesis, if they were split randomly. In that case, they would cancel each other out, and can would be considered as nothing more than white noise. However, research in experimental psychology suggests that, in violation of Bayes' rule, most people tend to over-react to unexpected and dramatic news events. The question then arises whether such behavior matters at the market level.

The term over-reaction implies a comparison to a reaction level that would be considered to be appropriate. As a well-established norm used to be treated the Bayes' rule which prescribed the correct reaction to new information. Nevertheless. Kahneman et al. (1982), in their paper "intuitive Prediction: Biases and Corrective Procedures", show that Bayes' rule does not provide a suitable characterization of how individuals should react to new information. When individuals revise their beliefs, they tend to overweight recent information and underweight prior data. According to Kahneman et al. (1982), people tend to make predictions according to the following simple rule: "The predicted value is selected so that the standing of the case in the distribution of outcomes matches its standing in the distribution of impressions". Even though this rule violates the basic statistical principal that extremeness of predictions must be moderated by considerations of predictability, DeBondt (1985) has provided considerable evidence that support that the actual expectations of professional security analysts and economic forecasters display the same overreaction bias. In fact, one of the earliest observations about the subject was made by Keynes (1964) where he characteristically states: "… day-to-day fluctuations in the profits of existing investments, which are obviously of an ephemeral and nonsignificant character, tend to have an altogether excessive, and even an absurd, influence on the market". Researchers have been making statements on this phenomenon of abnormal behavior around that time. Williams (1956) in his work "Theory of Investment Value" states that "prices have been based too much on current earning power and too little on long-term dividend paying power". Later, Arrow (1982) recognized that the work of Kahneman et al. (1982) clearly typifies this excessive reaction to new information flow that appears to be a common characteristic of all securities and futures markets. Two specific examples to which Arrow was referring are the excess volatility of security prices and the so-called price earnings ratio or value anomaly as we mentioned before.

The excess volatility issue has been investigated in detail by Shiller (1981). He concluded that, as far as the last century is concerned, there is just not enough variation in dividends that could rationally justify the observed aggregate price movements. In addition, Kleidon (1981) concluded that stock price movements are strongly correlated with the following year's earnings changes. A combination of the results of Shiller (1981) and Kleidon (1981) can lead to the emergence of a clear pattern of overreaction. Even though there has been observed a trend in dividends, investors still seem to attach disproportionate importance to short-run economic developments.

Up to this point the over-reaction hypothesis has received some positive feedback from researchers and starts to take shape as a realistic consistently-occurring phenomenon. Nevertheless, it is still reasonable to question, how does this process survive arbitrage? Shouldn't rational traders take advantage of this effect? Russel and Thaler (1985) attempted to answer such questions. After research, they concluded that the existence of some rational agents is not sufficient to guarantee a rational expectations equilibrium in an economy with some "noise" traders.

DeBondt & Thaler (1985), in their paper "Does Stock Market Overreact?", attempted to elaborate on the subject. Based on the fact that, if stock prices systematically overshoot, then their reversal should be predictable from past data alone with no use of any accounting data such as earnings, they suggested two main hypotheses:

a) Extreme movements in stock prices will be followed by subsequent price movements in the opposite direction.
b) The more extreme the initial price movement, the greater will be the subsequent adjustment.

They insisted on the importance of the over-reaction effect due to the fact that it may apply in many other contexts. Thus, based on a model originally proposed by Beaver and Landsman (1983), they concluded that their results are indeed consistent with the over-reaction hypothesis. As they report, over the last half-century, loser portfolios of 35 stocks outperform the market by, on average, 19.6%, thirty-six months after portfolio formation. Winner portfolios, on the other hand, earn about 5% less than the market, so that the difference in cumulative average residual between the extreme portfolios equals to 24.6%.

Furthermore, their results had more notable aspects. The over-reaction effect appeared to be asymmetric in the sense that it was much larger for losers than for winners. In addition, most of the excess returns are realized in January, a fact that is consistent with the seasonality effects that have been discussed earlier.

Their over-reaction hypotheses, in essence, predicts that stocks with more extreme return experiences will result in more pronounced subsequent price reversals. Thus, a non-challenging way to generate more extreme observations is to lengthen the portfolio formation period. The same fact can be verified by comparing the test period performance of less versus more extreme portfolios for any given formation period. Their results confirm the prediction of the over-reaction hypotheses. The subsequent price reversals, measured by the difference between the cumulative average residuals of the extreme portfolios, follow in movement the cumulative average residuals. This as the subsequent price reversals grow, so do the cumulative average residuals.

All in all, the results of their study have interesting implications for most anomaly effects discussed earlier such as the size effect, the January effect and the value effect. Blume and Stambaugh (1983), Keim (1982) and Reinganum (1981) have studied the interaction between the size and January effects. Their findings strongly suggest that the small firm effect can be also considered as a "losing firm" effect based on the turn-of-the-year. The results of DeBondt an Thaler (1985) strengthen this approach since according to them losers earn exceptionally large January returns while winners do not. Nonetheless, companies in the extreme portfolios do not systematically differ with respect to market capitalization.

Roll (1983) credits the tax-loss selling phenomenon for the occurrence of the January effect. But even after purging the data of tax-loss selling effects, Reinganum (1983) finds a considerably smaller January seasonal effect related to company size. This result may be due to his particular definition of the tax-loss selling measure. His measure is related to the securities' relative price movements over the last six months prior to portfolio formation only. Thus, if many investors choose to wait longer than six months before realizing losses, the portfolio of small firs may still contain many losers.

DeBondt and Thaler (1985) findings controvert the hypothesis that the tax-loss selling phenomenon is responsible for the occurrence of the January effect. To begin with, if in early January selling pressure disappears and prices rebound to equilibrium levels, the loser portfolio should not rebound once again in the second January of the test period since it outperforms the market. The same can be applied to each subsequent year. Secondly, if prices rebound in January, why is that effect so much larger in magnitude than the selling pressure that caused it during the final months of the previous year? If that is indeed the case, then investors may wait for years before actually realizing losses.

Considering the value effect, the results of DeBondt and Thaler (1985) support the hypothesis that high value stocks are overvalued whereas low value stocks are undervalued. However, this argument implies that the value effect is also a January effect for the most part.

In conclusion, evidence provided by DeBondt and Thaler (1985) support the prediction of the overreaction hypothesis that portfolios of prior losers are found to outperform portfolios with prior winners. Thirty-six months after portfolio formation, the losing stocks have earned about 25% more than the winners, even though the latter are significantly riskier.

2.6.4 Trading Rules

Another barrier to the prevalence of the Efficient Market Hypothesis is the existence of filter rules or trading rules. A filter rule consists of a set of orders that determine the actions to be taken when shares rise or fall in value by a specific percentage. If markets are consistent with the weak form of EMH, then the patterns suggested by such filter rules should not be able to yield excess returns.

Researchers have tested empirically a plethora of filter rules in order to be ascertained whether such rules can yield excess profits, at least in comparison with the level of profits that the EMH would dictate. Most of these tests rely on historical data.

One of the first researchers to test such a filter rule was Basu (1977). Basu focalized on the hypothesis that claims that low P/E securities will tend to outperform high P/E stocks. Essentially, prices of securities and biased, and the P/E ratio is one of the indicators of this bias. This hypothesis originates on the value effect, and Basu based on his findings proposed some trading filters that could provide excess profits during the period examined.

In order to test the hypothesis at hand. He formed at least two portfolios consisting of securities with similar P/E ratios for any given period within his sample. Then he compared the risk-return relationships as well as the performance in returns of these portfolios. In addition, in order to test the consistency of the Efficient Market Hypothesis in his sample, he compared the returns of the low P/E portfolio with the returns of randomly selected securities that have the same amount of risk.

His research provides evidence that disprove the EMH. The two portfolios with the lowest P/E achieved higher returns on average than the two portfolios with the highest P/E. He also clearly shows (in Table 1, Basu 1977) that as one moves from the low towards the high P/E portfolios, the average annual rates of return decline. The higher returns on the lower P/E portfolios were not associated with higher levels of systematic risk, as the capital market theory would suggest. This became clear after taking into consideration Jensen's measure (differential return). The low P/E portfolios still earned more than they should according to their level of risk, while the high P/E portfolios earned less than they should according to their level of risk respectively. In addition, these results are also statistically significant when assuming normality. The correlation coefficients (Fisher Index) show that all of the P/E portfolios are well diversified, since it's value between each portfolio and the market is above 95%. As a result, the Sharpe measure (reward-to-variability) also shows that the performance of the low P/E portfolios is superior to that of their high ratio counterparts. A similar property was verified by various researchers.

Friend and Blume (1970) and Black, Jensen and Scholes (1972) showed that the differential returns are on average non-zero and are inversely related to the level of systematic risk. Low-risk portfolios usually earn significantly more than what is predicted by the model and also high-risk portfolios earn significantly less than what is predicted by the model, which directly contradicts the theory.

After establishing these results, Basu (1977) proceeds to test whether the performance of the lower P/E portfolio, after adjustments for portfolio-related costs and tax effects, is superior to the performance of portfolios composed by randomly selected securities with the same overall level of risk.

His results offer directly comparable return of portfolios, but in order to adjust returns to transaction costs, since as he considers them to be related to the type of investor, Basu divides the investors into four different types. The first two groups mainly consist of investors who enter the market for some pre-specified reason other than speculation, while the other two include speculators and traders whose goal is to profit from the market's reaction to new information. The importance if this distinction lies to the fact that the effective transaction costs for each case are different. In addition to this hypothesis, Basu also introduces three more assumptions. First, the marginal costs off search and information processing for the portfolios were assumed to be ¼th of 1% per annum. Second, capital gains and dividends are assumed taxable annually at 25% and 50% rate respectively. Lastly, in the evaluation of profitability of a tax-paying trader investing in the lowest P/E portfolio instead of a random portfolio, the effect of tax deferral by trading the random portfolio at the end of the period rather than annually was deducted from the portfolio with the lowest P/E.

The results suggest, that even after the adjustments for portfolio-related costs and taxes, the mean incremental returns could have been earned by acquiring the portfolio with the lowest P/E instead of a random one with the same risk level, for each one of the four investor categories. In fact, by investing in the lowest P/E portfolio, the first two categories of investors could have achieved 2% to 3% more returns per annum than the associated random portfolios of equivalent risk. Even though the rest of the traders could have also achieved higher returns by a small percentage, the differences are generally not significantly different from zero.

All in all, the results are consistent with the view that P/E ratio information is not fully reflected in security prices, at least not as fast as the semi-strong form of market efficiency would dictate. Instead, it seems that this phenomenon persisted throughout the duration of the studied period. Securities trading appeared to be consistently inappropriately priced, creating opportunities for every type of investor to achieve excess returns. Nevertheless, obstacles such as transaction and search costs moderate the ability of investors to exploit market's inefficient pricing of securities. Subsequently, investors are hindered from earning the abnormal returns which are significantly greater than zero. Thus, the hypothesis that capital markets are efficient in the sense that security price behavior is consistent with the semi-strong version of the fair game model cannot be rejected easily.

Nevertheless, in contrast with the belief that new information is instantaneously included in the security prices, there seem to be lags and frictions in the adjustment process. As a result, publicly available P/E ratios seem to possess part of the information which may catch the investor's attention when forming or revising portfolios.

Based on those findings, Basu proposed a trading filter that would allow investors to achieve excess profits at least within the frames of the studied period. The filter dictates the purchase of securities with low P/E ratio during the first days or weeks of every new year, since according to the literature, the same securities would yield excess profits during the last months of the year.

3. Testing the Efficient Market Hypothesis

Earlier we discussed that market efficiency is divided into three forms based on the type of information that is being reflected on the assets price. Now we will take a more elaborate look on the most prominent tests that are available in the literature for testing each one of the three forms on EMH.

3.1 Testing the weak form of EMH

3.1.1 Runs Test

The runs test is one of the most popular methods used to test the Efficient Market Hypothesis. A run is essentially a series of changes in the price of an asset moving in the same direction. This will result in one of the following three possible outcomes: positive for upward price movement, zero for no price change, and negative for downward price movement.

As we mentioned before, it is better to regard the random walk model as a special case of the more general expected return model in order to make a more detailed specification of the economic environment. Thus, we have the basic model of market equilibrium to be a "fair game" expected return model, with random walk arising when certain environmental conditions are met so as to provide repeats of one-period returns distributions through time. With this concept in mind, we should expect violations of the pure independence assumption of the random walk model.

Among the various researchers that have provided examples for such cases are Osborne (1962) and Fama (1965). In their work, they note that large daily price changes tend to be followed by large daily price changes. Nevertheless, the signs of the successor changes appear to be random. This indeed imply the denial of the random walk model but not of the Efficient Market Hypothesis. This may be due to the fact that when new information arrives to the market, it cannot always be immediately evaluated precisely. As a result, we will have an over-adjustment or under-adjustment to the information. But since we have evidence that indicate that the price changes on days following the initial large change are random in sign, we can safely assume that the initial large change represents an unbiased adjustment of the price effect caused by the flow of new information. This explanation is sufficient for the expected return efficient market model.

Niederhoffer and Osborne (1966) document two cases that arise from complete randomness. In the first case, the data imply that reversals (pairs of consecutive price changes that have different sign) appear to be up to three times as likely as continuations (pairs of consecutive price changes that share the same sign). The second case is characterized by continuations that appear to be more frequent after a continuation rather than after a reversal. For example, let (+|++) indicate the case where we have a positive price change that followed a continuation of positive price changes. Then events such as (+|++) or (-|--) appear to be more frequent than events such as (+|+-) or (-|-+).

In the same paper, Niederhoffer and Osborne (1966) also attempt to offer explanations for these phenomena based on the structure of New York Exchange. In more detail, as Niederhoffer and Osborne describe, there are three specific major type of orders than a trader may place on a given stock: (a) buy limit – which implies the order to buy at a specific price or lower, (b) sell limit – which implies the order to sell at a specific price or higher, (c) buy or sell at market – which implies the order to buy or sell at the lowest selling price or highest buying price respectively, offered by the market). The trader maintains a book with unexecuted price limit orders. It is obvious that unexecuted sell limit orders are at highest prices than unexecuted buy orders. Let us also be reminded that the smallest non-zero price change allowed is 1/8 of the original price.

Now let’s assume that there are more than one unexecuted sell limit orders at the lowest price of any preceding order. Thus, a transaction at this price that is initiated by an order to buy at market, can only be followed by either a transaction at the same price (when the following order is to buy) or a transaction at a lower price (if the following order is to sell). For consecutive price increases to occur, the consecutive market orders to buy exhaust the market limit orders to sell for a specific price. In essence, the excessive tendency towards reversal for consecutive significant (larger than 1/8) price change can arise as a result of stacking unexecuted buy and sell limit orders.

Niederhoffer and Osborne (1966) in their paper provide a more elaborate and extensive version of the explanation. Their findings suggest that an event of (+|++) occurs with higher frequency than that of (+|+-), and this is due to the fact that limit orders tend “to be concentrated at integers, halves, quarters and odd eights in descending order for preference”. The frequency of the event (+|++), which usually requires for sell limit orders to be exhausted at least two consecutive higher prices, reflects the absence of sell limit orders at odd eights more than what the event (+|+-) would reflect. The event sequence (+|+-) would imply that sell limit orders at a single price have been exhausted and so more or less reflects the average bunching of limit orders at all eights.

Even though these results are extremely interesting, the type of dependence uncovered does not imply market inefficiency. It appears to be a direct result of the ability of traders to place buy or sell limit orders and it does not offer any space for the existence of profitable trading rules based on these orders. In the same manner, the fact that trades are concentrated around integers, halves, even eights and odd eights is an interesting fact about the trader’s behavior, but it does not imply market inefficiency.

On the other hand, this analysis implies market inefficiency but not within the framework of the weak form. The content of a trader’s book is very important information about the possible future behavior of prices. This information is only available to the trader, and thus the trader has inside information about the possible outcome of the stock's price. Also, when a specialist is asked for a quote, he may give away information but he is prevented by law from reveling the full contents of his or her book. Thus, the trader seems to have monopolistic access to an important block of information while at the same time he uses that information to turn a profit. In this sense, this result may disprove the existence of market efficiency in its strong form. Tests for the strong form will be discussed later in this chapter.

Run tests have been conducted from various researchers in an attempt to support or disprove the EMH in its weak form. Sharma and Kennedy (1997) applied similar tests on the Bombay Stock Exchange, New York Stock Exchange and London Stock Exchange, concluding that stocks on the Bombay Stock Exchange indeed followed a random walk. Abeysekera (2001) conducted a study on the Colombo Stock Exchange but his results were not consistent with the weak form of EMH. Robinson (2005) studied the Jamaican Stock Exchange and his finding had results similar to Abeysekeras', which implied market inefficiency. Omran and Farrar (2006) work with the Stock Markets of Egypt, Jordan, Morocco, Turkey and Israel, and his findings were mixed with some Stock Markets presenting decisive evidence against the weak form of EMH, while other Stock Markets provided unclear results. Hamid et al. (2010) applies the tests in Asia Pacific countries, rejecting the weak form of EMH. Thomas and Kumar (2010) studied the Indian Stock Market concluding market inefficiency while Aumeboonsuke (2012) elaborating on the stock markets of Malaysia, Indonesia, Philippines, Thailand, Singapore and Vietnam, arrived at a similar result.

3.1.2 Autocorrelation Tests

A bit part of the literature is concerned with tests of serial covariances of returns. Let X_t be a fair game, with unconditional expectation zero and serial covariance:

$$E(\tilde{x}_{t+\tau}\tilde{x}_t) = \int_{x_t} x_t E(\widetilde{x_{t+\tau}}|x_t) f(x_t) dx_t$$

Where f represents a density function. But we assumed that x_t is a fair game, thus:

$$E(\tilde{x}_{t+\tau}|x_t) = 0$$

From this follows that for all lags, the serial covariances between lagged values of a fair game variable are zero. As a result, observations of a fair game variable are linearly independent. But this does not imply that the serial covariance of one-period returns would also be zero. In the weak form tests, as fair game variable we have:

$z_{j,t} = r_{j,t} - E\left(\tilde{r}_{j,t}\middle|r_{j,t-1}, r_{j,t-2}, \ldots\right)$ (9)

But if we take the covariance between r_{jt} and $r_{j,t+1}$ is:

$$E\left(\left[\tilde{r}_{j,t+1} - E\left(\tilde{r}_{j,t+1}\right)\right]\left[\tilde{r}_{jt} - E\left(\tilde{r}_{jt}\right)\right]\right) =$$

$$\int_{r_{jt}}\left[r_{jt} - E\left(\tilde{r}_{jt}\right)\right]\left[E\left(\widetilde{r_{j,t+1}}|r_{jt}\right) - E\left(\widetilde{r_{j,t+1}}\right)\right] f\left(r_{jt}\right) dr_{jt}$$

While (9) does not imply that $E\left(\widetilde{r_{j,t+1}}\middle|r_{jt}\right) = E(\tilde{r}_{j,t+1})$. In essence, in the fair game efficient markets model, the deviation of the return for t+1 from its conditional expectation is a fair game variable. On the other hand, the same does not apply for the conditional expectation since it can depend on the observed return for time t.

In the random walk literature, this problem is not encountered since it is assumed by the theory that the expected returns are stationary through time. In practice however, this implies estimating serial covariances by taking cross products of deviations of observed returns from the overall sample mean return. Thus, the fact that this procedure does not affect in a major way the results of the covariance tests is accidental.

Nevertheless, it is still challenging to state which level of serial correlation would imply the existence of a trading rule, strong enough to yield considerable profits. Alexander (1961) brings the first evidence of the existence of such rules. In his paper, he tests a great variety of trading rules, but the most extensively examined can described as follows:

If the price of a security moves up at least y%, buy and hold the security until its price moves down at least y% from a subsequent high, at which time simultaneously sell and go short. The short position is maintained until the price rises at least y% above a subsequent low, at which time one covers the short position and buys. Movements less than y% in either direction are ignored. Such a system is known as a y% filter. After extensive tests, in his final paper Alexander concludes:

"In fact, at this point I should advise any reader who is interested only in practical results, and who is not a floor trader and so must pay commissions, to turn to other sources on how to bear buy and hold. The rest of this article is devoted principally to a theoretical consideration of whether the observed results are consistent with a random walk hypothesis".

Eventually, Alexander finds some evidence in his results that do not support the independence assumption of the random walk model, but a random walk model is not mandatory for market efficiency. Market Efficiency under the sub martingale model would support the existence of filter rules that can turn profit but still cannot beat buy-and-hold strategies. The same result is supported by the findings of Fama and Blume (1966) who also compare the profitability of various filters to a buy-and-hold strategy for individual stocks.

Nevertheless, taking a more elaborate look on the filters of both Alexander and Fama & Blume, evidence of inconsistency can be found even for the sub martingale model. More specifically, very small filters of 0.5% to 1.5% indicate that it is possible to devise a very short-term strategy that will on average outperform buy-and-hold. If we exclude/ignore transaction costs, the earnings of the very short-term strategy at hand will be minimal but, due to the frequency of transactions, they will accumulate considerable amount of profit in the long run, outperforming the buy-and-hold. Of course, when we take into consideration transaction costs, their advantage over the buy-and-hold disappears.

All in all, such tests do produce empirically noticeable results that do not agree with the Efficient Market Hypothesis. But, even in cases where the results are statistically significant, the economic consequences are so minimal that can be hardly used to declare a market inefficient.

3.1.3 Variance Ratio Tests

The Variance Ratio Tests is essentially a tool that allows us to examine the weak form of EMH by investigating whether the variance of the time series at hand is stationary. There are many different tests that one can use (Lo & MacKinlay 1988, Chen and Deo 2006). The basic principle of the test is to obtain and compare the variance of a subset of the time series with the variance of a different similarly sized subset. Essentially, it tests if our time series follow a random walk. Lo and MacKinlay (1988) proposed two statistics for testing an individual VR estimate which are robust under homoscedasticity and heteroskedasticity. In practice, it is customary to examine the VR statistics for several k values. The null is rejected if it is rejected for some k value. However, as stressed by Chow and Denning (1993), this sequential procedure leads to an oversized testing strategy. In this context, multiple VR tests have been suggested such as multiple comparison tests (Chow and Denning, 1993), and Wald-type joint tests (Richardson and Smith 1989, Cecchetti and Lam 1994). Even though the individual Lo-MacKinlay and multiple VR tests are quiet powerful testing for homoscedastic of heteroskedastic nulls, it is critical to note that these tests are asymptotic tests in that their sampling distributions are approximated by their limiting distributions. Indeed, the sampling distribution of the VR statistic can be far from normal in finite sample, showing severe bias and right skewness. These finite sample deficiencies may give rise to serious size distortions or low power, which can lead to misleading inferences. This is especially true when the sample size is not large enough to justify asymptotic approximations (Cecchetti and Lam, 1994). To circumvent this problem, some alternatives have been proposed, such as Chen and Deo (2006) with a power-transformed VR statistic, Wright (2000) with exact VR tests based on rank and sign, Whang and Kim (2003) with subsampling method, and Kim (2006) with bootstrap method, among others.

In a later chapter we will see in detail how can one carry out such a test in practice.

3.2 Testing the semi strong form of EMH

So far, we have discussed the methods that concern the tests for the weak form of Efficient Market Hypothesis, which contains information such as past price history of stocks. Now we will look into methods that allow for testing of the semi strong form of EMH, which is concerned with where current prices fully reflect all obviously publicly available information. Each test however is focusing in a specific kind of financial information generating event that can have as a result a change in stocks price. As a result, every test, even if it has positive results for the EMH, only supports the model specified. The main idea is that as such evidence accumulate for various models, the concept of the semi strong form of EMH will be accepted is valid.

However, the events have to be impactful enough to the prices to make a difference. Thus, only a few major types of information generating financial events are taken into consideration.

One of the first and most important papers on the subject was written by Fama, Fisher, Jensen and Roll (1969), where the researchers studied the adjustment of stock prices to the new information provided by events such as splits. The only real implication of splits should be the multiplication of the number of stocks without increasing the total value of the stocks at hand. Thus, no new information that affect the price should be generated. Fama, Fisher, Jensen and Roll assume that a split also represents a set of fundamentally important information. It is the "hidden" information that lead the company to the decision of the split in the first place. The main focus of the paper is to examine unusual behavior of stocks prices around the split dates, and then to elaborate on what extend the splits or other fundamental factors are responsible.

The paper is elaborating on the following model:

$$\tilde{r}_{j,t+1} = a_j + \beta_j \tilde{r}_{M,t+1} + \tilde{u}_{j,t+1}$$

Where $r_{j,t+1}$ is the rate of return on asset j for month t, $r_{m,t+1}$ is the corresponding return on a masker index M, a_j and β_j are parameters that can vary from security to security, and $u_{j,t+1}$ is a random disturbance.

In this model, the abnormal or unusual behavior caused by the stocks splits would be reflected in the estimated regression residuals for the months surrounding the splits. For every split, the month in which the effective date of a split occurs is defined as O. Month 1 and month -1 will represent the months following and preceding the split accordingly. The average residual over all the split securities for month m would be defined as:

$$u_m = \sum_{j=1}^{N} \frac{\tilde{u}_{jm}}{N}$$

Where $\tilde{u}_{jm}$ is the sample regression residual for security j in month m and N is the number of splits within the month. The researchers also define the cumulative average residual U_m as:

$$U_m = \sum_{k=-29}^{m} u_k$$

The average residual u_m can be interpreted as the average deviation (in month m relative to split month) of the returns of split stocks from their normal relationships with the market. In the same way, U_m can be interpreted as the cumulative deviation from month -29 to month m. Finally, in their paper the researchers also define $u_m^+, u_m^-, U_m^+ \, and \, U_m^-$ as the average and cumulative average residuals for splits followed by increased (+) and decreased (-) dividends respectively. As an "increase", we take the case where the percentage change in dividends on the plit share in the year after the split is greater than the percentage change for the New York Stock Exchange as a whole, while a "decrease" would be the case of a relative dividend decline.

The results of the Fama, Fisher, Jensen and Roll (1969) show that the cumulative average residuals rise in the 29 months prior to the split, are uniformly positive for all three dividend categories. This cannot be solely attributed to the splitting process, since in only about ten per cent of the cases that the time between the announcement and the effective day of the split is greater than four months. In reality, it appears that the decision of a firm to proceed to a split occurs in unusually good times, i.e. during periods when the prices of their stocks would have increased more than it would be implied by their normal relationship with general market prices, which itself probably reflects a sharp improvement, relative to the market, of the earnings of the firms in the time span preceding the split.

These results are being depicted in the following graphs that are taken directly from Fama, Fisher, Jensen and Roll (1969).

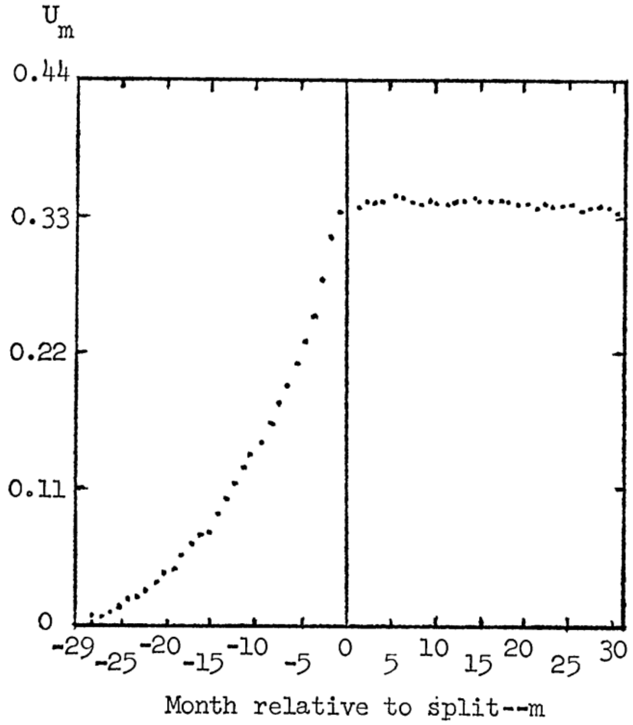


Figure 1: Cumulative average residuals of all splits.

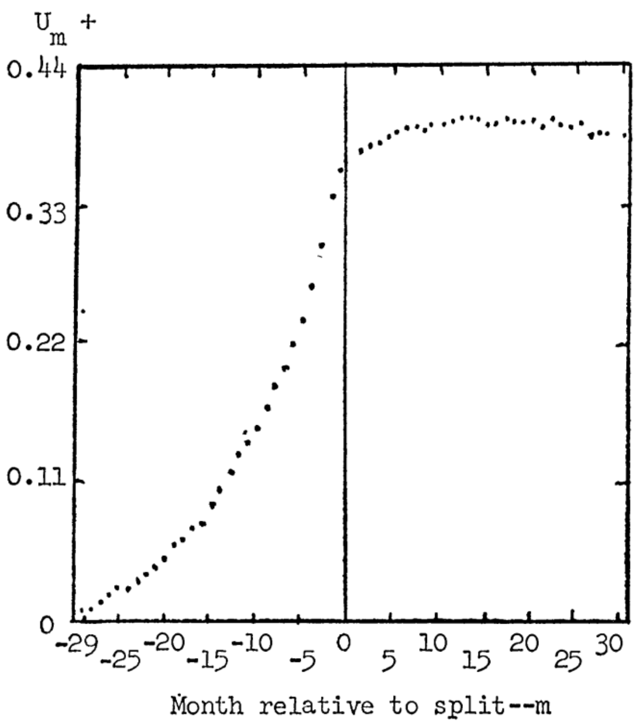


Figure 2: Cumulative average residuals for dividend increased.

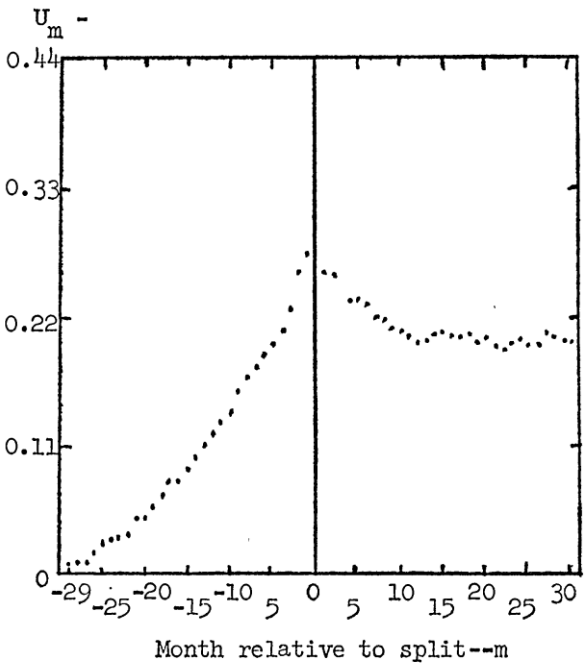


Figure 3: Cumulative average residuals from dividend decreases.

Notice that after the split month, there is little no to further movement in U_m, which is the cumulative average residual for all splits. This result is surprising, considering the fact that 71.5% of all splits experienced greater percentage dividend increases in the year after the split occurred in comparison to the average of all securities on the New York Stock market Exchange.

With this in mind, Fama, Fisher, Jensen and Roll (1969) suggest that when a split is announced, the market interprets it as a signal of the positive expectations of the company's directors about the course of future earnings and the company's ability to be able to maintain dividend payments to a higher level. All in all, the large price increases preceding the split could be a sign of the alteration in expectations concerning the future potential of the firm, rather than an immediate result of the actual split.

If this hypothesis is correct, then the behavior of returns subsequent to splits should be substantially different for the cases where the dividend increase materializes than for the cases where it does not. This hypothesis is indeed being verified by the data. The cumulative average residuals for the increased dividends move upwards even after the split is declared, which is consistent with the hypothesis that when the split is declared, there is a price adjustment in anticipation of future dividend yields. In the case of stocks associated with decreased dividends, the hypothesis still appears to be valid. The cumulative average residuals for these stock rise in the period preceding the split, but then fall dramatically within a short period after the split, since the anticipated dividend increase does not meet the expectations. It is also visible that after a year of the split, the cumulative average residual has returned to the same level as it was five months preceding the split, which can be assumed to be the earlier time that reliable information about a split is likely to reach the market. Thus, we see that by the time it has become obvious that the anticipated dividend increase is not going to be materialized, the apparent effect of the split has faded away and the stock's returns have reverted to their normal relationship with the market returns.

Last but not least, we have to mention even though post-split returns are different depending on whether the dividend increases materialize or not, and despite the fact that most of the split securities do experience increase in dividends, when we examine all the split securities together, there is no movement in the securities average residuals. As a result, the market appears to deliver unbiased forecasts of the implication of a split for future dividends, and these forecasts contribute in a quick adjustment of the price of the securities, which occurs within the first month that follows the split. Fama, Fisher, Jensen and Roll (1969) come to the conclusion that the market is indeed efficient in the sense of the weak form of EMH, at least with respect to the information flow relative to splits.

In order to test for EMH in its weak form, each test can focus in a different type of information generating financial event. In this manner, Ball and Brown (1968) chosen the annual earnings announcements as the information generating event at hand. In their study, they use residuals from a time series regression of the annual earnings of a company on the average earnings of all their firms to classify the firm's earnings for a given year as having increased or decreased relative to the market. Subsequently, they use residuals from regressions of monthly common stock returns on an index of returns to compute the cumulative average return residuals separately for the earnings that increased and similarly for those that decreased. Ball and Brown (1968) come to the conclusion that only a small percentage (10% to 15%) of the total information of the annual earnings announcement has not been anticipated by the month of the announcement, as reflected in the stock's price.

Waud (1968) has used a similar method of residual analysis, having as a subject of study the announcements of discount rate changes by Federal Reserve Banks, and their effects. In this case, the residuals are nothing but the deviations of the daily returns on the S&P 500 index from the average daily return. His findings suggest the existence of a statistically significant announcement effect on stock returns for the first trading day that follows an announcement. Nevertheless, the magnitude of adjustment is minimal, never exceeding the 0.5%. It is very important to note, that as Waud mentions in his paper, it appears that the market anticipates the announcements, and thus stocks prices are adjusting a day prior to the announcement. Waud is lead to this conclusion by the consistent non-random patterns of the signs of average return residuals on the days immediately preceding the announcement. Thus, this observation is particularly interesting and appears to provide some support to the theory behind the Efficient Market Hypothesis.

Scholes (1969) studies the same subject and uses as an information generating financial event large secondary offerings of common stock and on new issues of stock. In his findings one can see that on average, secondary offerings of common stock are coupled with a decline with magnitude of one to two percent in the cumulative average residual returns of the corresponding common stocks. Since the magnitude of the price adjustment appears to be unrelated to the size of the issue, Scholes concludes that the adjustment at hand occurs due to the negative information implied by the fact that somebody is trying to sell a large block of a firm's stock. Another important fact that arises from Scholes' research is that the value of information in a secondary also depends on some extend to the identity of the vendor. The largest negative cumulative average residuals occur where the vendor is the corporation itself or one of its officers. The second largest would be the investment companies but the distance between the two is considerable. We also have to keep in mind that the identity of the vendor is not generally known to the public at the time of the secondary offerings of common stock. In addition, company insiders need only to report their trading activities in their own company's stock to the Securities and Exchange Commission within six days after a sale. After the period of six days, the market appears to have already fully adjusted to the new information provided by the secondary offerings, a fact that is indicated by the random behavior of the average residuals.

Scholes also applies the same method on another sample that consists of new issues of common stock. Similar to the results of Fama, Fisher, Jensen and Roll (1969), the cumulative average residuals rise in the months preceding the new security offering (suggesting that parties tend to issue new security offerings after favorable recent events). The months following the new offering, the average residuals behave randomly, indicating that the information contained in the new offering is now fully reflected in the price.

To sum it up, the available evidence arising from the research of events considering public announcements on common stock returns are consistent with the semi-strong form of Efficient Market Hypothesis.

3.3 Testing the strong form of EMH

Tests concerning the strong form of Efficient Market Hypothesis deal with whether all available information is fully reflected in prices. In other words, they delve into the possible existence of individuals with monopolistic access to private information that can achieve higher expected trading profits than the rest of the traders in the market.

Niederhoffer and Osborne (1966), after extensive research, came to the conclusion that specialists on the New York Stock Exchange apparently use their monopolistic access to information concerning unfilled limit orders to generate monopoly profits. In addition, Scholes (1969) also provides evidence which indicate that officers of corporations sometimes have monopolistic access to information about their own firms.

Another relevant subject that has concerned the literature is who are the people in the investment community that have monopolistic access to information. The most comprehensive studies are those of Sharpe (1965), Sharpe (1966), Treynot (1965), Jensen (1968) and Jensen (1969).

Focusing on Jensen's work, he uses a risk-return framework to evaluate the performance of a set of mutual funds over a ten-year period. The main question he attempts to answer is if mutual fund managements have any special access to information which allows them to earn returns above the norm. Jensen approaches the question from a plethora of angles and concludes that even though his results apply to only one segment of the investment community, they are striking evidence in favor of the efficient markets model. He specifically states:

"Although these results certainly do not imply that the strong form of the martingale hypothesis holds for all investors and for all time, they provide strong evidence in support of that hypothesis. One must realize that these analysts are extremely well endowed. Moreover, they operate in the securities markets and financial communities. Thus, the fact that they are apparently unable to forecast returns accurately enough to recover their research and transactions costs is a striking piece of evidence in favor of the strong form of the martingale hypothesis – at least as far as the extensive subset of information available to these analysts is concerned."

4. Data Snooping

When performing a test on weak-form EMH, just as with every other statistical test, tests for weak-form EMH have a chance of making an error. The researcher could make the mistake of either rejecting the null hypothesis when it is true - type 1 error – or the chance of failing to reject the null hypothesis when it is false – type 2 error. Typically, the researcher sets the probability of type 1 error equal to 0.05 for a single test. The runs test, which is commonly used to test if the market is weak-form efficient, does not require any parameters, as with other tests and thus would have a type 1 error rate equal to 0.05. Tests that require the investigation of accurate representation of the data are biased by data snooping (Lo & MacKinlay 1990, Romano & Wolf 2005, White 2000) and thus the probability of type 1 error is not an accurate representation of the type 1 error rate. Take the Ljung-Box tests (autocorrelation test) for example. The test requires the researcher to input the lag period when calculating the test statistic. Another example is the Lo & MacKinlay tests (variance ratio test), which requires a holding period in order to calculate the test statistic. Looking at the ACF (Autocorrelation Functions) and PACF (Partial Autocorrelation Functions) charts, modeling the indices of interest, and then determining the parameters for these tests constitute data snooping (White 2000).

Data snooping is also known as data mining. Data mining has received has positive connotations in the past. It is a great tool of extracting valuable relationships from mass data. Nevertheless, most of the negative feedback arises from the fact that often, naïve practitioners can mistake the spurious for the substantive. Numerous of such cases are recorded in the literature. Various researchers have examined issues of model selection in the context of specification searches, with specific attention to issues of inference. Leading position in the subject had Leamer (1978, 1983), with his work being focused on the dangers of data snooping and the methods that are possibly useful in order to evaluate the strength of the relationship obtained by data mining. Nevertheless, little has been suggested to properly account for the effects of data snooping up until White (2000).

White (2000) baptizes his technique the bootstrap reality check (BRC). The BRC provides simple and straightforward procedures for testing the null that the best model encountered in a specification search has no predictive superiority over a given benchmark model, permitting account to be taken of the effects of data snooping. In essence, the BRC seeks to control the simultaneous rate of error under the null hypothesis, but his results are also relevant for controlling average error rates.

Often, one would like to identify further outperforming strategies, apart from the one that is best in the sample. While the BRC algorithm of White (2000) does not address this question, it could be modified to do so. Romano & Wolf (2005) provided a method that extends the BRC in the sense that it can also identify strategies that beat the benchmark but are not detected by the BRC. They developed a stepwise multiple testing method in which the BRC constitutes the first step. Further outperforming strategies can be detected in subsequent steps, while maintaining control of the FWE (familywise error rate), which is defined as the probability of incorrectly identifying at least one strategy as superior. The FWE is the prevalent measure for data mining. The crucial difference between the two methods is that if some hypotheses are rejected in the first step, the method of Romano & Wolf (2005) does not stop where the BRC would and potentially rejects further hypotheses in subsequent steps. Thus, the stepwise multiple testing method offers improved power without sacrificing the asymptotic control of the FWE. In order to reject hypotheses in every given step, Romano & Wolf (2005) constructed a joint confidence region for the set parameters that pertains to the set of null hypotheses not rejected in previous steps. The joint confidence region was determined by an appropriate bootstrap method, depending on whether the observed data are independent and identically distributed random variables or a time series.

In addition, Romano & Wolf (2005) also proposed the use if studentization whenever is possible. Among other reasons, Romano and Wolf mention that they prefer studentization because it results in a more even distribution of power among the individual tests.

5. Methodology

5.1 Tests for the weak form of EMH

5.1.1 Runs Tests

Earlier we delved into the definition and purpose of runs tests. In this chapter we will elaborate on the methodology that the writer used to carry out his tests of the weak form of the Efficient Market Hypothesis. As a reminder, runs tests are "a succession of identical symbols which are followed or preceded by different symbols or no symbols at all" (Siegel, 1956). According to Mobarek and Keasey (2000), runs tests which are non-parametric tests, have an advantage over parametric tests because runs tests ignore the properties of the distribution of the series. The null hypothesis for runs tests is: H_0: The series is a random series. The formula for runs tests has been given by Wallis and Roberts (1956) as:

$$m = \frac{N(N+1) - \sum_{i=3}^{i} n_i^2}{N}$$

Where m is the total expected number of runs, N is the total number of observations, and n_i is the number of observations in each category i. For Large number of observations (N>30), the sampling distribution of m is approximately normal and the standard error of m is given by:

$$\sigma_m = \left(\frac{\sum_{i=1}^{3} n_1^2 [\sum_{i=1}^{3} n_i^2 + N(N+1)] - 2N \sum_{i=1}^{3} n_i^3 - N^3}{N^2(N-1)}\right)^{\frac{1}{2}}$$

And the Z-statistic to test the hypothesis is

$$Z = \frac{R \pm 0.5 - m}{\sigma_m}$$

Where R= actual number of runs, m=expected number of runs and 0.5 = continuity adjustment. The sigh is negative if R≥m and positive otherwise.

A lower-than-expected number of runs indicate that the market's overreaction to information, subsequently reversed, while higher number of runs reflect a lagged response to information. Either situation would suggest an opportunity to create excess returns (Poshakwale, 1996). The runs test converts the total number of runs into a Z statistic. For large samples, the Z statistic gives the probability of a difference between the actual and expected number of runs. If the Z value is greater than or equal to 1.96, then the null hypothesis is rejected at the 5% level of significance (Sharma & Kennedy, 1977).

5.1.2 Autocorrelation Tests

Autocorrelation or serial correlation tests are used to determine whether the observations in time series data are correlated with their lags or not (Kutner, Nachtsheim, Neter & Li, 2005). The null hypothesis for autocorrelation tests is H_0: There is zero autocorrelation at the first k autocorrelations ($\rho_1 = \rho_2 = \cdots = \rho_\kappa = 0$). For a large sample, the Ljung and Box (1978) Q statistic follows the chi-square distribution with k degrees of freedom:

$$Q_{LB} = N(N+2)\sum_{j=1}^{k}\frac{\rho_j^2}{N-j}$$

Where N equals the number of observations. The ρ_j is the jth autocorrelation and the k represents the number of autocorrelations. Ignoring the sign, a high Q_{LB} statistic implies that the autocorrelation is high and it indicates that the stock price is related to the previous prices. This evidence would be against the weak-form efficient market hypothesis. Therefore, if the Q_{LB} statistic value is greater than or equal to the critical value obtained from the chi-squared table, then the null hypothesis is rejected at the 5% level of significance.

5.1.3 Variance Ratio Tests

According to Lo and Mackinlay (1988), the variance ratio tests assumes that the variance of the increment sin the random walk series is linear in the sample interval. If a series follows a random walk process, the variance of its q-differences would be q times the variance of its first differences.

$$Var(p_t - p_{t-q}) = qVar(p_t - p_{t-1})$$

The Variance Ratio VR(q), is

$$VR(q) = \frac{\frac{1}{q}Var(p_t - p_{t-q})}{Var(p_t - p_{t-1})} = \frac{\sigma^2(q)}{\sigma^2(1)}$$

For a sample size of nq+1 observations $(p_0, p_1, \dots p_{nq})$,

$$\sigma^2(q) = \frac{\sum_{t=q}^{nq}(p_t - p_{t-q} - q\hat{\mu})^2}{h}$$

Where

$$h = q(nq + 1 - q)\left(1 - \frac{q}{nq}\right)$$

and

$$\hat{\mu} = \frac{1}{nq}\sum_{t=1}^{nq}(p_t - p_{t-1}) = \frac{1}{nq}(p_{nq} - p_0)$$

$$\sigma^2(1) = \frac{\sum_{t=1}^{nq}(p_t - p_{t-1} - \hat{\mu})^2}{nq - 1}$$

Lo and Mackinlay (1988) proposed the test-statistics for homoscedasticity as:

$$Z(q) = \frac{VR(q) - 1}{[\varphi(q)]^{\frac{1}{2}}} \approx N(0,1)$$

And the test-statistics for heteroskedasticity as:

$$Z^*(q) = \frac{VR(q) - 1}{[\varphi^*(q)]^{\frac{1}{2}}} \approx N(0,1)$$

Where φ(q) is the asymptotic variance of the variance ratio under the assumption of homoscedasticity, and φ*(q) is the asymptotic variance of the variance ratio under the assumption of heteroskedasticity.

$$\varphi(q) = \frac{2(2q - 1)(q - 1)}{3q(nq)}$$

$$\varphi^*(q) = \sum_{t=j+1}^{q-1}\left[\frac{2(q - j)}{q}\right]^2 \widehat{\delta(j)}$$

Where

$$\hat{\delta}(j) = \frac{\sum_{t=j+1}^{nq}\left(p_t - p_{p-1} - \hat{\mu}\right)^2\left(p_{t-j} - p_{t-j-1} - \hat{\mu}\right)^2}{\left(\sum_{t=1}^{nq}(p_t - p_{t-1} - \hat{\mu})^2\right)^2}$$

The null hypothesis for the variance ratio test is $H_0 =$ The series is a random walk. According to the equations of Z(q) and Z*(q), a high variance ratio will result in a high value of Z test-statistic. If the Z value is greater or equal to 1.96, then the null hypothesis is rejected at the 5% level of significance and this evidence would be against the weak-form of the efficient market hypothesis.

5.1.4 Expected Returns

For our calculations, we also need the expected return of the portfolio assisting of all the Swedish stocks for the period defined. To extract this return, we use a simple model similar to the one proposed by Fama and French (1992). Our model adopts the following form:

$$r_i = \beta_0 + \beta_1 * RiskFreeRate_t + \beta_2 * Inflation_t + \beta_3 * ExhangeRate_t + \beta_4 * Spread_t + e_t$$

$RiskFreeRate_t$: As risk free rate the 3-month short term Swedish interbank ratio for time t is used.

$Inflation_t$: We compute the inflation rate as $\frac{CPI_t - CPI_{t-1}}{CPI_{t-1}}$, where CPI_t, CPI_{t-1} are the Swedish Consumer Price Index of year t and year t-1 respectively.

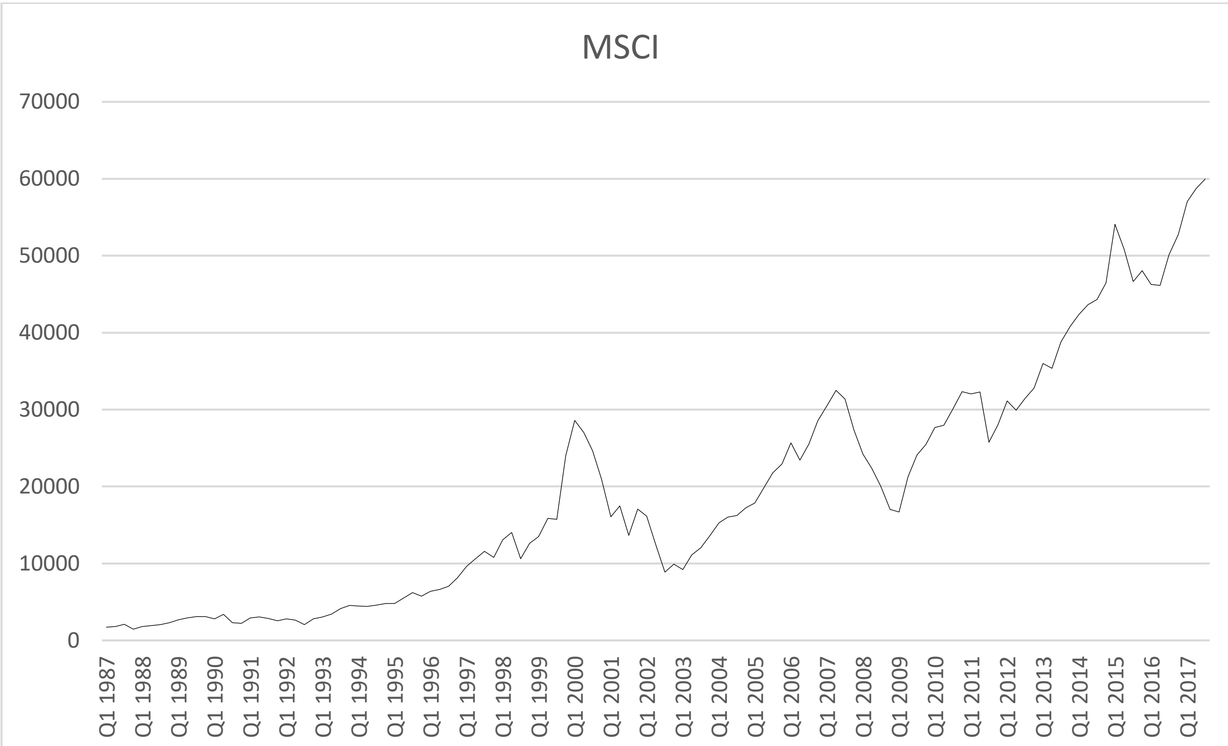
$ExchangeRate_t$: Is the Swedish Krona to United States Dollar Ratio at time t.

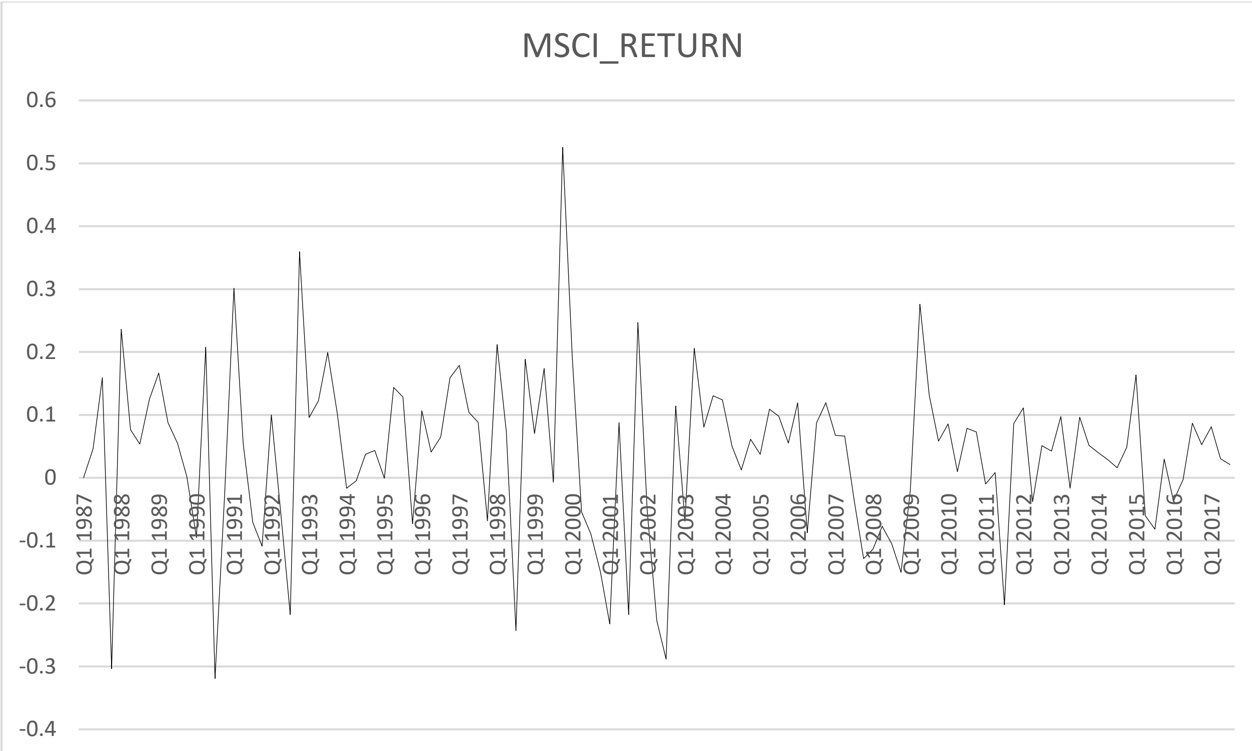
$Spread_t$: As spread we define the difference between the long term (10-year) Swedish interbank interest rate minus the short term (3-month) Swedish interbank interest rate for time t.

6. Data Description

The aim of the empirical part of this paper is to test the validity of the weak form of the EMH using the methods described in the previous chapter. The data used for the tests are an MSCI index that includes all the stocks that were actively traded in the Stockholm Stock Exchange during the period between January 1987 to September 2017. The stocks count up to a total of 789 stocks, and are recorded in quarterly frequency.

The graphs below show the price and the returns of the MSCI index for Sweden during the period examined.





The index which is used for the conduct of this study is called MSCI index and was extracted by the MSCI data base. MSCI (formerly known as Morgan Stanley Capital International) is an American provider of equity, fixed income, hedge fun stock market indexes, and equity portfolio analysis tools. Data of other indexes (e.g. OMSX 30) were also available but the MSCI index was chosen over others due to the fact that it is the older among all the relevant indexes, providing a larger sample.

The index is calculated as:

$$IndexPrice = \sum_{s\in I,t} \frac{EndofDayNumberOfShares_{t-1} * PricePerShare_t * InclusionFactor_t * PAF_t}{FXrate_t}$$

Where:

- $EndOfDayNumberOfShares_{t-1}$ is the number of shares of security s at the end of day t-1.

- $PricePerShare_t$ is the price per share of the security at time t.
- $PricePerShare_{t-1}$ is the price per share of security s at time t-1.
- $InclusionFactor_t$ is the inclusion factor of the security s at time t. The inclusion factor can be one or the combination of the following factors: Foreign Inclusion Factor, Domestic Inclusion Factor, Growth Inclusion Factor, Value Inclusion Factor, Index Inclusion Factor
- PAF_t is the Price Adjustment Factor of the security s at time t.
- $FXrate_{t-1}$ is the FX rate of the price currency of security s vs USD at time t-1. It is the value of 1 USD in foreign currency.

For our calculations, we also used the exchange rate between USD and Sweden Kronor, the Swedish inflation rate, the 10-year Swedish Interbank Rate, the 3-months Swedish Interbank Rate, the Swedish Gross Domestic Product index and the Swedish Consumer Price Index. Data were provided by the Reuters Database.

7. Results

In this chapter, the most important results of the empirical work will be presented and briefly discussed. The tests have been carried out with the use of the software environment called "R". The lines of coding used to produce these results are also provided.

To begin with, we present a small summary of the data.

DATE	MSCI_RETURN	EXCHANGE_RATE	INFLATION_RATE
Q1 1987: 1	Min. :-0.31923	Min. : 5.307	Min. :-1.431e-02
Q1 1988: 1	1st Qu.: -0.03045	1st Qu.: 6.504	1st Qu.: 7.839e-05
Q1 1989: 1	Median: 0.05244	Median: 7.208	Median: 4.131e-03
Q1 1990: 1	Mean: 0.03778	Mean: 7.398	Mean: 5.551e-03
Q1 1991: 1	3rd Qu.: 0.10523	3rd Qu.: 8.097	3rd Qu.: 8.898e-03

Q1 1992: 1	Max.: 0.52543	Max. :10.885	Max.: 4.764e-02
(Other):117			

SPREAD	SWEDEN_INTERBANK_3M	GDP
Min.: -9.2100	Min.: -0.591	Min.: 272927
1st Qu.: 0.0250	1st Qu.: 1.958	1st Qu.: 444436
Median: 0.9780	Median: 3.880	Median: 643569
Mean: 0.7506	Mean: 4.942	Mean: 663313
3rd Qu.: 1.5312	3rd Qu.: 8.155	3rd Qu.: 854524
Max.: 3.2510	Max. :20.000	Max. :1169670

SWEDEN_INTERBANK_3M_RATE

Min.: -0.00591

1st Qu.: 0.01957

Median: 0.03880

Mean: 0.04942

3rd Qu.: 0.08155

Max.: 0.20000

Following, we defined our expected returns model and calculated the coefficients and residuals of the model.

Call:

Lm (formula= MSCI_RETURN~RISK_FREE + INFLATION_RATE + SWEDISH_KRONA_TO_USD + SPREAD, data = data)

Residuals:				
Min	1Q	Median	3Q	Max
-0.40879	-0.06284	-0.00202	0.06696	0.46393

Coefficients:	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	0.086648	0.093454	0.927	0.35573
RISK_FREE	0.134960	0.392329	0.344	0.73146
INFLATION_RATE	0.475369	1.578024	0.301	0.76376
SWEDISH_KRONA_TO_USD	-0.010689	0.011545	-0.926	0.35643
SPREAD	0.027849	0.009101	3.060	0.00274 **

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.1276 on 118 degrees of freedom

Multiple R-squared: 0.08145, Adjusted R-squared: 0.05032

F-statistic: 2.616 on 4 and 118 DF, p-value: 0.03862

Here we can see the estimates, Standard Errors, t-values and Probabilities of our estimates. We see that all of our estimates are statistically significant except the SPREAD. Nevertheless, we will not omit it from the rest of the analysis for the sake of completion.

The tests are applied on the variable EXCESS_RETURN which is nothing but the fitted values of our model.

The results of the runs test performed are the following:

```
        Exact runs test

data:  EXCESS_RETURN

Runs = 22, p-value = 4.434e-14

alternative hypothesis: two sided
```

H_0: *The sequence was produced in a random manner.*

H_a: *The sequence was not produced in a random manner.*

For a large-sample runs test, the test statistic is compared to a standard normal table. We can also come to a conclusion by looking at the p-value. The p-value is lower than the significance level a when a is 10%, 5% and 1%. Thus, *we do reject the null hypothesis – the sequence was produced in a random manner*.

Next follow the results of the autocorrelation tests. First is the Box-Ljung test:

```
        Box-Ljung test

data:  EXCESS_RETURN

X-squared = 52.48, df = 1, p-value = 4.345e-13
```

H_0: The data are independently distributed (i.e. the correlations in the population from which the sample is taken are 0, so that any observed correlations in the data result from randomness of the sampling process).

H_a: The data are not independently distributed, they exhibit serial correlation.

For a large sample, the Q statistic follows the chi-squared distribution. If the Q statistic value is greater than or equal to the critical value from the chi-squared table, then the null hypothesis is rejected at the 5% level of significance. Indeed, our Q statistic (which equals to 52.48) is greater than the critical value for significance level 5%, thus *we reject the null hypothesis – the data are not independently distributed.*

We also performed the Box-Pierce Test, and as expected the results agree with the Ljung-Box test:

```
        Box-Pierce test

data:  EXCESS_RETURN

X-squared = 51.221, df = 1, p-value = 8.253e-13
```

Once again, with such a low p-value we always reject the null hypothesis.

In addition, we use a tool from our next group of tests, the Variance Ratio Tests. The Auto.VR(x) test is provided by R and we make use of it as additional verification to our results. When applied to the data, the test yields the following results:

```
Auto.VR(EXCESS_RETURN)

$stat: Automatic variance ratio test statistic

[1] 19.5596

$sum: 1+ weighted sum of autocorrelation up to the optimal order

[1] 11.23428
```

$$H_0: The\ sequence\ is\ serially\ uncorrelated.$$
$$H_1: The\ sequence\ is\ serially\ correlated.$$

The automatic variance ratio test statistic is calculated as:

$$VR = \frac{\sqrt{\frac{T}{l}}[\widehat{VR}(l)-1]}{\sqrt{2}} \to^{d} N(0,1).$$

The d over the arrow represents the convergence in distribution of VR to standard normal as T, l and T/l approach infinity. The l is the lag truncation point.

Thus, we compare our automatic variance ratio test statistic with a standard normal table. We find that our statistic is greater than the standard normal t-statistic, and as a result *reject the null hypothesis – our sequence is not serially uncorrelated.*

The last tool that we will take advantage of is the variance ratio test provided by Lo and MacKinlay (1998). The results of the test are the following:

```
>Lo.Mac(EXCESS_RETURN,2)

$Stats

        M1              M2

k=2     7.151805        5.078564

> Lo.Mac(EXCESS_RETURN,4)

k=4     9.346376        6.673726

> Lo.Mac(EXCESS_RETURN,8)

k=8     8.312807        6.272202

> Lo.Mac(EXCESS_RETURN,16)

k=16    4.545363        3.703804
```

H_0: The Variance Ratio VR(q) equals 1, our sequence follows a random walk.

H_1: The variance Ratio VR(q) does not equal 1, our sequence does not follow a random walk.

For every k, at 5% significance level, we reject the null hypothesis, thus *our sequence does not follow a random walk.*

8. Summary and Conclusions

In Chapter 1, we began our analysis with a small introduction to the term of Efficient Market Hypothesis. In the following Chapter, we talked about the concept as well as the basic mathematical framework that surrounds the subject. We also delved deeper into the meaning of the Efficient Market Hypothesis, elaborating on the forms that it may take relevant to the information that the market provides. Subsequently, we mentioned the path of evolution that the theory followed from the beginning of its formation until recent years. Then we focalized into the weaknesses of the theory and the critic that crumbles its foundation. Chapter 3 elaborates in detail on the tests that one can use to prove or disprove the forms of the Efficient Market Hypothesis. In Chapter 4, the term of data snooping was discussed, as well as its importance to the results and implications of the research. Chapter 5 presented the methodology that was used to produce the results of our research.

Chapter 7 has investigated the behavior of stock returns by testing RWH in the Swedish Stocks Market. The specific objective of the chapter was to test the weak form of market efficiency in the Swedish Stocks Market. Towards the end, parametric and non-parametric tests are used to analyze quarterly data on a single market index that contains all actively traded stocks listed in the Swedish Stocks Exchange, which count up to 789 different stocks. The results of both parametric and non-parametric tests rejected the Efficient Market Hypothesis in its weak form.

Nevertheless, there are plenty of factors that can be modified and possibly affect the end results.

Due to the great length of our data set, the test results could be different if we divided the period into sub-periods, and then applied the tests to the sub periods a hand. That would reveal possible periods that affected our results negatively in respect with the consistency of our hypothesis towards the weak form of the Efficient Market Hypothesis.

Another factor that plays a vital role in the formation of our results is the very nature of the expected returns model that we decided to use. Following the work of Fama and Frence (1988), we used as our depended variable the returns of the stocks of the Swedish market and we run regressions on macroeconomic variables such as exchange rates, inflation rate and spreads.

Of course, our study only focusses to a very specific version of a model that can take a plethora of forms. In a similar manner, one can take into consideration more factors of the same caliber, such as Monetary Growth M1, M2, M3 and elaborate on what magnitude do the newly introduced factors affect the behavior of the model, the results that the model produces and subsequently the effect that is has considering the consistency of the model towards the weak form of the Efficient Market Hypothesis.

Also, one has to keep in mind that the results are always detained by the data. As mentioned on Chapter 4, the special characteristics of the data set that we focus on may or may not emphasize the quest of our research. A lot of the findings that we face during our analysis are a direct result of the nature of the data and not necessarily due to the weakness of the theory that is being tested.

References

- Alexander S. S. (1961). Price Movements in Speculative Markets: Trends or Random Walks? Industrial Management Review 2, 7-26
- Aumeboonsuke, V., Dryver, A. L. (2014). The importance of using a test of weak-form market efficiency. International Review of Economics and Finance 33, 350–357
- Basu, S. (1977). Investment Performance of Common Stocks in Relation to Their Price-Earnings Rations: A Test of the Efficient Market Hypothesis. The Journal of Finance 32, 663-682
- Campbell et al. (1996). The Econometrics of Financial Markets. Princeton University Press
- Cuthbertson, K., Nitzsche, D. (2004). Stocks, Bonds and Foreign Exchange. Wiley Publications.
- Damodaran, A., (2002). Investment Valuation: Tools and Techniques for Determining the Value of Any Asset. Wiley Publications
- Fama et al. (1969). The Adjustment of Stock Prices to New Information. International Economic Review 10, 1-21
- Fama, E., F. (1965). The Behavior of Stock-Market Prices. Journal of Business 38, 34-105
- Fama, E., F. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. The Journal of Finance 25, 383-417
- Fama, E., F. (1991) Efficient Capital Markets: II. Journal of Finance 46, 1575-1617
- Hamid et al. (2010). Testing the Weak form of Efficient Market Hypothesis: Empirical Evidence from Asia-Pacific Markets. International Research Journal of Finance and Economics 58
- Jegadeesh N., Titman S. (1993). Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency. Journal of Finance Volume 48m Issue 1, 65-91.
- Jensen, M., C. (1968). Problems in Selection of Security Portfolios: The Performance of Mutual Funds in the Period 1945-1964. The Journal of Finance 23, 389-416
- Jensen, M., C. (1969). The Pricing of Capital Assets, and The Evaluation of Investment Portfolios. The Journal of Business 42, 167-247
- Kendall, M., G., Hill, A., B. (1953). The Analysis of Economic Time-Series Part I: Prices. Journal of the Royal Statistical Society 116, 11-34
- Kothari, S., P., Lewellen, J., Warner, J., B. (2004). Stock Returns, Aggregate Earnings Surprises, and Behavioral Finance. MIT Sloan Working Paper No. 4284-03
- Lintner, J. (1965). The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets. The Review of Economics and Statistics 47, 13-37
- Ljung, G., M., Box, G., E., P. (1978). On a measure of lack of fit in time series models. Biometrika 65, 297-303
- Lo, A., W., MacKinlay, A., C. (1988). Stock Market Prices do not Follow Random Walks: Evidence from a Simple Specification Test. The Review of Financial Studies 1, 41-66

- Lo, A., W., MacKinlay, A., C. (1989). Data-Snooping Biases in tests of Financial Asset Pricing Models. The Review of Financial Studies 3, 431-467
- Niederhoffer, V., Osborne, M., F., M. (1966). Market Making and Reversal of the Stock Exchange. Journal of the American Statistical Association 61, 897-916
- Omran, M., Farrar, S., V., (2006). Tests of weak form Efficiency in the Middle East emerging markets. Studies in Economics and Finance 23, 13-26
- Poshakwale (1996). Evidence on Weak Form Efficiency and Day of the Week Effect in the Indian Stock Market. Finance India Vol.X, No.3, 605-616
- Robinson, J. (2005). Stock price behavior in emerging markets: Tests for weak form market efficiency in the Jamaican stock exchange. Social and Economic Studies 54, 51-69
- Samuelson, P., A. (1965). Proof That Properly Anticipated Prices Fluctuate Randomly. Industrial Management Review 6, 41
- Sewell, M. (2011). History of the Efficient Market Hypothesis. UCL
- Sharma, J., L., Kennedy, R., E. (1997). A Comparative Analysis of Stock Price Behavior on the Bombay, London, and New York Stock Exchanges. The Journal of Financial and Quantitative Analysis 12, 391-413
- Sharpe, W., F. (1964). Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk. He Journal of Finance 19, 425-442
- Sharpe, W., F. (1966). Mutual Fund Performance. The Journal of Business 39, 119-138
- Tsay, R., S. (2010). Analysis of Financial Time Series. Wiley Publications.
- Tzavalis, E., Petralias, A. (2009). Investments. AUEB Publishing
- Wallis and Roberts (1956). Statistics: A New Approach. Illinois: Free Press