Tesi doctoral presentada per En/Na

Jordi ANDREU CORBATÓN

amb el títol

"Market Indices: Bases, Biases and Beyond"

per a l'obtenció del títol de

Doctor per la Universitat de Barcelona

Barcelona, 23 de juliol de 2009.

Facultat d'Economia i Empresa
Departament de Matemàtica Econòmica, Financera i Actuarial

UNIVERSITAT DE BARCELONA

A mi padre

Siempre hay esperanza...

“Marx, la broma ha terminado”.

Groucho Marx, *El mundo está loco*. (En referencia al Crack del 1929)

MARKET INDICES: BASES, BIASES AND BEYOND

UNIVERSIDAD DE BARCELONA
DOCTORADO EN ESTUDIOS EMPRESARIALES
DEPT. DE MATEMATICA ECONOMICA, FINANCIERA Y ACTUARIAL
BIENIO 2003-2005

TESIS DOCTORAL

MARKET INDICES: BASES, BIASES AND BEYOND

DIRECTOR

Dr. ANTONIO ALEGRE ESCOLANO

DOCTORANDO

JORDI ANDREU CORBATON

AGRADECIMIENTOS

Quisiera expresar mi agradecimiento al Dr. Antonio Alegre por acompañarme en el final de este largo camino que ha sido mi tesis doctoral. Recordar también a mis tutores durante los dos primeros años de doctorado el Dr. Dídac Ramírez y el Dr. Salvador Torra, pues es en parte suyo el resultado del tercer capítulo. Quisiera hacer llegar a estas tres personas mi gratitud por su dirección y orientación siempre desde el respeto más profundo a mis opiniones. Expresamente, reflejar aquí la ayuda recibida del Dr. Salvador Torra en la obtención de las bases de datos y hacer extensivo a través de él mi agradecimiento a la Bolsa de Barcelona y a Cristina Urbano, del grupo GVC Gaesco.

Agradecer el apoyo a los compañeros del Departament d'Economia y del Departament de Gestió d'Empreses y especialmente a los amigos del área de Finanzas, gracias por su afecto y por hacer más corto y llevadero este viaje.

Al Dr. Màxim Borrell, por su inestimable ayuda y consejos, que me han hecho crecer como investigador y persona. Mi agradecimiento infinito al Profesor Sebastián Cano, por el trabajo que hemos compartido a lo largo de su y mi tesis, al Profesor Daniel Liviano, por su ayuda en la parte empírica y al Profesor Javier H. Gutiérrez.

A todas las personas que han compartido estrechamente mi vida durante esta larga travesía, desde ese día en que llegué a Barcelona queriendo ser doctor. Gracias por su comprensión, amor y paciencia, por entender que no dispuse siempre del tiempo necesario para ellas, y por levantarme cada vez que me caía. Finalmente, agradecer a mi familia, a mis padres y a mi hermana su fe en mí. Y a Sandra, por compartir su sonrisa conmigo.

Aunque esta tesis ha sido un trabajo personal, no habría sido posible sin todas estas personas que han dejado su huella en mi vida y en mi trabajo. Desear que vean su ayuda y sus ideas reflejadas en la versión final de esta tesis. Todos los errores, omisiones o defectos que puedan aún existir en las próximas páginas a pesar de sus esfuerzos por ayudarme, son tan sólo míos.

CONTENTS

INTRODUCTION

Market Indices are perhaps one of the most well known concepts in finance. It is usual to see them in daily news or find them without difficulty in newspapers, television or radio debates. Market Indices have also a crucial role in the professional financial field: hundreds of institutional investors, pension funds or investment banks use, follow and create market indexes. The origins of indexing are located around 1880 when the first equity indexes were created. Since then, indexes have spread over the world developing a huge industry. The Modern Portfolio Theory (MPT), the CAPM or the efficiency analysis contributed to this expansion, and theoretical and empirical studies concluded passive investments (or indexing) were winning strategies. Indexing grew exponentially supported by lower management costs, simplification of the manager selection, superior performance in average, and the EMH and CAPM fever. Indices appeared everywhere, trying to cover not only equity markets, but also Fixed Income products, commodities, Real Estate or Hedge Funds. In the last 20 years indexed institutional assets have grown 40% annually.

Having in mind the importance of indexing, it is really shocking not to find a complete and deep analysis of market indexes. When I started this PhD thesis, I realized with astonishment everybody talks about passive investment, but it was impossible to find indexing literature. No books or articles about index construction were available. Regarding this reality, the main objective of this thesis was to contribute with a complete theoretical and empirical analysis of market indexes. In the following pages I try to concentrate a serious consideration about indexing. A deep literature analysis is presented and completed with an empirical study in the last article. Now, near the end of the PhD adventure, I can say with happiness some authors have detected the same lack in the literature. In the last years, S. A. Schoenfeld and D. Broby tried to answer some of the questions I am going to ask here.

The thesis presented here selected the modern approximation instead of the traditional thesis structure. It tries to reward brevity, applicability and published work. For that reason, each of the chapters has an article structure, presenting the abstract, introduction, the paper development, conclusions, appendixes if necessary and references. Regarding references, I must note here only cited references are presented at the end of each paper. Finally, it is necessary to outline a brief version of the third article, which was accepted for publication in Applied Financial Economics, and different versions of the first article are under consideration in some journals.

A PhD modern approach has advantages but also problems. One is time. I have not been able to complete the theoretical and empirical analysis of indices I wished at the beginning of the journey. I think an exhaustive theoretical study is provided, unfortunately, the empirical part is shorter than expected. It is time to say a complete empirical analysis regarding statistical characteristics of market indexes is still incomplete. After two years of looking for an international data base, I finally found it during 2008. I have been working with the data base since then, but it has been impossible to provide here a finished study.

The final structure of my PhD thesis is as follows:

◊ Introduction.

◊ Some Traditional Financial Ideas Revisited: the Market, Efficiency and Indexing under new Financial Approaches.

  In this first article, a revision of financial fundamental pillars is provided. This revision is absolutely essential to understand theoretical and empirical bases of indexing and hypotheses underlying index construction.

◊ Market Indexes in detail. Biases revealed.

  A deep Market Index analysis is presented here. After an historical revision and a discussion about functions of an index functions and desirable characteristics, market index's biases are provided and studied.

◊ Refunding Market Indexes using VaR.

  Using the theoretical framework of Minimum Risk Indices, these indices are generated for the Spanish, American and Argentinean Stock Markets. Minimum Risk Indices using VaR provide market indices with less risk and with higher profitability in some cases, due to the partial elimination of the efficiency bias.

◊ Conclusions

CHAPTER 1. SOME TRADITIONAL FINANCIAL IDEAS REVISITED: THE MARKET, EFFICIENCY AND INDEXING UNDER NEW FINANCIAL APPROACHES

ABSTRACT

Financial theory is built on ideas such as the market, zero-sum games, efficiency or rationality. The conclusion of traditional approaches is that passive investing is optimal, and for that reason, a huge range of indexes have been developed. In this paper I revisit these ideas and conclude that the market is a minus-sum game, inefficiencies are almost always present and traditional asset pricing models should be reformulated. Indexing, as the passive investment approach tool, is seen not to be optimal and suffers from biases that affect its profitability and risk. These biases are detected in the paper and classified for future research. Finally, Ecology in Financial Markets is presented and used to give a detailed but qualitative analysis of a financial system in ecological terms.

1. INTRODUCTION

Financial theory has grown to include ideas such as the market, zero-sum games, efficiency and passive investment. Considering the importance of these issues, it might be assumed universally accepted definitions exist for them, however, the evidence suggests otherwise. After a thorough review of financial theory it is necessary, in our opinion, to debate essential financial concepts again. As a good friend once said to us, mentioning the philosopher Edgar Morin, "it is time the sciences rethought themselves[1]", and in my opinion, economics and finance should not be an exception. I would like to think the work presented here contributes someway to the rethink proposed by Morin.

In this paper I propose the revision of some fundamental pillars of financial theory. The first pillar I look at is the concept of the market, which is analysed and broken down into its constituent parts. An *Analytical Financial System* where only one asset is negotiated and a theoretical *Global Financial System* where all financial products are negotiated are presented for operative and classification purposes. The second pillar the analysis considers is the financial system as a non zero-sum game. The reasons for this consideration are clear. Commissions, fees, and advisors' or analysts' salaries are extracted from the game and have impact on it. The third pillar the article revisits is the concept of efficiency. The Efficient Market Hypothesis (EMH) derives from perfect competitive markets and assumptions regarding investors' rationality, non-correlation in irrational trades and arbitrage. Although efficiency as a concept has been accepted for decades, it seems difficult to sustain complete efficiency nowadays. People are not fully rational, they do not follow von Neumann-Morgenstern or Bayes rules in their predictions, they deviate from rationality using correlated strategies and arbitrage is limited due to transactional and holding costs or legal limitations. Empirically the efficiency theory has other problems. Predictability of prices, random walk hypothesis rejections, forecasting power of inflation and E/P ratios and periodical anomalies have been extensively documented. The fourth basic financial pillar is

1 Pensar Europa, *(1992) Conference, Barcelona.*

the CAPM. Modifications of basic hypotheses can include the introduction of transaction costs, multi-period, heterogeneous beliefs, information asymmetry or investors not following the mean-variance approach and these show very different results from those which most people are familiar with. Basically I can summarize them by saying that not all investors do hold the same portfolio and that the market portfolio is not located in the efficient frontier.

Financial theory, as based on these four pillars before, concludes that an investor can not consistently beat the market (the total market return), thus meaning that a passive investment strategy (buying and holding each security of the total market proportional to their capitalization) is the proper security mix that all investors should hold. To simplify this objective, market indexes were created to be followed for passive investors. The revision proposed here brings us to two different conclusions: the market portfolio, in contrast to what is defended in traditional finance, is not the proper mix of stocks which each investor must select, and a passive strategy following a market index is affected by profitability biases. I classify these into five groups: the sample bias (how the index proxies the market), the construction bias (how an index is calculated), tracking error, the efficiency bias (separation of the market portfolio from the efficient frontier), and the active bias (opportunity cost).

After analysing in detail the world which new finance is facing, I present here an alternative vision of financial systems that is based on Ecology. Ecology in Financial Markets seems a promising point of view which is still in its infancy but which can bring new ideas and approaches to financial analysis. I look at two lines of research within Ecology in Financial Markets. The first one is Biologically Inspired Algorithms (BIA), born metaphorically from diverse ecological sources, which provides academics with new tools such as Neural Networks, Artificial Immune Systems, Evolutionary Computation, and Social Systems. The second one is what I have called Financial Ecology, a branch of research that studies a Financial System as an ecosystem where different organisms and species live and compete for survival. A simple ecological qualitative analysis is provided in the article. I divide a financial system into four groups of organisms: Producers (who use financial systems for their own activities rather than speculation), Providers (who permit actor's access to the ecosystem), Organizers (who delimit and define financial systems' regulations) and Predators (who look for financial trading benefits in the financial market).

The structure of the paper is as follows. In Section 2 I revisit the traditional financial ideas that gave birth to indexing and analyze in detail the Market, Efficiency and the CAPM. This analysis is completed in Section 3 where problems with theoretical hypotheses are commented on and where market index's biases are also presented. Section 4 discusses Ecology in Financial Markets as a growing and promising approach to finance, paying attention to Biological Inspired Algorithms and to Financial Ecology. Section 5 gives my conclusions, Section 6 establishes future lines of research and Sections 7 and 8 contain the appendix and the references.

2. THE FOUNDATIONS OF INDEXING: TRADITIONAL FINANCIAL IDEAS

Traditional and modern financial theory has developed to include key ideas such as the market, efficiency, active or passive management. After lengthy research into financial knowledge and literature, I believe it is necessary to start a debate about some essential issues that are sometimes taken for granted. In the remainder of this article, I try to develop this analysis in order to get to the foundations of market indexes.

2.1. BREAKING THE MARKET DOWN

The market is the mother of our economic system and, of course, of finance. The market is everywhere, in every article, website or book about finance. But what is the market? Are all of us really speaking about the same thing when we say market? I try to clarify this concept below.

The *market* seems an ethereal individual governing all the actors[2] in the financial field. Market is a very general word with different meanings and nuances[3] that must be taken into account. Three different meanings can be given:

(1) The market is the physical or not physical 'place' where assets are negotiated. I call this idea the POE market, from physical or electronic market.

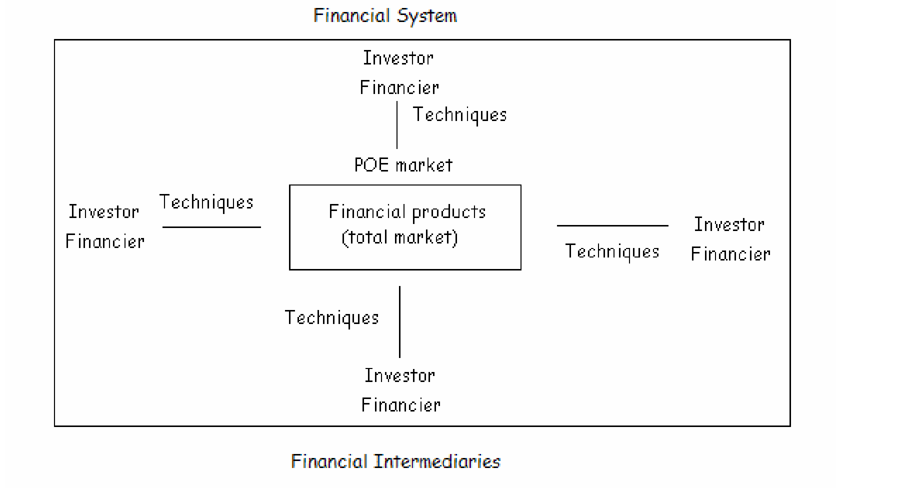
(2) The market is a means by which the exchange of financial products takes place as a result of buyers and sellers being in contact with each other, either directly or through mediating agents or institutions. This meaning of market is equivalent to financial system. A financial system is composed of products, investors (or financiers), POE markets, techniques and financial intermediaries.

(3) Market is a word used to refer to all the products that are traded in a POE market. This meaning of market is synonymous with total market.

Figure 1 enables us to better understand these ideas. The financial system [meaning 2] is composed of investors (and actors with funding necessities, i.e. financiers), who put their money into a POE market [meaning 1] (or who try to collect funds using the financial system) which is run, controlled, and intermediated by financial intermediaries[4]. Investors (and financiers) use more or less simple techniques (a thumb rule, a GARCH analysis, etc.) to buy or sell a combination of products, buying the total market as a whole [meaning 3].

2 Actor is defined here as every individual or institution participating in a financial system in one way or another. It is an overarching concept that includes investors, brokers, dealers and regulators among others.
3 Similar definitions can be found in The New Encyclopaedia Britannica, The New Palgrave or The Blackwell Encyclopaedia of Political Thought among others.
4 Here Financial Intermediaries is a general term that includes financial intermediaries and market supervisors.

Figure 1
The Market Idea broken down



Source: Authors' own

As it has been seen, there are 5 key elements in this scenario that must be studied: (a) Products, (b) Investors (or Financiers), (c) POE Market, (d) Techniques and (e) Financial Intermediaries.

(a) Products: Products are assets that are traded in a POE market. They can be securities, futures, options, commodities, etc. All the products of a similar category, treated as a block, are usually called the total market or simply the market [meaning 3]. It is shown in Figure 2. We have a group of assets with a price at a specific moment in time. This group of stocks will be different over time, because some will disappear and others will appear due to dynamics in the economy, appearing from new IPOs or from the primary market. There are lots of POE markets existing simultaneously, with a greater or lesser degree of relationship among them according to complementary of investors' asset allocation strategies[5] or legal restrictions among other possible reasons.

5 For example, the Bond POE market is related directly to the stock POE market. They share some investors and money at different points in time.

Figure 2
Total Market

Total market at t=1

$(x_{1,1};P_{1,1})$	$(x_{1,2};P_{1,2})$	$(x_{1,3};P_{1,3})$
$(x_{1,4};P_{1,4})$	...	$(x_{1,n};P_{1,n})$

Total market at t=t

$(x_{t,1};P_{t,1})$	$(x_{t,2};P_{1,2})$	$(x_{t,3};P_{t,3})$
$(x_{t,4};P_{t,4})$	...	$(x_{1,z};P_{1,z})$

Note: Where xij, is the total number of assets of type j at moment i, and Pij is the price of the asset j, at moment i. The assets will be different over time because some will appear and disappear due to economy dynamics, giving a total of n in t=1, and a total of z in t=t.
Source: Authors' own

(b) Investors (or Financiers): Financiers are actors who use the financial arena to collect money to cover their financial necessities. They are, for example, companies offering stocks or governments creating bonds. Investors, in contrast, are actors who trade in the financial arena, buying and selling the products that are available in that market, putting their money to work. This general concept lies behind different kinds of people and institutions with different necessities, strategies and goals. This idea of heterogeneity is absolutely essential to this analysis and will let us understand the market as a great mixture of different kinds of investors (and financiers), as opposed to the traditional homogeneous investor proposed by lots of financial models, who is totally rational and able to solve a complex stochastic optimization model [Shiller (2003)]. Investors are not always the same in either the same or in different markets. Some will appear in a market, whereas others will abandon a POE market for multiple reasons or will move between different POE markets. Investors can win or lose and they learn from their actions. I make a careful analysis of this point in part four.

(c) POE market: A POE market is equivalent to meaning (1) which I gave in the definition of market. It is the physical or electronic place where interaction between all participants in the financial system occurs.

(d)Techniques: These are the financial tools used by all the actors in the financial system to valuate assets, establish strategies, and determine prices and conditions in trades.

(e) Financial Intermediaries: These are institutions that deal with investors and financiers in a financial system. They rule the POE market, controlling it and acting as commissioners, agents or advisers. As we have seen, it is a broad concept that includes a lot of actors, not necessarily only brokers or dealers.

After defining the market properly and debating its usual meanings, it is necessary to continue the analysis. Two concepts are defined below: the analytical market and the global market.

Analytical Financial System (or Analytical Market)

I define an *Analytical Financial System* [or Analytical Market according to meaning (2)] as a financial system with a POE market where only one type of asset is negotiated.

There is not just one Analytical Market, there are lots of analytical markets in the same or in different countries. Some will share the same POE market, financial intermediaries, techniques, even investors and financiers, because some actors play simultaneously in different Analytical Markets. Some are similar to each other, whereas others are complementary (such as stocks and futures for example) or maintain more or less direct relations with each other. A Global Financial System is a general definition that must permit us to analyse a huge and complex reality as a unique market.

Global Financial System (or Global Market)

I define a *Global Financial System* (or Global Market according to meaning 2) as a theoretical financial system that includes all products, POE markets, techniques, financial intermediaries, investors and financiers. It is possible to define a Global Market within certain boundaries, such as a country or monetary region, or without boundaries, such as a Global Worldwide Market. There is no Global Market at present because there is no free access for all investors in all markets and there are legal restrictions and costs constraints, however, in the information technology era it seems preposterous not to think that we will head in that direction.

Before I continue with the article some additional considerations need to be taken into account. In the following sections, an analysis will be made first of an Analytical Market at time t (static), and then of an Analytical Market during a period of time (dynamic). After this, the same analysis will be made of a Global Market.

2.2. MARKET AS A ZERO-SUM GAME

One of the most familiar ideas in finance is that the market (that is, the financial system, meaning 2) is a zero-sum game[6]. Imagine an Analytical Market at a specific moment in time (t). Imagine a unique investor able to buy the total market, so he will assume total risk and obtain what is called the total market return (or simply market return). In an Analytical Market there is not only one investor, there are many, however, they can be summed up as acting as a unique investor holding the total market. Some investors will obtain higher returns than the market return, but this is offset by others who will earn less, and this distribution is totally random. This is the idea of a zero-sum game [Sharpe (1991)]. Sharpe demonstrates that in a market with active and passive investors (index followers), each type of investor holds the same holdings, so they will generate the same performance on average, that is to say, passive investors will get the market return, and active managers, who as a whole have the other part of the market, will obtain the market return too. In the words of Sharpe, "before costs, the return on the average actively managed dollar will equal the return on the average passively managed dollar, and, after costs, the average return actively managed will be less than the average return in a passive strategy". This

6 Game Theory is an interesting field that is out of our scope. For an introduction see Binmore (2008).

idea is very simple, but hides some interesting assumptions and problems that I will deal with later.

The zero-sum game idea is the same if we apply it in a dynamic framework, or if we apply it in a global market (statically or dynamically). What investor wins comes from another's losses, so passive investing is the most rational strategy to follow.

2.3. EFFICIENCY AND THE EMH

Efficiency is essential in finance and derives from the concept of perfect competitive markets. The first authors to deal with efficiency were Bachelier (1900), Kendall (1953), and Samuelson (1965) with his mathematical proof. Efficiency must be seen from two different points of view.

(1) *Operational efficiency* is related directly to how a POE market works, to transactional costs (brokerage commissions and spreads) and to competition between participants in the financial system. Without a good level of operational efficiency it is impossible to reach information efficiency.

(2) *Information efficiency* is related to the use of information in a financial system, and particularly, in determining prices in a market.

Jensen (1978) asserts that a market is efficient if it fully and correctly reflects all relevant information in determining asset prices. Formally, the market is said to be efficient with respect to an information set (Ω_t) if security prices are unaffected by revealing that information to all participants. Moreover, efficiency implies that it is impossible to make economic profits by trading on the basis of this information. Timmermann and Granger (2004) introduce search technologies (S_t) and forecasting models (M_t) to the definition. Thus, a market is efficient if it is impossible to make economic profits trading on the signals produced by a forecasting model that has been defined using predictor variables in the information set and selected using a search technology.

Fama (1970) focuses his analysis on the information set and proposes the Efficient Market Hypothesis (EMH). In his original paper market efficiency can be analysed at three different levels[7].

Weak level: This applies if the information set only comprises past and current asset prices. A market is efficient at a weak level if it is impossible to use the information set to obtain gains. In that case, Technical Analysis and other techniques that use past prices do not have forecasting ability. The market does not to have a memory.

Semi-strong level: This applies if the information set comprises all publicly available information. A market is efficient at a semi-strong level if it is impossible to use the information set to obtain gains. So, all the techniques that use public information, including fundamental analysis, are useless.

Strong level: This applies if the information set comprises all information, including private information. Again, a market is efficient at a strong level if even insiders can not benefit from the information set.

If information is distributed, it appears as a random variable and is incorporated directly into prices; it is possible to conclude that price formation is random and cannot be forecasted. This argument is based on some important assumptions:

7 These three levels of efficiency are analysed from a slightly different point of view in Fama (1991), and the aforementioned return predictability, event studies and private information analysis.

(1) Investors rationality[8]: When investors are rational they value securities for their fundamental value[9]. Moreover, they are able to analyse all the alternatives and establish a ranking of preferences among them.

(2) If (1) is not fulfilled, irrational investors trade randomly and cancel each other out because their actions are not correlated.

(3) Arbitrage:[10] If case (1) or (2) fails, there are rational arbitrageurs who eliminate mispricing.

What follows from all these definitions is that EMH rules out the possibility of trading systems based on current information that sets expected returns in excess of the equilibrium expected returns. Put more simply, an average investor can not consistently beat the market (the total market return). This analysis is useful for any Analytical or Global market, from a static or dynamic perspective.

Concluding, the EMH reinforces that information analysis is useless. Due to the random characteristic of prices, the optimal strategy for all investors is to get the market return (the average return of the total market). This can be done using a passive buy and hold strategy, where each investor must replicate the total market in his portfolio, as can be seen in Figure 3, and here is where the importance of indexes appears.

8 Gulko (2004), however, states that this argument would not be necessary for efficiency.
9 This is the net present value of its future cash flows discounted using their risk characteristics and probabilities.
10 Perfect arbitrage is defined as an activity where informed traders use inefficiencies to generate operations without risk and with a sure return.

Figure 3
Optimal Strategy in Traditional Finance

Market is a zero game
EMH
CAPM, APT

⇓

Efficient Market
Impossible to beat
Impossible to predict

⇓

Optimal Strategy
Buy and hold the total market
Total market indexes with capitalization weightings (tracking error)
Sampled indexes (sample and construction biases, tracking error)

Source: Authors' own

2.4. THE CAPM, EFFICIENCY AND MARKET

The classical CAPM [Sharpe (1963, 1964), Lintner (1965) and Mossin (1966)] is defined in a dynamic one period Global Market that does not includes financial products alone, and has been one of the most important approaches in finance. One of the attractions of the CAPM is that it offers powerful and intuitive relations between expected return and risk. The CAPM makes Markowitz's original efficient frontier (EF) linear and identical for all individuals [Lintner (1965), Tobin (1958)], although this result is related to the assumptions of the model, as can be seen in part 3, basically with the assumption of identical probability judgements (homogeneous expectations). The difference between investors is the amount invested in the risky asset (the market portfolio) and the risk-free asset. Again, as with the zero-sum game and the EMH, the CAPM concludes that the market portfolio[11] (i.e. buying and holding each security of the total market proportional to their capitalization) is the proper mix among risky securities for every investor

11 The tangency optimal portfolio derived from Tobin's (1958) separation theorem, complemented in Lintner (1965), Mossin (1966).

[Mossin (1966)]. But this support of passive investing goes one step further with the CAPM: a selected portfolio (an index that represents the total market) is used to calculate risks and expected returns in equilibrium. Market indexes, that is, proxies of the market, thus become essential to traditional financial thinking. Given that this matter is of vital importance, it is odd to say the least that no authors have carried out an accurate analysis of them. I aim to redress this with this study.

2.5. THE APT MODEL

The Arbitrage Pricing Theory was developed by Ross (1976), and Roll and Ross (1984). This model develops an interesting idea that reinforces some traditional concepts: investors' rational behaviour exhausts arbitrage opportunities. The model describes security returns as a single-or multifactor model, giving results that are compatible with the CAPM approach, but with less restrictive assumptions regarding investors and the market. What is quite new and useful in this approach is that factors are not known a priori, and the market is not imposed as the most important factor. The principal components approach is an appropriate technique for finding an approximate factor structure [Chamberlain and Rothschild (1983)]. Factor analysis has also been used since the first developments [Roll and Ross (1980)].

2.6. INDEXES AND PASSIVE INVESTMENT

Traditional financial approaches are clear about one thing: passive investing is the optimal strategy a rational investor must follow if he is in a efficient zero-sum game market and CAPM hypotheses are fulfilled (I call it a perfect market). A passive strategy means buying and holding the total market, and because this is impossible, we need something that represents (or proxies) the total market: a market index.

Most of the indexing development has been done to proxy a concrete stock analytical market. Indexes were born directly from the blind belief in the market as a zero-sum game (efficiency and modern portfolio theory) and they rapidly spread around the world. Two important issues must be outlined:

First, a good market index must be composed of all the possibilities or investments that are available in a selected market. In this line of thinking total market indices are the market, and are really carrying out the main function we want for an index. Total market indices were impossible to construct some years ago because computational technologies were not as powerful as they are today. A good sample of a population must be representative of the population, and consequently index builders have developed sampled-indexes able to control 60%-90% of the capitalization of the market[12] or 60-90% of the traded stocks.

Second, a passive investor wants to obtain the average return,[13] as if he or she holds the whole market. The average arithmetic return of the market and the arithmetic return of the index will be the same only if the total market index is calculated using capitalization weighting.

Although market indexes have been developed traditionally to proxy a specific stock analytical market, some investigations and developments have been made in other directions. Some attempts include Fixed-Income indexes, Hedge Funds or Real Estate indexes that proxy different analytical markets. In addition, the Global Market has not been free of attacks from indexers. There are some apparently global indexes trying to cover a more or less global stock market [Schoenfeld (2004)]. Finally, it is important to analyse whether indexes have been developed by professionals, by investment banks as products to offer to special clients or as benchmarks. Indexes symbolize some institutions, represent countries, zones, sectors, industries, style or social policies and are built with alternative weighting criteria and float adjustments. These indexes can be interesting to some investors,

12 For example, the S&P500 includes about 75% of the US companies by capitalization.

13 $$R_{t,m}=\frac{(x_{t,1}\cdot P_{t,1}+x_{t,2}\cdot P_{t,2}+\ldots+x_{t,n}\cdot P_{t,n})-(x_{t-1,1}\cdot P_{t-1,1}+x_{t-1,2}\cdot P_{t-1,2}+\ldots+x_{t-1,z}\cdot P_{t-1,z})}{(x_{t-1,1}\cdot P_{t-1,1}+x_{t-1,2}\cdot P_{t-1,2}+\ldots+x_{t-1,z}\cdot P_{t-1,z})}$$

or can cover some commercial necessities, but are quite removed from the foundations and main objectives of an index.

3. TRADITIONAL FINANCIAL IDEAS UNDER REVIEW

3.1. THE MARKET: A MINUS-SUM GAME

The market is not a zero-sum game: what some people lose does not go completely to others, and this is true in a static or dynamic framework and for an analytical or global market.

First of all, there are commissions and fees coming from brokers, dealers and stock exchanges. There are advisors, analysts and fund managers that live for and from the investors’ profits or losses. If the whole market belonged to a single investor only, commissions would be almost irrelevant, but this is not the case. Adding together all investors’ activities will not give us the same result. Market is a minus-sum game, and set up so that most traders must lose money [Elder (1993)].

Secondly, the idea of a zero game can be interesting from a theoretical point of view, or from a traditional or simple analysis of the average investor. But it is not of interest at all from the point of view of a private investor, institution, or fund manager. In a minus-sum game there are two types of actors: financial intermediaries that always win, and investors, who can be classified as winners and losers[14]. However, is it possible to analyse the distribution of winners and losers to draw more specific conclusions? Certain items can permit us to establish specific characteristics of the winners that separate them from the losers, as we show in point 4 of this paper.

14 Winners: investors who will obtain higher returns than the market return taking into account commissions and costs. Losers: investors who will obtain lower returns than the market return.

Finally, in Sharpe (1991) there are two key issues: i) active managers cannot outperform the market without some investors paying added costs via inferior performance. Indeed, this is the point which is going to be developed. ii) Passive investors build their portfolio using the benchmark index which is assumed to be an accurate representation of the total market. As we will see subsequently, only a few indexes can be assumed to really represent the total market, and even supposing they do, if a market is not efficient, a passive strategy is not optimal at all.

3.2. EFFICIENCY ON THE ROPES

Efficiency is one of the key issues in finance; therefore, it is not difficult to say it creates controversy. Efficiency, as a theoretical idea, provides an interesting framework in which to build some theoretical models, however, in practice, as in other applications of perfect competitive markets to reality, there are some problems that must be outlined.

First of all operational efficiency is not the same in all markets. It is very easy to see that transactional costs and brokerage commissions are very different across markets and countries. This has an important effect on efficiency. High transactional costs and technological or legal restrictions in trade make informational efficiency difficult, but also, in contrast, make it too expensive or impossible to use even large predictable patterns are detected [Timmermann and Granger (2004)]. Likewise, a market could be very inefficient, but this inefficiency can not be exploited due to its own inefficiency if the mispricing does not go above or below certain limits [Pontiff (2006)].

The second point to analyse is information efficiency. From a theoretical point of view, there are some problems with the EMH hypothesis. These problems have recently been analysed by Behavioural Finance [Shleifer (2000) for example]:

(1) Investor full rationality is difficult to sustain. People are not fully rational and underprice and overprice securities, constantly overreacting to extreme bad or good news [Debondt and Thaler (1985, 1987)][15] . Many react to irrelevant information or trade using noise [Barber and Odean (2000)]. Investors (especially individuals but also institutions) follow analysts' recommendations or their neighbours' trading system. Some also hold losing positions and are unable to sell due to regret feelings [Odean (1998)] or buy and sell managed funds actively and expensively. Investors follow fashionable gurus and behave more like children running after a ball than as rational optimizers. Sometimes these investors are called noise traders [Kyle (1985) or Black (1986)] but it is not clear if there is any smart money in contrast. In other words, people deviate from standard decision making models in some important areas. First, in their attitude toward risk: regarding trading as a gamble does not follow von Neumann-Morgenstern's rationality model. Second, individuals systematically violate Bayes rule and other maxims of probability theory in their predictions of uncertain outcomes [Kahneman and Tversky (1973)] and make systematic errors of judgment and probability assessment [Kahneman, Slovic and Tversky (1982)] based on psychological biases [see Edwards (1968), Lo *et al.* (2005) or Lo (2004) for more information]. What is clear is that Homo Economicus has started to die out in finance.

(2) If investors are not fully rational they trade randomly and cancel each other out, because they are not correlated. The psychological evidence does not seem to support this point. People, noise traders and also smart money deviate from rationality, but deviate in the same way, using correlated strategies in price bubbles or panic movements, moving sometimes as a herd. This problem becomes more severe when investors trade by listening to rumours or by following their neighbours, and in some theoretical models and empirical notes it can be seen that these movements cannot be eliminated through arbitrage [Shiller (1984)].

(3) Imagine the first two assumptions fail. EMH has the last weapon of arbitrage to prevent prices from moving away from fundamental values. Arbitrageurs, as

15 Zarowin (1989) states that this overreaction could be derived from size effects.

the last rational investors, exploit little inefficiencies, earn money, and maintain efficiency [Keane (1986)]. But, arbitrage, in the real world is not so powerful and easy to carry out. The effectiveness of arbitrage relies firstly on the availability of close substitutes. Sometimes securities do not have obvious substitutes, or they are far from perfect. Secondly, arbitrage has transactional[16] and holding costs[17], as well as legal limitations, that make this activity costly and risky. Finally, arbitrageurs are finite, in number and in wealth, because real arbitrage implies money and risk. With a real risk in arbitrage activities, and risk aversion, the arbitrageurs' capability to control prices near to fundamentals is limited. This limited capability is not able to prevent mispricing from becoming worse in the long term [Zhang (1999), Shleifer (2000), Shiller (2003) or Delong *et al.* (1990)]. Finally, remember most arbitrageurs do not use their own money and are usually evaluated in the short term. Your activity is limited when you have to defend your job and an arbitrage activity becomes negative. In a world of imperfect arbitrage, financial markets rather than being self-correcting, exhibit what is referred to as hysteresis effects [Hutchinson (1999)]. Moreover, irrational investors may not necessarily be eliminated by the market or arbitrage and they could survive [Hirshleifer and Luo (2001)].

Anomalies

In the previous paragraphs the paper studies some of the theoretical challenges of the EMH. It is time to look at efficiency empirically. The majority of empirical studies have been applied to stock markets, so I will start by using this analytical market. According to Jensen's (1978) definition of efficient markets, a trading strategy producing consistent risk-adjusted economic gains, after properly defined transaction costs and over a sufficient long period of time, is evidence against the EMH, at least, in a specific context and period of time. During the last thirty years, a lot of anomalies have been reported, and EMH is also on the ropes empirically. In an attempt to bring together this enormous literature, I classify them using Fama's three levels of information [Fama (1991)]: a) return predictability, b)

16 Transactional costs have received attention from theories of mispricing.

17 Holding costs and idiosyncratic risk assumed by an arbitrageur have been analysed by a few authors, for example, Pontiff (2006).

event study and, finally, c) private information. For further references see Figure 4a, 4b, and 4c in the appendix.

(a) *Return predictability*: Financial returns seem predictable in the short-term and perhaps in the long term[18]. Moreover, some theoretical models [Shiller (1984), Summers (1986)] make it possible to predict long-horizon returns from past returns. Looking at the literature in detail, return predictability was associated with the Random Walk Hypothesis (RWH). If a time series did not reject RWH, it was considered not predictable, and weak efficiency in the market was assumed[19]. As a result of this idea, a lot of studies have been carried out and since the first Cowles and Jones's (1937) test; other methodologies have been used for that purpose: the ADF, GPH, LOMAC and RS tests are the most common examples.[20] Countless empiric analyses have been done on RWH and from the earliest studies[21] to the most recent[22], different methodologies and alternative tests have been applied to US and international data, and to emerging markets.[23] Despite this huge amount of literature, the results have been mixed, and acceptance and rejection of RWH are almost equally represented. Discussion has opened up, and methodologies to control thin trading (at least in emerging markets), non-linearities and structural breaks have been developed to try to create a framework for comparing results. Finally, another key point in the efficiency battle has been volatility analysis. According to this, stocks seem to show an excess of volatility facing efficient markets [Shiller (1989)].

Leaving these controversial results on one side for a while, these are not the only ways in which EMH has been challenged. There are lots of articles about the forecasting power of inflation, short-interest rates, E/P ratios, or other financial

18 Lo and MacKinlay (1988), Conrad and Kaul (1988), Poterba and Summers (1988), Fama and French (1988), Fama (1991) or Fisher (1966) for a complete discussion of these results.

19 Low-dimensional non-linear deterministic systems may generate price behaviour as apparently random series. [See Hinich and Patterson (1985), Summers (1986), Hsieh (1991), Scheinkman and Lebaron (1989), Frank *et al.* (1988), Opong *et al.* (1999), Yadav (1999) and Jasic and Wood (2006) for non-linearity studies].

20 See Figure 4b for further information.

21 See for example Kendall (1953), Osborne (1959, 1962), Roberts (1959), Working (1960), Larson(1960), Cowles (1960), Alexander (1961, 1964), Granger and Morgenstern (1963), Mandelbrot(1963), Fama (1965a, 1965b, 1970), Samuelson (1965), Leroy (1973), Cowles and Jones (1973), Rubinstein (1976), Lucas (1978).

22 Fama and French (1988), Poterba and Summers (1988), Lo and MacKinlay (1988), Kim *et al.* (1991), Richardson and Stock (1989), Richardson (1993), Richards (1997), Balvers *et al.* (2000).

23 See Figure 4c for further information about RWH in emerging markets.

variables [see Solnik (1993) among others]. Moreover, the existence of anomalies such as the January effect, weekend effect or small firm effect [Fama (1991)], profitable momentum, tactical asset allocation or technical trading strategies is well documented [Jegadeesh and Titman (1993), (2001), Chan, Jegadeesh and Lakanishok (1996), Arshanapalli *et al.* (2004), Brock *et al.* (1992) and Blume *et al.* (1994) for example]. Finally, recent articles [Leung *et al.* (2000)] have shown how forecasting the signs of how an index is going to move could be easier and more profitable than all the strategies mentioned up to now.

(b) *Event studies*: Sometimes prices do not seem to adjust quickly to information. Event studies show how good news, bad news, corporate movements or other similar events affect prices. Some of the most puzzling problems start in event studies. Some authors have analysed how changes in dividends [Charest (1978)], mergers [Asquith (1983)], takeovers [Keown and Pinkerton (1981)], inclusions of stocks in an index [Shleifer (1986)] or earning reports can affect prices [Jegadeesh and Titman (1993)].

(c) *Private information and insider trading*: The use of private information is usually analysed using mutual funds. Some studies conclude that mutual funds have and use private information [Henriksson (1984), Chang and Lewellen (1984)], but others conclude that pension plans earn less than passive funds [Beebower and Bergstrom (1977), Ippolito and Turner (1987)]. These contradicting results could be due to some differences in the methodology used that may have been solved recently. Seyhun (1998) or Jeng *et al.* (1999) show the ability of insider traders to get abnormal returns. What is clear is that laws against insider trading have been made for a reason.

In conclusion, both perfect efficiency and perfect inefficiency are difficult to assume from a theoretical or empirical point of view. Marginal efficiency [Zhang (1999)] or adaptive efficiency seems a more plausible approach. Anomalies will constantly appear in the market, some will disappear after becoming public if the market is adaptively efficient [Timmermann and Granger (2004) and Dimson and Marsh (1999)], simply because they are spurious, or because they attract sufficient capital to exploit the predictable pattern, but others will never disappear because they are part of the financial system, and bound by legal constraints or acts of investors, and will create pockets of inefficiency that can be exploited [Daniel and Titman (1999)]. These pockets of inefficiencies will be greater in less developed markets, where operational efficiency is lower or where there are higher legal constraints, but will also be present in developed countries if they are born out of investors' irrational biases. If a market is not efficient, economic theories on security pricing such as CAPM or APT may become too weak, and psychological concepts or ecological models may do a better job.

Although most authors have analysed efficiency and anomalies in stock markets, it is possible to find efficiency studies in other analytical markets such as bond markets,[24] future bond markets,[25] derivatives markets,[26] exchange rates,[27] real estate markets,[28] and with stock market indices[29]. Conclusions from these studies are also not clear. Some papers find evidence of efficiency in a concrete market while others reject the efficiency hypothesis, concluding that there are arbitrage opportunities, delays in price responses or overreactions.

3.3. THE CAPM, PROBLEMS IN REALITY

Using CAPM original literature [Sharpe (1963, 1964), Lintner (1965) and Mossin (1966)] and posterior one [Markowitz (2005), Hutchinson (1999), Elton and Gruber (1991), Haugen (2001), Lundgren (2005) and Perold (2004)] CAPM assumptions have been synthesized as follows:

1. Transactions costs and other illiquidities can be ignored (a frictionless market where arbitrage plays an important role).

24 [Rendleman *et al.* (1979), Hotchkiss and Ronen (2002), Halpern and Rumsey (1997), Hall and Miles (1992)]
25 [Tse (1999), Becker *et al.* (1996)]
26 [Zhang and Lai (2006), Cassese and Guidolin (2004), Bell *et al.* (2007)]
27 [Liu and He (1991), Larsen and Lam (1992), Bekaert (1996) or Jasic and Wood (2006)]
28 [Kleiman *et al.* (2002)]
29 [Brock *et al.* (1991), Frank *et al.* (1988) and Opong *et al.* (1999)]

2. Assets are infinitely divisible and perfectly liquid.

3. The only two decision parameters are risk and return (in terms of expected return and standard deviation), and all investors hold mean-variance efficient portfolios.

4. Investors are risk averse.

5. All investors operate in the asset market as price takers, and have the same investment opportunities.

6. The investor's horizon is identical for all investors (one period).

7. All investors are well informed, and have the same (and correct) beliefs about means and variances.

8. Everybody can lend or borrow at the risk-free rate, or in a more extended CAPM approach, investors can buy or sell short without limits.

9. The market (total market idea) is a Global market. Moreover, it includes all risky assets available, and not just traded financial assets.

The results of the basic CAPM are absolutely related to its assumptions, so if some of them are not fulfilled[30] the results will be in deep trouble. In particular: i) There are transaction costs that make market a minus-sum game and limit arbitrage activities. Taxes affect investors and their investment decisions because an optimal management involves realization of losses and deferral of capital gains

30 Haugen (2001).

[Constantinides (1983)]; ii) Shares are almost, but not infinitely divisible; iii) Not all investors use mean and variance to establish their portfolios and almost nobody maintains optimal portfolios. Moreover, investors could only choose between portfolios using a mean-variance approach if probability distributions were normal, or in case normality fails, the utility function was quadratic [Tobin (1958)]. Normality is not usually the case in financial data, and a quadratic function is difficult to maintain[31]; iv) Investors are risk averse, but risk seeking behaviours have been reported and risk preference is not stable over time and in specific individual situations [Lo (2004)]. Some personality assessment methods are being used to examine the relationships between specific personality traits and risk-taking in different domains [Nicholson *et al.* (2002)]; v) Not all investors act as price takers, at least not in all analytical stock markets. In addition, not all investors have the same opportunities or assets to invest in; vi) The investment horizon is absolutely different for each investor; vii) Not all investors are well informed, and they do not have the same predictions for expected returns, volatilities and correlations. If this was the case, trade will be almost insignificant and the markets will collapse [Grossmann and Stiglitz (1980)]. Unfortunately (or fortunately) information is not spread homogeneously and asymmetric information exists. Indeed, people have different expectations about returns and risk, and they evolve over time because individuals process new information, sometimes incorrectly [Debont and Thaler (1987)]; viii) Investors can not lend and borrow at the same risk-free rate, and there are limitations to selling short; ix) Finally, the market portfolio in the CAPM is a hybrid portfolio embracing all risky assets available and not just traded financial assets, (including shares, bonds, other financial and non financial assets, consumer durables, real estate and human capital). However, this point is missed sometimes, the market portfolio being thought to comprise stocks only [Hutchinson (1999), Fama and French (2004)].

31 It is difficult to maintain the utility function reaches a maximum at some wealth, and it is really contrary to human behaviour that when wealth increases, the propensity to take on risk decreases.

Modifications of the CAPM

As it has been seen previously, assumptions in the CAPM model are quite unrealistic. To complete the basic model, the CAPM has been extended in various ways.

From a conceptual point of view, some of the best-known extensions include for example: heterogeneous beliefs for investors [Lintner (1969), Merton (1987), Haim *et al.* (2006)], short sales without limit instead of risk-free lending and borrowing [Black (1972)]), non-markable assets [Mayers (1973)], multiple periods [Fama (1970b), Merton (1973), Elton and Gruber (1975), Stapleton and Subrahymanyam (1977), and Breeden (1979)], irrational investing [Solnik (1974), Stulz (1981), Adler and Dumas (1983)], consumption orientation [Breeden (1979), Rubinstein (1976)], inflation [Friend *et al.* (1976)], multibeta approaches [King (1966), Gomez and Zapatero (2003)], or weaker assumptions relying on arbitrage pricing [Ross (1976)]. Conditional versions [Andersen *et al.* (2003) or Soydemir (2005)], downside betas in the D-CAPM [Estrada (2002)], GARCH models [Fraser *et al.* (2004)], Stochastic Betas [Joslova and Philipov (2005)] or the Best-Beta CAPM [Zou (2006)] are perhaps the latest approaches in this never-ending race. It is important to note here that when the traditional hypotheses of the CAPM are not fulfilled, investors hold different portfolios, therefore, the market portfolio is not the optimal and only portfolio an investor must hold [Lintner (1969)] and some other results are a little different from the traditional CAPM. Even Sciubba (2006) has shown believers of the CAPM can be driven to extinction in an evolutionary framework.

Empirically, the CAPM has been challenged too. Although the early tests rejected the Sharpe-Lintner version [Douglas (1968), Black, Jensen and Scholes (1972), Miller and Scholes (1972), Blume and Friend (1973)], betas more ore less seemed to explain expected returns in lots of papers from different countries, data bases, data periods and markets, so the popularity of the CAPM grew rapidly. Moreover, posterior modifications provided improvements in some cases increasing the reputation of the CAPM [Fama and McBeth (1973), Stambaugh (1982), Reilly and

Wright (1988), Chen (2003), Balyers and Huang (2004), Soydemir (2005), Hyde *et al.* (2005), Fletcher and Kihanda (2005), Fernandez (2006) and Zou (2006) to mention only a few]. To conclude, despite these controversial results, the CAPM fails to describe average realized stock returns since the 1960s with US or UK data [Jagannathan and Wang (1996), Strong and Xu (1997), Campbell and Vuolteenaho (2004) or Fama and French (2006)]. This is not the only problem the CAPM has to face: i) There is no consensus with respect to the index, time frame or data frequency that must be used for estimation and beta's instability has been a problem since the first analyses [Blume (1975), (1979)]. ii) Betas seem to be more useful for explaining expected returns in individual stocks than in portfolios [Bartholdy and Peare (2005)]. iii) The CAPM is unable to explain the Value premium[32] or Small Firm effect[33].

At this point it is necessary to recall that, on balance, not all investors hold the same portfolio, so the market portfolio is not located in the efficient frontier. Even if security market prices efficiently reflect the best estimation of future cash flows, and even if all investors rationally optimize the relationship between risk and expected return, which as we have seen, is not the case, the market portfolio is likely to be inefficient.[34] The conclusion is obvious: the market portfolio, in contrast to what is defended in traditional finance, is not the proper mix of stocks that must be selected by an investor, and a passive strategy cannot be an optimal strategy. Moreover, the "market" as represented by an index is not perhaps the only important factor that explains returns, as is defended in the CAPM, and that therefore there should be discussion regarding the role of market indexes as proxies of the market and factors in asset pricing models.[35]

32 Excess of expected return for value stocks for US and international stocks was documented in Fama and French (1992, 2006).
33 Excess of expected return for small stocks.
34 Similar approaches in Haugen and Baker (1991), Haugen (2001) and Markowitz (2005).
35 Only some authors test the CAPM using a range of market portfolios [Stambaugh (1982)]

3.4. THE APT MODEL

The APT model is based on two key assumptions: rationality and arbitrage. Leaving aside for a while the problems of these hypotheses which have been outlined above, the APT model helped to develop the blind belief in market efficiency and rationality that has dominated traditional finance. Empirically and theoretically the APT plays a better role than the CAPM, and is the second most widespread model among academics and practitioners. According to previous conclusions in this paper, the APT, which does not force a market index to become an essential part of the model, seems a better approach when analyzing the market.

3.5. INDEXES, USES AND ABUSES

Indexes

In a perfect market,[36] a passive strategy is optimal. Passive means buying and holding the market portfolio using usually a total market index or a well sampled capitalization weighted index. In such a situation, an investor will hold an efficient portfolio. But indexes have gone one step further. Nowadays some cover a selected market, sector or industry. They are calculated using different weightings, or by trying to reflect an investment strategy or allocation policy rather than the market. And finally, there are consultants who decide which stocks enter or exit an index according to more or less clear criteria, and not necessarily according to representativeness. None of these apparently inoffensive items have been treated at length in either the academic or professional fields, but they are indeed important, because they create biases in the performance that a passive investor is trying to obtain. Using the deviation source from the market return and the efficient frontier, I classify these biases as:

36 Efficient zero-sum game market, where hypotheses of the CAPM are fulfilled.

1. *The sample bias*: This is the difference between an index return and the real total market return due to the selection of the sample of stocks that make up an index. It is a statistical bias between population and sample. It could also be called a *representation bias*[37.] One important sample bias is the *selection bias* which could be crucial in some indexes. Stock inclusions or exclusions do not always follow criteria looking for a better sample of the total market.[38] Depending on the selection criteria, an index can be more an active rather than a *passive investment*.

2. *The construction bias*: This is the difference between the index return and the real total market return due to the use of different weighting criteria (price, capitalization or GDP) or methodologies (Laspeyres, Paasche or Geometric mean among others) in the construction of an index.

3. *Tracking error*: This is the difference between the index return and the return obtained by a passive investor following the index due to commissions and turnover costs. Indexes are calculated without fees or transaction costs, but it is clear that a passive investor does not have this privilege if he or she tries to form their own portfolio.

Efficiency

Financial markets are not perfect, and other biases must be analysed.

4. *Efficiency bias*: If a market is not perfect, the market portfolio can not be optimal or efficient. *Efficiency bias* is the difference between the market portfolio return (which is not optimal in this case) and the return that the efficient frontier pays for the assumed risk. The efficiency bias can also be defined in terms of risk.

37 Something similar is presented in Hedge Fund Index analysis in Liew (2003).

38 For example, in any S&P500 exclusion there are 10 to 20 candidates to replace it. The decision of inclusion will depend on reasons as if the candidate is on an underweighted or overweighted industry, if it comes from a recent IPO, etc. [Gastineau (2002b)]

This is the difference between the market portfolio risk and the efficient frontier risk according to the obtained return.

5. *Active bias*: If a market is not perfect some of the traditional hypotheses do not fit reality. If transaction costs create tax harvest strategies, investors play away from the mean-variance approach or are affected by psychological biases, and pockets of inefficiency can appear in the market. In this situation an active strategy could be more profitable than a passive investment strategy. In that context, active bias is defined as the difference between the real market return and the best opportunity for investment using the best active strategy (opportunity cost). The optimal strategy in a marginally efficient market could be a combination of indexing and active management, but this remains to be verified. Once this theoretical bias has been established, another question is whether it can really be profitable and usable. Both aspects will be analysed subsequently.

Figure 5 sum up the strategy in an imperfect market. Not taking into account all the biases detected can generate undesirable results in the performance measurement of active and passive strategies and in the uses of indexes for asset pricing models. In the CAPM, a capitalization-weighted index is useless for explaining reality if investors and markets go away from the equilibrium predicted by the model. In that case, the predictability of the model can be seriously affected by the used market index depending on the sample, selection criteria, construction methodology or active strategy it hides behind [Bartholdy and Peare (2005)].

Figure 5
Optimal Strategy in a Marginally Efficient Market

Market is a minus-zero game EMH fails CAPM hypotheses are not fulfilled

⇓

Not Efficient Market Not Impossible to beat Not Impossible to predict

⇓

Optimal Strategy
Actives strategies could be useful
Total market indexes with capitalization weightings would not be optimal investments (tracking error, efficiency and active biases)
Sampled indexes (sample and construction biases, tracking error, efficiency and active biases)

Source: Authors' own

4. ECOLOGY IN FINANCIAL MARKETS

So far I have summarized principal traditional ideas in finance. We have also seen criticisms of the traditional framework and the implications when hypotheses of the models are not fulfilled. A lot of literature has been written and will be written on the areas I have commented on. Nevertheless, we are living in a key moment that will bring us a new paradigm in finance. I call this new and increasing paradigm *Ecology in Financial Markets or the Biological Approach* and I will break it down into its constituent parts in the following paragraphs. This new paradigm can be included in a more general framework called Ecological Economy that has been growing in acceptance from the 70s[39].

The application of evolutionary ideas to economics is not totally new. Petty, Cantillon or Quesnay, who established the basis for classical economy, presented

39 See Cuerdo and Ramos (2000) for further information.

economic arguments related directly to nature. Thomas Malthus used biological arguments in his studies, Schumpeter's (1939) business cycles are similar to natural selection, Wilson (1975) applied principles of competition, reproduction, and natural selection to social interactions[40], and Lo put forward other applications in economics (2004, 2005). But until recently, the *Biological Approach* has not consistently been used in finance. I divide it into two important and different lines:

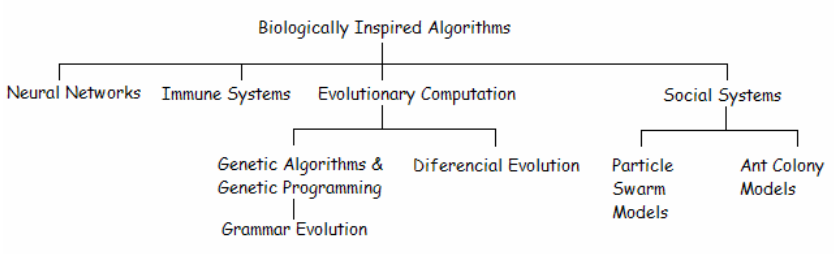
The first line is inspired by the biological systems and their functions to create mathematical applications capable of solving problems in a similar way to nature's way of solving them. I call this approach *Biologically Inspired Algorithms*. The second line sees a financial system as an ecosystem where species live and compete for survival. I call that approach *Financial Ecology*.

4.1. BIOLOGICALLY INSPIRED ALGORITHMS

Biologically Inspired Algorithms (BIA) are born metaphorically from diverse sources including evolution, models of social interaction among organisms, biological neurons and natural immune systems. In recent years, taking natural processes as its inspiration, a considerable literature has emerged, and complex ever changing financial markets where populations compete for profit and survival are also there to exploit the power of BIA. It is not the aim of this article to exhaustively analyse all these methodologies, I only intend to give a general overview of BIA.

40 Sociobiology (named by Wilson) has been a productive line of research, see Barkow *et al.* (1992), Pinker (1993, 1997), Crawford and Krebs (1998), Buss (1999), Gigerenzer (2000) and Trivers (1985). For an interesting review of evolutionary psychology, see Barrett, Dunbar and Lycett (2002).

Figure 6
Biologically Inspired Algorithms (BIA)



Source: developed by the author on the basis of Brabazon and O'Neill (2006)

As Figure 6 shows, BIA can be divided into four main groups:

a) *Neural Networks*: Artificial Neural Networks[41] is a modelling methodology inspired by the workings of the human brain and can be constructed for purposes of prediction, classification and clustering. With this non-parametric modelling tool there are three main items to bear in mind: i) connection topologies, ii) training methods, and iii) learning algorithms. These have recently been applied to financial markets with interesting results [see, for example, Baestanes et al. (1994), Hutchinson *et al.* (1994), Gately (1996), Ruggiero (1997), Zirilli (1997), Franses and Van Homelen (1998), Hu *et al.* (1999, 2004), Yao and Tan (2000) or Fernandez-Rodriguez *et al*. (2000)].

b) *Immune Systems*: Artificial Immune Systems are inspired by the natural immune system. The processes used in these applications are a negative selection, based on the ability of the immune system to discriminate between self and non-self, and the clonal expansion-selection algorithm, inspired by the clonal selection and affinity maturation process of B cells. A more detailed explanation and financial applications can be found in Brabazon and O'Neill (2006).

41 General information can be found in Patterson (1996) and Torra (2004).

c) *Evolutionary Computation*: In Evolutionary Computation (or Evolutionary Algorithms) solutions evolve according to rules of survival, crossover and mutation. Some interesting applications in this group are Genetic Algorithms, Genetic Programming, Grammar Evolution and Differential Evolution. Genetic Algorithms[42] adopt a population unit of analysis where each member encodes a potential solution. Evolution in the population is simulated by means of pseudo-natural selection according to the quality of the solution and these solutions are disturbed by crossover and mutation in an attempt to uncover better solutions. Genetic Programming[43] (similar to Genetic Algorithms) is applied directly to solutions and not to indirect encoding solutions. Grammar Evolution[44] can evolve computer programs, rule sets or sentences in any language using molecular biologists' ideas as the genotype and phenotype. Differential Evolution[45] is inspired by the field of evolutionary computation and social algorithms. Good solutions are evolved and affected by mutation, recombination and fitness-based selection. A substantial body of literature applies evolutionary methodologies to financial markets with promising results [Bauer (1994), Deboeck (1994), Colin (1994), Neely *et al.* (1997) or Varetto (1998)].

d) *Social Systems*: Social Systems (or Social Models) exhibit flexibility, robustness, self-organization and communication between individuals. The two most popular models are Particle Swarm Models and Ant Colony Models. Particle Swarm Models (or Particle Swarm Optimization Algorithms) present particles which encode solutions moving in an n-dimensional search space in an attempt to uncover better solutions. Each particle has a memory of the best location it has found and also knows the best location found to date by all the particles in the population. The size and direction of each particle's movement is a function of its own history and the social influence of its peer group [Kennedy *et al.* (2001)]. Ant Colony Models are a family of population-based, optimisation and clustering algorithms that are metaphorically based on the activities of social insects. Each individual follows a limited set of rules but the interaction of the activities of these individuals gives rise to a complex emergent, self-organised structure [Bonabeau

42 More information is available in Holland (1975), Goldberg (1989) and Mitchell (1996).
43 See Koza (1992, 1994).
44 See O'Neill (2001) and O'Neill and Ryan (2001, 2003).
45 See Storn and Price (1995, 1997) or Price (1999).

et al. (1999)]. Financial applications of social models can be found in Brabazon and O'Neill (2006).

In conclusion BIA's seem a great tool for modelling financial markets. First of all, this is because financial systems and prices defined in POE markets are affected by multiple unpredictable factors in a non-linear way [Brock *et al.* (1991)]; Second, BIA's are solutions to optimize problems that are common in finance (e.g. portfolio selection and management); and finally, BIA's can model a financial system as an ecosystem where different species are interacting, an idea I am going to develop in the next section. What is clear is Biological Inspired Algorithms remain an enormous field to explore, and will be potentially useful for financial analysis and forecast.

4.2. FINANCIAL ECOLOGY

We live in a world where traditional financial hypotheses are not fulfilled. Financial systems are a minus sum game where efficiency usually fails. Financial markets are complex ever changing systems ruled by irrational active and passive investors and arbitrageurs. Each investor (including arbitrageurs) holds a different portfolio according to his beliefs, necessities, fears and expectations. He tries to find satisfactory rather than optimal solutions and is affected by his behavioural biases [Simon (1955), Lo (2004)]. An Analytical Financial System as well as a hypothetical Global Financial System will never be completely information efficient (the best of cases would be adaptively efficient) and efficiency will evolve over time according to certain parameters that include actors and other economic and legal variables. Marginal efficiency and anomalies have allowed us to rethink the role of indexes in our world because passive investments are not always optimal, even if the market is efficient. The study and quantification of different index biases is essential to developing correct investment strategies, a point that I leave for future research. In spite of this, and until this detailed research is done, I present here an alternative vision of financial systems. Taking into account ideas and approaches from ecology, I develop a market structure along with its rules of evolution which has been inspired by other pioneering studies [Hirshleifer (1977), Farmer and Lo (1999), Farmer (2002), McGoun (1997), Frankfurter (2006) or Niederhoffer (1997)]. My approach is simple, incomplete and quantitative, but the ecological vision of financial systems is still in its infancy. My intention is not to define an extensive model but to give some ideas and reflections on the parallels between ecology and finance. I might summarize my aim by saying that the following paragraphs will raise more questions than answers.

4.2.1. FINANCIAL ORGANISM ECOLOGY IN A DYNAMIC GLOBAL FINANCIAL SYSTEM: DEFINITIONS

I see a *Global Financial System* as an ecosystem where competition determines evolution of the market[46.] As in Biology, a *financial ecosystem is a system formed by the interaction of a community of organisms with their environment*. It is necessary then to study in detail these two elements: environment and organisms.

Environment Analysis

The *Financial Environment* is a dynamic aggregation of all conditions and elements which make up the surroundings and influence the development and actions of individuals. I call these elements ecological factors. Ecological factors are elements related to POE markets, products, actors and techniques (parts of the Global Financial System defined in point 2) that act on organisms. It is possible to classify ecological factors as: i) External vs. Internal; ii) Abiotic[47] vs. Biotic[48]; iii) Limiting[49] vs. Non-limiting; and iv) Global vs. Regional. For further information see Figure 7.

46 General accepted biological definitions are used here. See Dajoz (2001) or Smith and Smith (2001).
47 Characteristics related to the environment.
48 Factors related to interactions among individuals.
49 Limiting factors limit growth, size or distribution of organisms by their presence or absence [Liebig (1840)].

Figure 7
Ecological Factors in a Global Financial System

POE markets	
Regulations, technologies, legal boundaries	External, Abiotic, Limiting, Global and Regional.
Products	
Availability, liquidity, legal conditions	External, Abiotic, Limiting, Global and Regional.
Actors	
Relationships	External, Biotic, Limiting, Global and Regional
Personal characteristics	Internal and Biotic
Techniques	
Uses, technology, models	External, Biotic, Limiting, Global and Regional

Source: Authors' own

Although ecological factors directly affect organisms living in the ecosystem, and also the environment itself, few studies analyse them in detail. Historically, some attempts have tried to cover ecological factors from the traditional financial point of view. There are studies relating liquidity or legal constraints to efficiency or technologies to trade, but perhaps with the exception of Behavioural Finance and its specialization in internal ecological factors, few studies have analysed ecological factors from an evolutionary or ecological point of view. An interesting field is developing for academics and researchers and in my opinion, mathematical models or empirical studies matching ecological factors with species living in a Global Financial System will be of considerable interest. How will trade be affected by regulations, changes in technologies, legal boundaries, liquidity or legal constraints to exchanges of financial products? How will efficiency be affected by the number of traders and species, their evolution as a group or their strategies? How will the probability of survival be determined by relationships between species, personal characteristics (including behavioural biases) or products used in each strategy?[50] These questions and others can be dealt with using a new ecological approach.

50 Survival analysis seems particularly useful for these studies. See Cox and Oakes (1984) for an introduction to Survival analysis.

Organism Analysis

Once ecological factors have been analysed, it is necessary to pay attention to populations living in the environment organized and controlled by ecological factors. Traditional and more modern approaches trying to classify organisms (basically investors) can be found in Figure 8.

Figure 8
Traditional and Modern Classifications of Investors

Classification 1	Institutional Vs Individual Investors
Classification 2	Domestic Vs Foreign Investors
Classification 3	Smart Money Vs Noise traders
Classification 4	Active Vs Passive Investors
Classification 5	Holders, Rebalancers, Valuators and Shifters
Classification 6	Producers Vs Speculators

Note: Classification 1 and 2 distinguish investors according to the information they have [asymmetrically informed traders, Kyle (1985)] and the restrictions they face (foreign investors have legal constraints in some markets for example). Classification 3 divides investors following Black (1986) between smart money (arbitrageurs) and noise traders that forecast with error prices and volatilities. Noise traders are not necessarily individuals, institutions can act as noise traders too [Hughen and McDonald (2005)]. Classification 4 divides investors according to their beliefs regarding efficiency. Classification 5 is presented in Leibowitz (2005). Holders follow a buy and hold strategy (they are passive investors or active investors with psychological biases that stop them changing positions). Rebalancers (usually institutions or funds) rebalance portfolios constantly to follow their portfolio allocation strategies. Valuators trade according to whether they believe something is cheap or expensive (active investors or arbitrageurs). Finally, Shifting occurs when a fundamental change in asset allocation is required usually due to monetary needs. All the groups presented can act as shifters at one moment or another. Classification 6 can be found in Zhang (1999). Producers participate in the market for their own needs rather than speculation (hedgers, tourists, financiers, among others) while Speculators are professional participants in the market (hedge funds, arbitrageurs, traders, etc.)

Source: Authors' own

I propose a different approach, providing a more extensive classification of all organisms living in a financial ecosystem.

Organisms of a population can be classified into groups in which species can be distinguished.[51] Species in financial markets are not only different because of their characteristics, but basically because of their behaviour regarding investments. What really helps us to distinguish between groups and species are activities regarding investment. Despite being a member of one species, an individual can act at a specific point of time as if it were another species, thus using an investment strategy that is typical of another species.

I start the organism analysis with a simple example. Imagine our Global Financial System is an Ecosystem located in a distant place. The ecosystem is quite distinct. It is placed at the end of a river which forms a little but growing delta. The river brings sediment and food for organisms living in the delta, but also determines delta's evolution, size and form. The river's flow, despite its fundamental importance for the configuration of the delta, is only one of the aspects of a more general phenomenon that defines the delta's form and physical evolution, that is, the climate. This includes the winds, temperatures, rains and the tides. The delta is also affected by another important element: organizers. Organizers live in the countryside, near the delta, taking care of it and defining its structure by making canals or moving earth, just as beavers do when they build their dams. Producers are other organisms to take into account. They use the delta for different activities creating thereby foods for other species. However, they do not live in the delta. There is a wide variety of Plants in the delta too. Some are few in number, concentrated in one particular area, while others are widespread, some even venturing into the countryside or establishing themselves on the banks of the river. All the plants feed from the sediment and minerals provided by the river and by the activities of the producers' and predators, and are affected by the climate. They provide refuge and a place where predators can stay and live during their feeding activities, thus they can be called providers too. In the the waters of the delta, defined by organizers, producers, providers and the climate, we can find fish, the main food source for predators. The population and behaviour of the fish depend directly on the organizers, producers, providers and the climate. Predators are the most extensive and interesting group inhabiting the ecosystem. It is a

51 That is, the taxonomic category composed of individuals possessing common characteristics and behaviour distinguishing them from other categories of individuals of the same taxonomic level. See Lo (2005b) for a financial approach to species classification.

general group including a great many of species that have one thing in common: they eat fish, although each species uses different techniques to hunt them.

In our example, the river, the flow, the sun and the winds, that is, the climate as a whole is the real economy, while the delta ecosystem is clearly the Global Financial System or, if we think geographically, the Global POE market. The Organizers are those actors or organisms living off the real economy or political institutions and defining and affecting the delta with their activities. They are the Government and Public Organizations creating laws to control and govern the Global Financial System, determining boundaries or limits and thus establishing the delta's existence and characteristics. Producers are organisms that use the Global Financial System for their own needs rather than speculation [Zhang (1999)]. In this group we can find financiers searching for funds, hedgers, tourists buying and selling currencies, etc. Producers, with their activities, are generating food for providers and predators. Fish in our analogy can be interpreted as profits, and their presence is directly affected by the real economy, that is, by organizers', producers' and providers' activities. Providers are another large group of organisms that feed from minerals (the trade and manipulation of financial products). Minerals depend directly on the climate (economic cycle) and other organisms' activities. Producers and also predators generate waste through their activities that can be interpreted as minerals. Among the providers we find organisms owning markets (companies that provide and control stock, futures or derivatives markets), brokers, dealers and market makers (providing access to the market or running it), and also information or software providers. In this group I finally include advisors, commissioners and other financial experts that sell their knowledge to other participants.[52] The producers' main objective is to stabilize the delta for the other organisms' activities, providing platforms for hunting or hiding (providing platforms to trade, access to the market, analytical tools for investment strategies, advice, etc.) meaning that providers is a highly appropriate name for them. Finally we have predators. Predators (investors) pursue fish (profits) and can be classified according to their hunting strategies (investment strategies). Traditionally two essential species have been defined:

52 In the first version of this paper I called them parasites. Because of the negative meaning of this word, I have not used it here, but there is a note about parasitism in the species relationship section.

passive and active investors. Passive investors are bears waiting for their prey to jump out of the water in a specific part of the delta because they believe the chance of finding a fish is the same wherever they wait (this is perhaps true for a perfect market but not always with real markets). Active investors think differently. They are convinced that fish are not equally distributed throughout the delta and try to find strategies to profit from that situation. Among the passive investors we find different strategies as total market exposure, index followers or buy and hold strategies. Active investors use infinite strategies too: arbitrage,[53] contrary, momentum and technical or fundamental strategies for example.

All actors in a Global Financial Market can be associated with one of these four global groups (producers, providers, organizers, predators). In more detail, if we separate the activities of each actor in a financial market, we can match each activity with one species.[54] A qualitative analysis of organisms (such as the one here) can be useful but models covering and explaining the dynamics and structure of this ecosystem are required. Models covering this approach are partial and incomplete. The learning process, competition, survival, mutation, the effect of ecological factors over species and other essential ideas of financial ecology have not yet been included in existing models.[55] I leave model development for further research, given that this is a huge field of investigation.

Species relationships

Ecologically it is possible to talk about three basic types of relationship among groups and species: *competence, predation and mutualism*.

Competition is a type of interaction where individuals of the same or different species look for the same scarce resource or commodity. In our opinion

53 I consider them active investors because real arbitrage assumes risk and does not use the market portfolio strategy.
54 Sophisticated financial intermediaries such as banks or insurance companies have different strategies in their normal activities (advice, access to markets, information sellers, own investment, etc.) and each activity can be associated to a group.
55 Figlewski (1978, 1982), Luo (1995, 1998, 2001, 2003), Niederhoffer (1997) or Patel *et al.* (1991) can be cited as pioneering studies.

competition is the most important relationship in a financial ecosystem. Organizers compete for power in political spheres, providers compete among themselves for trade, predators compete for benefits and some kind of competition can also be established among producers. Competitive relationships open new possibilities for parallelisms. Can intraspecies and interspecies ecological models help us to analyse a financial ecosystem? Can territorial or temporal competition be defined in similar terms to those used in ecology?

Predation means one species feeding on another species. It should not be a common interaction in a financial ecosystem but it exists.[56] Parasitism can be analysed as slight predation and it is more common than predation. For example, when providers do not provide real value they act as parasites feeding from predators causing them illnesses (reduction in profits) or death (elimination from the market). Predation, parasitism and Epidemiology (the study of illnesses) must be used to analyse inefficiencies, understood as symptoms of market power, abuses, illegal or barely legal activities, and as indicators of quality in the performance of actors in the ecosystem, both in developed and emerging markets.

Finally, *mutualism* means two species obtaining profits from *mutual* interaction.

On the empirical hand, it is fundamental to understand an ecosystem, but is more useful to modelize it. Classical mathematical models developed by Lotka and Volterra, Nicholshon and Bailey or Hassell and Varley may be extraordinarily useful for these purposes [see Dajoz (2001) for more information].

4.2.2. FINANCIAL ORGANISM ECOLOGY IN A DYNAMIC GLOBAL FINANCIAL SYSTEM: STATUS AND EVOLUTION

The state of the financial ecosystem evolves over time and is the result of the interaction between the climate (as an abiotic factor) and organisms (as biotic

56 Predation includes activities as the elimination of naïve investors, abuses, or insider practices.

factors). The rules that define evolution and survival in the ecosystem are summarized below. The ideas presented here complete and extend the Adaptive Market Hypothesis presented in Lo (2004, 2005).

1) Individuals act in their own self-interest to ensure their own survival.

2) Individuals make mistakes and do not act always rationally.

3) Individuals learn, adapt and change their feeding activities using trial and error. They select strategies and develop heuristics according to experience, positive and negative reinforcement.

4) Individuals do not have the same propensity to learn, adapt and change their feeding activities. They are affected by psychological biases[57] and other ecological factors that define their propensity to change and adapt.

5) Individuals do not hold the same portfolio because they use different investment strategies.

6) Investment strategies fit differently to the market, so they bring different probabilities of survival. Investment strategy fitting is not related necessarily to rationality.

7) The survival probability associated with each investment strategy is not constant over time. It evolves according to the ecosystem's evolution.

57 Trivers (1985, 1991), Waldman (1994), Hirshleifer (1999) or Daniel and Titman (1999).

8) Organisms establish relationships among themselves, and these relationships are not stable over time.

9) Ecological factors define natural selection and the survival of organisms.

10) Organism interaction drives adaptation, innovation and mutation. I see adaptation as changes in investment strategies, and innovation and mutation as the creation of new investment strategies.

11) Surviving individuals transmit, directly or indirectly, their investment strategies to offspring (reproduction). Offspring can be analysed as new investors entering the Global Financial Market over time.

12) Non surviving individuals disappear from the market.

13) The state of the ecosystem determines future evolution.

Some interesting conclusions can be drawn analysing these ideas regarding financial ecosystem evolution: i) Behavioural Finance becomes a key issue in financial analysis because individuals commit errors, are affected by behavioural biases or learn and adapt their investment strategies; ii) Investment strategies provide different survival probabilities. This should be analysed carefully from an empirical point of view, but it seems clear from a conceptual point of view. Moreover, profitability is cyclical in each investment strategy (and in each financial product) according to market evolution; iii) Ecological factors (both abiotic and biotic) define natural selection and the evolution of groups and species in our financial ecosystem. The number of individuals (inside each group or species) changes over time according to the characteristics of the ecosystem. Changes or rebalances in some ecological factors affect population presence or absence; iv) Natural selection and organism interaction define evolution and marginal efficiency in the market; v) The equity risk premium is not constant, it varies with market participants' characteristics (demographics or behavioural biases) and ecosystem status; vi) Individual risk preferences are not stable over time; vii) Active strategies add value if they provide a better risk-return relation or outperform the average return of the market and this is only possible by adapting strategies in the face of changes in the financial ecosystem.

5. CONCLUSIONS

The market, in contrary to what is defended in traditional finance, is not a zero-sum game. There are commissions, fees, advisors or brokers that live off the market and off investors. EMH is not fulfilled in our actual world. Financial markets are very different from markets in the 1970s. The world has changed, and what seemed plausible thirty years ago is difficult to imagine now. The evolution of technologies, direct market access through internet by individual traders, and the increase in money invested directly through these channels may have changed the market and its efficiency. But market efficiency is also affected by investor rationality, correlations in irrational trades and limitations in arbitrage, elements that are seen nowadays from a different perspective. It has been shown empirically that anomalies and other inefficiencies have been growing too. The CAPM has played an important role in finance, but tells us more when assumptions are not fulfilled. The market portfolio is not the proper mix among risky securities for everybody when assumptions are not sustainable in the real world, contrary to what is defended by traditional finance. The conclusion then is clear: a passive strategy is not optimal and is not located in the efficient frontier, and index followers could be supporting a suboptimal strategy if they do not pay attention to index biases. Market Indexes present biases depending on how they are constructed, selected, and replicated (e.g. sample and construction biases, tracking error). In addition, other biases appear when the market is really inefficient (e.g. efficiency and active biases). Taking into account these biases is vitally important, and makes us rethink the meaning of outperformance or underperformance of the market.

After analysing in detail the circumstances which finance is facing, I have presented Ecology in Financial Markets as a promising discipline that can bring new ideas and approaches to financial analysis. Two lines of research within this discipline were looked at: Biologically Inspired Algorithms (BIA), (providing tools as Neural Networks, Artificial Immune Systems, Evolutionary Computation, and Social Systems) and Financial Ecology, which studies a financial system as an ecosystem where different organisms live and compete for survival. A qualitative analysis of a financial system has been provided in our article using this approach. I divide a financial ecosystem in four groups of organisms (Producers, providers, organizers and predators) and provide different rules that define evolution and survival in the ecosystem.

6. FUTURE LINES OF RESEARCH

A deep analysis of indexes is required in order to study and quantify index biases. A good knowledge of these biases is essential to define investment strategies and, according to the results, market indexes should be revised. Secondly, the ecological approach seems a promising discipline to be exploited. More detailed analysis of financial systems as ecosystems and mathematical models applying these ideas should be encouraged. How predation, competition and mutualism relationships affect the financial system’s evolution must be analysed and mathematical models from ecology should be used in this objective. I will take steps in this direction with my future work.

I would like to thank Dr. Maxim Borrell, Dr. Salvador Torra and Professor Sebastian Cano for their comments/opinions regarding this paper.

7. APPENDIX

Figure 4a
Empirical Anomalies

Return predictability through autocorrelation	Lo and MacKinlay (1988)
Volatility studies	Shiller (1989)
Forecasting power of financial variables	Solnik (1993)
Periodical anomalies and firm effects	Fama (1991)
Profitable strategies (momentum, technical, etc.)	Jegadeesh and Titman (2001)
Event studies	Jegadeesh and Titman (1993)
Insider trading studies	Seyhun (1998)

Figure 4b
Random Walk Methodologies

First tests	Cowles and Jones (1937)
Augmented Dickey-Fuller test (ADF)	Dickey and Fuller (1979)
GPH fractional integration test	Geweke and Porter-Hudak (1983)
Single Variance Ratio test (LOMAC)	Lo and MacKinlay (1988)
Thin trading methodologies	Miller *et al.* (1994)
Rank and Signs Variance Ratio test (RS)	Wright (2000)
Non-linearity Analysis	Brock *et al.* (1991),

Figure 4c
The Random Walk Analysis in emerging markets.

Africa	Dickinson and Muragu (1994), Olowe (1999)
Turkey	El-Erian and Kumar (1995), Antoniou and Ergul (1997), Darrat and Benkato (2003), Buguk and Bronser (2003)
Gulf Markets	Abraham *et al.* (2002), Butler and Malaikah (1992)
Latin America	Urrutia (1995), Ojah and Karemera (1999), Grieb and Reyes (1999)
Asia	Alam *et al.* (1999), Darrat and Zhong (2000), Lima and Tabak (2004)

Note: only some authors have been selected as a reference for each anomaly.

8. REFERENCES

Abraham, A.; Seyyed, F.; Alsakran, S. (2002): "Testing the Random Walk Behavior and Efficiency of the Gulf Stock Markets".*The Financial Review*, v.37, pp.469-480.

Adler, M.; Dumas, B. (1983): "International Portfolio Choice and Corporation Finance: A Synthesis". *Journal of Finance*, v.38, pp.925-984.

Alam, M.; Hasan, T.; Kadapakkam, P. (1999): "An Application of Variance-Ratio Test of Five Asian Stock Markets". *Review of Pacific Basin Financial Markets and Policies*, v.2, pp.301-315.

Alexander, S. S. (1961): "Price Movements in Speculative Markets: trends or random walks". *Industrial Management Review*, n.2, pp.7-26.

Alexander, S. S. (1964): "Price Movements in Speculative Markets: trends or random walks, n.2", in P. Cootner, ed., *The Random Character of Stock Market Prices*. Cambridge, MA: MIT Press.

Andersen T.G.; Bollerslev, T.; Diebold, F.; Wu, J. (2003): "Realized Beta: Persistence and Predictability". *PIER Working Paper Archive.*

Antoniou, A.; Ergul, N. (1997): "Market Efficiency, Thin trading and Non-linear Behaviour: Evidence from and Emerging Market". *European Financial Management*, v.3, pp.175-190.

Arshanapalli, B.; Switzer, L. N.; Hung, L.T.S. (2004): "Active versus passive strategies for EAFE and the S&P500". *Journal of Portfolio Management*, v.30, iss.4, pp.51-60.

Asquith, P. (1983): "Merger bids, uncertainty and stock holder returns”. *Journal of Financial Economics*, v.11, pp.51-83.

Bachelier, L. (1900/reprint 1964): "Theory of speculation ". Cootner Ed, *Random Character of Stock Market Prices*, MIT Press, Cambridge.

Baestanes, D. E.; Van Den Bergh, W. M.; Wood, D. (1994): *Neural Network Solutions for Trading in Financial Markets*. London: Pitman Publishing.

Balvers, R.; Wu, Y.; Gilliland, E. (2000): "Mean reversion across national stock markets and parametric contrarian investment strategies". *Journal of Finance*, v.55, pp.745-772.

Balyers, R.; Huang, D. (2004): "Money and the (C) CAPM: Theory and Evaluation". *Department of Economics, West Virginia University, Working Papers: 04-10.*

Barber, B.; Odean, T. (2000): "Trading is hazardous to your wealth: the common stock investment performance of individual investors". *Journal of Finance*, v.55, pp.773-806.

Barkow, J.; Cosmides, L.; Tooby, J. (1992): *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*.. Oxford: Oxford University.

Barrett, L.; Dunbar, R.; Lycett, J.(2002): *Human Evolutionary Psychology*. Princeton, N. J.; Princeton University Press.

Bartholdy, J.; Peare, P. (2005): "Estimation of Expected Return: CAPM vs. Fama and French". *International Review of Financial Analysis*, v.14, iss. 4, pp. 407-427.

Bauer, R. (1994): *Genetic algorithms and Investment Strategies*. New York: Wiley.

Beckaert, G. (1996): "The time variation of risk and return in foreign exchange markets: a general equilibrium perspective". *The Review of Financial Studies*, v.9, pp.427-470.

Becker, K.G.; Finnerty, J. E.; Kopecky, K.J. (1996): "Macroeconomic News and the Efficiency of International Bond Futures Markets". *Journal of Futures Markets*, v.16, iss.2, pp.131-145.

Beebower, G. L.; Bergstrom, G. L. (1977): "A performance analysis of pension and profit-sharing portfolios: 1966-75". *Financial Analysts Journal*, v.33, pp.31-42.

Bell, A.R.; Brookds, C.; Dryburgh, P. (2007): "Interest rates and efficiency in medieval wool forward contracts". *Journal of Banking and Finance*, v.31, iss.2, pp.361-380.

Binmore, K. (2008): *Game Theory: a Very Short Introduction.* Oxford University Press.

Black, F. (1972): "Capital Market Equilibrium with Restricted Borrowing". *Journal of Business*, v.45, iss.3, pp.444-455.

Black, F. (1986): "Noise". *Journal of Finance*, v.41, pp.529-543.

Black, F.; Jensen, M.C.; Scholes, M. (1972): "The Capital Asset Pricing Model: Some Empirical Tests", in Jensen, M.C. ed. *Studies in the Theory of Capital Markets*, New York: Praeger.

Blume, M. E. (1975): "Betas and Their Regression Tendencies". *The Journal of Finance*, v.30, iss.3, pp. 785-795.

Blume, M.E. (1979): "Betas and Their Regression Tendencies: Some Further Evidence". *The Journal of Finance*, v.34, iss.1, pp. 265-267.

Blume,M.; Easley, D.; O'Hara, M. (1994): "Market Statistics and Technical Analysis: The Role of Volume", *Journal of Finance*, v.49,iss.1, pp.153-181.

Blume, M.; Friend, I. (1973): "A New Look at The Capital Asset Pricing Model". *Journal of Finance*, v.28, iss.1, pp.19-33.

Bonabeau, E.; Dorigo, M.; Theraulaz, G. (1999): *Swarm Intelligence: From Natural to Artificial Systems.* Oxford: Oxford University Press.

Brabazon, A.; O'Neill, M. (2006): *Biologically Inspired Algorithms for Financial Modelling*. New York: Springer.

Breeden, D. (1979): "An intertemporal Asset Pricing Model With stochastic Consumption and Investment Opportunities". *Journal of Financial Economics*, v.7, pp.265-296.

Brock, W.; Hsieh, D.; LeBaron, B. (1991): *Nonlinear Dynamics, Chaos and Instability: Statistical Theory and Economic Evidence*. Cambridge, MA: MIT press.

Brock, W.; Lakanishok, J.; Lebaron, B. (1992): "Simple Technical Trading Rules and the Stochastic Properties of Stock Returns", *The Journal of Finance*, v.48, n.5, pp.1731-1764.

Buguk, C.; Brorsen, B.W. (2003): "Testing weak-form market efficiency: Evidence from the Istanbul Stock Exchange". *International Review of Financial Analysis*, v.12, pp.579-590.

Buss, D. (1999): *Evolutionary Psychology: The New Science of the Mind*. Boston: Allyn and Bacon.

Butler, K.; Malaikah, S. (1992): "Efficiency and Inefficiency in Thinly Traded Stock Markets: Kuwait and Saudi Arabia". *Journal of Banking and Finance*, v.16, pp.197-210.

Campbell, J.Y.; Vuolteenaho, T. (2004): "Bad Beta, Good Beta". *American Economic Review*, v.94, iss.5, pp.1249-1275.

Cassese, G.; Guidolin, M. (2004): "Pricing and Informational Efficiency of he MIB30 Index Options Market: an analysis with High-Frequency Data". *Economic Notes*, v.33, iss.2, pp.275-321.

Chamberlain, G.; Rothschild, M. (1983): "Factor structure, and mean-variance analysis on large asse markets". *Econometrica*, sept.

Chan, L.; Jegadeesh, N.; Lakonishok, J. (1996): "Momentum Strategies". *Journal of Finance*, v.51, pp.1681-1713.

Chang, E. C.; Lewellen, W. G. (1984): "Market timing and mutual fund investment performance". *Journal of Business*, v.57, pp.57-72.

Charest, G. (1978): "Dividend information, stock returns, and market efficiency-II". *Journal of Financial Economics*, v.6, pp. 297-330.

Chen, M.-H. (2003): "Risk and Return: CAPM and CCAPM". *Quarterly Review of Economics and Finance*, v.43, iss.2, pp. 369-393.

Colin, A. (1994): "Genetic algorithms for financial modelling" in Deboeck, G. (Ed), *Trading on the Edge: Neural, Genetic and Fuzzy systems for Chaotic and Financial Markets*, New York: Wiley.

Conrad, J.; Kaul, G. (1988): "Time-variation in expected returns". *Journal of Business*, v.61, pp.409-425.

Constantinides, G.M. (1983): "Capital Market Equilibrium with Personal Tax". *Econometrica*, v.51, iss.3, pp.611-636.

Cowles, A. (1960): "A Revision of Previous Conclusions Regarding Stock Price Behavior". *Econometrica*, v.28, pp.909-915.

Cowles, A.; Jones, H. (1937): "Some a Posteriori Probabilities in Stock Market Action". *Econometrica*, v.5, pp.280-294.

Cox, D.R.; Oakes. D. (1984): *Analysis of Survival Data*, Chapman & Hall/CRC.

Crawford, C.; Krebs, D., eds. (1998): *Handbook of Evolutionary Psychology: Ideas, Issues, and Applications*. Mahwah, N. J.; Lawrence Erlbaum.

Cuerdo Mir, M.; Ramos Gorostiza, J.L. (2000): *Economía y naturaleza. Una historia de las ideas*. Editorial Sintesis, Madrid 2000.

Dajoz, R. (2001): *Tratado de Ecología*. Madrid: Ediciones Mundi-Prensa.

Daniel, K.; Titman, S. (1999): "Market efficiency in an Irrational World". *Financial Analysts Journal*, nov/dec 1999, v.55, iss.6, pp.28-40.

Darrat, A. F.; Benkato, O. M. (2003): "Interdependence and Volatility Spillovers under Market Liberalization: the case of Istanbul Stock Exchange". *Journal of Business Finance and Accounting*, v.30, pp.1089-1114.

Darrat, A.; Zhong, M. (2000): "On Testing the Random-Walk Hypothesis: A Model-Comparison Approach". *The Financial Review*, v.35, iss.3, pp.105-124.

De Long, J. B.; Shleifer, A.; Summers, L.H.; Waldmann, R.J. (1990): "Noise trader risk in financial markets". *Journal of Political Economy*, v.98, pp.703-738.

Deboeck,G. (1994): *Trading on the Edge: Neural, Genetic and Fuzzy Systems for Chaotic Financial Markets*. New York: John Wiley and Sons.

Debondt, W.F.M.; Thaler, R.H. (1985): "Does the stock market overreact”. *Journal of Finance*, v.40, pp.793-805.

Debondt, W.F.M.; Thaler, R.H. (1987): "Further evidence on investor overreaction and stock market seasonality". *Journal of Finance*, v.42, pp.557-581.

Dickey, D.; Fuller, W. (1979): "Distributions of the estimators for autoregressive time series with a unit root". *Journal of American Statistical Association*, v.75, pp.427-431.

Dickinson, J.P.; Muragu, K. (1994): "Market Efficiency in Developing Countries: a Case Study of the Nairobi Stoch Exchange". *Journal of Business Finance and Accounting*, v.21, pp.133-150.

Dimson, E.; Marsh, P. (1999): "Murphy's law and market anomalies". *Journal of Portfolio Management*, v.25. iss.2, pp.53-69.

Douglas, G.W. (1968): *Risk in the Equity Markets: An Empirical Appraisal of Market Efficiency*. Ann Arbor, Michigan: University Miccrofilms, Inc.

Edwards, W. (1968): "Conservatism in human information processing", in Kleinmutz, B. (ed): *Formal Representation of Human Judgment*. New York: John Wiley and Sons.

Elder, A. (1993): *Trading for a Living*, Wiley.

El-Erian, M.; Kumar, M. (1995): "Emerging Equity Markets in the Middle Eastern Countries". *IMF Staff Papers*, v.42, 313-343.

Elton, E.J.; Gruber, M.J. (1975): *Finance as a Dynamic Process*. NJ: Prentice Hall.

Elton, E. J.; Gruber, M.J. (1991): *Modern Portfolio Theory and investment analysis*. New York: John Wiley and Sons.

Estrada, J. (2002): "Systematic Risk in Emerging Markets: The D-CAPM". *Emerging Markets Review*, v.3, iss.4, pp.365-379.

Fama, E. (1965): "Random Walks in Stocks Markets". *Financial Analyst Journal*, sept, pp.55-59.

Fama, E. (1965b): "The behavior of stock market prices". *Journal of Business*, v.38, pp.34-105.

Fama, E. (1970): "Efficient capital markets: A review of theory and empirical work". *Journal of finance*, v.25, pp.383-417.

Fama, E. (1970b): "Multi Period Consumption Investment Decision". *American Economic Review*, v.60, pp.163-174.

Fama, E. (1991): "Efficient Capital Markets. II". *The Journal of Finance*, v.46, n.5, pp.1575-1617.

Fama, E.; French, K. R. (1988): "Permanent and temporary components of stock prices". *Journal of Political Economy*, v.96, pp.246-273.

Fama E.; French, K.R. (1992): "The Cross-Section of Expected Stock Returns". *Journal of Finance*, v.47, iss.2, pp.427-465.

Fama, E.F.; French, K.R. (2004): "The Capital Asset Pricing Model: Theory and Evidence". *Journal of Economic Perspectives*, v.18, iss.3, pp.25-46.

Fama, E.F.; French, K.R. (2006): "The Value Premium and the CAPM". *Journal of Finance*, v.61, iss.5, pp.2163-2185.

Fama, E.F.; MacBeth, J.D. (1973): "Risk, Return and Equilibrium: Empirical Tests". *Journal of Political Economy*, v.81, iss.3, pp.607-636.

Farmer, D. (2002): "Market Force, Ecology and Evolution". *Industrial and Corporate Change*, v.11, iss.5, pp. 895-953.

Farmer, D.; Lo, A. (1999): "Frontiers of Finance: Evolution and Efficient Markets". *Proceedings of the National Academy of Sciences*, v.96, pp.9991-9992.

Fernandez, V. (2006): "The CAPM and Value at Risk at Different Time-Scales". *International Review of Financial Analysis*, v.15, iss.3, pp. 203-219.

Fernandez-Rodriguez, F.; Gonzalez-Martel, C.; Sosvilla-Rivero, S. (2000): "On the profitability of technical trading rules based on artificial neural networks: Evidence from the Madrid stock market". *Economic Letters*, n.69, pp.89-94.

Figlewski, S. (1978): "Market 'Efficiency' in a Market with Heterogeneous Information". *Journal of Political Economy*, v.85, pp.681-597.

Figlewski, S. (1982): "Informational Diversity and Market Behavior". *Journal of Finance*, v.37, pp.87-102.

Fisher, L. (1966): "Some new stock-market indexes". *Journal of Business*, v.39, pp.191-225.

Fletcher, J.; Kihanda, J. (2005): "An examination of Alternative CAPM-Based Models in UK Stock Returns". *Journal of Banking and Finance*, v.29, iss.12, pp. 2995-3014.

Frank, M.; Gencay, R.; Stengos, T. (1988): "International Chaos". *European Economic Review*, v.32, iss.8, pp.1569-1584.

Frankfurter, G. M. (2006): "The Theory of Fair Markets (TFM) toward a New Finance Paradigm". *International Review of Financial Analysis*, v.15, iss.2, pp.130-144.

Franses, P.; Van Homelen, P. (1998): "On forecasting exchange rates using neural networks". *Applied Financial Economics*, n.8, pp.589-596.

Fraser, P.; Hamelink, F.; Hoesli, M.; Macgregor, B. (2004): "Time-varying betas and the cross-sectional return-risk relation: evidence from the UK". *The European Journal of Finance*, v.10, iss.4, pp. 255-276.

Friend, I.; Landskroner, Y.; Losq, E. (1976): "The Demand for Risky Assets and Uncertain Inflation". *Journal of Finance*, v.31, n.5, pp.1287-1297.

Gately, E. (1996): *Neural Networks for Financial Forecasting*. New York: Wiley.

Geweke, J.; Porter-Hudak, S. (1983): "The estimation and aplication of long memory time series models". *Journal of Time Series Analysis*, v.4, pp.221-238.

Gigerenzer, G. (2000): *Adaptative Thinking: Rationality in the Real World*. Oxford: Oxford University Press.

Goldberg, D. E. (1989): *Genetic algorithms in Search, Optimization and Machine Learning*. Boston: Addison-Wesley.

Gomez, J-P.; Zapatero, F. (2003): "Asset Pricing Implications of Benchmarking: A Two-Factor CAPM". *European Journal of Finance*, v.9, iss.4, pp.343-357.

Granger, C.; Morgenstern, O. (1963): "Spectral Analysis of New York Stock Market Prices". *Kyklos*, v.16, pp.1-27.

Grieg, T.; Reyes, M.G. (1999): "Random walk tests for Latin American equity indices and individual firms". *Journal of Financial Research*, v.4, pp.371-383.

Grossman, S.; Stiglitz, J. (1980): "On the impossibility of Informationally Efficient markets". *American Economic Review*, v.70, pp.393-408.

Gulko, L. (2004): "Efficient Irrational Markets". *Journal of Portfolio Management*, Winter 2004, v.31, iss.2, p.64-72.

Haim, L.; Moshe, L.; Golan, B. (2006): "Capital Asset Prices with Heterogeneous Beliefs". *Journal of Business*, v.79, iss.3, pp. 1317-1353.

Hall, S.G.; Miles, D.K. (1992): "Measuring Efficiency and Risk in the Major Bond Markets". *Oxford Economic Papers*, v.44, iss.4, pp. 599-625.

Halpern, P.; Rumsey, J. (1997): "A Note on Market Efficiency, Institutional Practice, and Economic Constraints in the Experience of the Canadian Bond Market". *Journal of Banking and Finance*, v.21, iss.1, pp.113-123.

Haugen, R. (2001): *Modern Investment Theory*. New Jersey: Prentice Hall.

Haugen, R. A.; Baker, H.L. (1991): "The Efficient market inefficiency of Capitalization-Weighted Stock Portfolios". *Journal of Portfolio Management*, Spring 1991, v.17, iss.3, pp.35-40.

Henriksson, R. T. (1984): "Market timing and mutual fund performance: An empirical investigation". *Journal of Business*, v.57, pp.73-96.

Hinick, M.; Patterson, D. M. (1985): "Evidence on nonlinearity in daily stock returns". *Journal of Business and Economic Statistics*, v.3, pp.69-77.

Hirshleifer, D. (1999): "An adaptative Theory of Self-Deception". *Working paper*. Ohio State University.

Hirshleifer, J. (1977): "Economics from a Biological Viewpoint". *Journal of Law and Economics*, v.20, pp.1-52.

Hirshleifer, D.; Luo, G. (2001): "On the Survival of Overconfident Traders in a Competitive Securities Market". *Journal of Financial Markets*, v.4, pp.73-84.

Holland, J.H. (1975): *Adaptation in Natural and Artificial Systems*. Michigan: University of Michigan Press.

Hotchkiss, E.S.; Ronen, T. (2002): "The Information efficiency of the Corporate Bond market: an intraday Analysis". *Review of Financial Studies*, v.15, iss.5, pp. 1325-1354.

Hsieh, D. A. (1991): "Chaos and nonlinear dynamics: application to financial markets". *Journal of Finance*, v.46, pp.1839-1877.

Hu, M.; Shi, Y.; Eberhart, R. (2004): *Recent advances in particle swarm*, Portland, Oregon, 19-23 June, 2004, IEEE Press: New Jersey.

Hu, M.; Zhang, G.; Jiang, C.; Patuwo, E. (1999): "A cross-validation analysis of neural network out-of-sample performance in exchange rate forecasting". *Decision Sciences*, v.30, iss.1, pp. 197-216.

Hughen, J. C.; McDonald, C. G. (2005): "Who are noise traders?". *The Journal of Financial Research*, v. 28, n.2, pp.281-298.

Hutchinson, H. (1999): *Portfolio indexing. Theory and Practice*. New York: Wiley.

Hutchinson, J.; Lo, A.; Poggio, T. (1994): "A non parametric approach to the pricing and hedging of derivative securities via learning networks". *Journal of Finance*, v.49, pp.851-889.

Hyde, S.; Cuthbertson, K.; Nitzsche, D. (2005): "Resuscitating the C-CAPM: Empirical Evidence from France and Germany". *International Journal of Finance and Economics*, v.10, iss.4, pp. 337-357.

Ippolito, R. A.; Turner, J. A. (1987): "Turnover, fees and pension plan performance". *Financial Analysts Journal*, v.43, pp.16-36.

Jagannathan, R.; Wang, Z. (1996): "The conditional CAPM and the cross-section of expected returns". *Journal of Finance*, v.51, pp.3-53.

Jasic, T.; Wood, D. (2006): "Testing for Efficiency and Non-linearity in Market and Natural Time Series*". Journal of Applied Statistics*, v.33, iss.2, pp.113-138.

Jegadeesh, N.; Titman, S. (1993): "Returns to buying winners and selling losers: Implications for stock market efficiency". *Journal of Finance*, v.48, iss.2, pp.65-91.

Jegadeesh, N.; Titman, S. (2001): "Profitability of Momentum Strategies: An evaluation of Alternative Explanations". *Journal of Finance*, v.56, iss.2, pp.699-720.

Jeng, L.A.; Metrick, A.; Zeckhauser, R. (1999): "The profits to insider trading: A performance-evaluation perspective". *NBER Working paper #6913*. Cambridge, MA. NBER.

Jensen, M.C. (1978): "Some anomalous evidence regarding market efficiency". *Journal of Financial Economics*, v.6, pp.95-101.

Joslova, G.; Philipov, A. (2005): "Bayesian Analysis of Stochastic Betas". *Journal of Financial and Quantitative Analysis*, v.40, iss.4, pp.747-778.

Kahneman, D.; Slovic, P.; Tversky, A. (1982): *Judgment under Uncertainty: Heuristics and Biases.* New York:Cambridge University Press.

Kahneman, D.; Tversky, A. (1973): "On the Psychology of Prediction". *Psychological Review*, v.80, pp.237-251.

Keane, S. M. (1986): "The efficient market hypothesis on trial". *Financial Analysts Journal*, v.42, iss2. pp.58-63.

Kendall, M. G. (1953): "The analysis of Economic Time-Series, Part-I Prices". *The Journal of the Royal Statistical Society*, v.96, pp.11-25.

Kennedy, J.; Eberhart, R.; Shi, Y. (2001): *Swarm Intelligence*. San Mateo, California: Morgan Kaufman.

Keown, A.; Pinkerton, J. (1981): "Merger announcements and insider trading activity: An empirical investigation". *Journal of Finance*, v.36, pp.855-869.

Kim, M. J.; Nelson, C.R.; Startz, R. (1991): "Mean reversion of stock prices: A reappraisal of the empirical evidence". *Review of Economic Studies*, v.58, pp.5151-5280.

King, B. (1966): "Market and Industry Factors in Stock Price Behavior". *Journal of Business*, v.39, pp-139-140.

Kleiman, R. T.; Payne, J. E.; Sahu, A. P. (2002): "Random Walks and Market Efficiency: Evidence from International Real Estate Markets". *Journal of Real Estate Research*, v.24, iss.3, pp.279-297.

Koza, J. (1992): *Genetic Programming*. Massachusetts: MIT Press.

Koza, J. (1994): *Genetic Programming II*. Massachusetts: MIT Press.

Kyle, A. S. (1985): "Continuous Auctions and Insider trading". *Econometrica*, v.53, pp.1315-1335.

Larsen, C.; Lam, L. (1992): "Chaos and the foreign exchange market". In Lam, L.; Naroditsky, V. (eds): *Modelling Complex Phenomena*. New York:Springer.

Larson, A. (1960): "Measurement of Random Process in Futures Prices". *Food Research Institute*, v.1, pp.313-324.

Leibowitz, M. L. (2005): "Alpha Hunters and Beta Grazers". *Financial Analysts Journal*, v.61, iss.5, pp.32-39.

Leroy, S. (1973): "Risk Aversion and the Martingale Property of Stock Returns". *International Economic Review*, v.14, pp.436-446.

Leung, M.T.; Daouk, H.; Chen, A. (2000): " Forecasting stock indices: a comparison of classification and level estimation models". *International Journal of Forecasting*, v.16, pp.173-190.

Liebig, J. (1840): *Traité de chimie organique*. Paris: Baillère.

Liew, J. (2003): "Hedge Fund Index Investing Examined". Journal *of Portfolio Management*, Winter 2003, v.29, iss.2, pp.113-123.

Lima, E. J. A.; Tabak, B. M. (2004): "Tests of random walk hypothesis for Equity Markets: Evidence from China, Hong Kong and Singapore". *Applied Economics Letters*, v.11, iss.4, pp.255-258.

Lintner, J. (1965): "Security Prices, Risk, and Maximal Gains from Diversification". *Journal of Finance*, v.20, iss.4, pp. 587-615.

Lintner, J. (1969): "The aggregation of Investors Diverse Judgements and Preferences in Purely Competitive Security Markets". *Journal of Financial and Quantitative Analysis*, v.4, iss.4, pp.347-400.

Liu, C.Y.; He, J. (1991): "A variance-ratio test of random walks in foreign exchange rates". *Journal of Finance*, v.46, pp.773-785.

Lo, A. W. (2004): "The Adaptative Markets Hypothesis: market efficiency from an Evolutionary Perspective". *Journal of Portfolio Management*, v.30, pp.15-29.

Lo, A.W. (2005): "Reconciling Efficient Markets with behavioural finance: the adaptive markets hypothesis". *Journal of Investment Consulting*, v.7, pp.21-44.

Lo, A.W.; MacKinlay A. C. (1988): "Stock market prices do not follow random walks: Evidence from a simple specification test*". Review of Financial Studies*, v.1, pp.41-66.

Lo, A.W.; Repin, D.V.; Steenbarger, B. N. (2005): "Fear and Greed in Financial Markets: A Clinical Study of Day Traders". *American Economic Review*, v.95, pp.352-359.

Lucas, R. (1978): "Asset Prices in an Exchange Economy". *Econometrica*, v.46, pp.1429-1446.

Lundgren, T. (2005): "Assessing the Investment performance of Swedish Timberland: A Capital Asset Pricing Model Approach". *Land Economics*, v.81, iss.3, pp.353-362.

Luo, G. (1995): "Evolution and Market Competition". *Journal of Economic Theory*, v.67, pp.223-250.

Luo, G. (1998): "Market Efficiency and Natural Selection in a Commodity Futures Market". *Review of Financial Studies*, v.11, pp.647-674.

Luo, G. (2001): "Natural Selection and Market Efficiency in a Futures Market with Random Shocks". *Journal of Futures Markets*, v.21, pp.489-516.

Luo, G. (2003): "Evolution, Efficiency and Nose Traders in a One-Sided Auction Market". *Journal of Financial Markets*, v.6, pp.163-167.

Mandelbrot, B. (1963): "The Variation of Certain Speculative Prices". *Journal of Business*, v.36, pp. 394-419.

Markowitz, H.M. (2005): "Market efficiency: a theoretical distinction and so what?". *Financial Analysts Journal*, vol61. iss.5, pp.17-30.

Mayers, D. (1973): "Nonmarkable Assets and the Determination of Capital Asset Prices in the Absence of a Riskless Asset". *Journal of Business*, v.46, pp.258-267.

McGoun, E. G. (1997): "Hyperreal finance". *Critical Perspectives on Accounting*, v.8, pp.97-122.

Merton, R. (1973): "An Intemporal Capital Asset Pricing Model". *Econometrica*, 41, p. 867-887.

Merton, R. (1987): "A simple Model of Capital Market Equilibrium with Incomplete Information ". *Journal of Finance*, v.42, pp.483-510.

Miller, M.; Muthuswamy, J.; Whaley, R. (1994): "Mean reversion of standard and Poor's 500 Index Basis Changes: Arbitrage-Induced or Statistical Illusion". *The Journal of Finance*, v.49, pp.479-513.

Miller, M.; Scholes, M. (1972): "Rates of Return in Relation to Risk: A Reexamination of Some Recent Findings", in Jensen, M. C. (ed) *Studies in the Theory of Capital Markets*. New York: Praeger.

Mitchell, M. (1996): *An introduction to Genetic Algorithms*. Massachusetts: MIT Press.

Mossin, J. (1966): "Equilibrium in a Capital Asset Market". *Econometrica*, October, n.35, pp.768-783.

Neely, C.; Weller, P., Dittmar, R. (1997): "Is technical analysis in the foreign exchange market profitable? A genetic programming approach". *Journal of Financial and Quantitative analysis*, v.32, pp.405-426.

Nicholson, N.; Fenton-O'Creevy, M.; Soane, E.; Willman, P. (2002): "Risk Propensy and personality". *COR working paper.*

Niederhoffer, V. (1997): *Education of a Speculator*. New York: John Wiley and Sons.

Odean, T. (1998): "Are investors reluctant to realize their losses?" *Journal of Finance*, v.53, pp.1775-1798.

Ojah, K.; Karemera, D. (1999): "Random walks and market efficiency tests of Latin American emerging equity markets: A revisit". *The Financial Review*, v.34, pp.57-72.

Olowe, R. A. (1999): "Weak Form Efficiency of the Nigerian Stock Market: Further Evidence". *African Development Review*, n.1, pp.54-67.

O'Neill, M. (2001): "Automatic Programming in an Arbitrage Language: Evoluting Programs with Grammatical Evolution". *PhD thesis*, University of Limerick, Ireland 2001.

O'Neill, M.; Ryan, C. (2001): "Grammatical evolution". *IEEE Trans. Evolutionary Computation*, v.5, iss.4, pp.349-358.

O'Neill, M.; Ryan, C. (2003): *Grammatical evolution: Evolutionary Automatic Programming in an Arbitrary Language*. Boston: Kluwer Academic Publishers.

Opong, K.K.; Mulholland, G.; Fox, A.; Farahmand, K. (1999): "The behaviour of some UK equity indices: An application of Hurst and BDS test". *Journal of Empirical Finance*, v.6, iss.3, pp.267-282.

Osborne, M. (1959): "Brownian Motion in the Stock Market". *Operations Research*, v.7, pp.145-173.

Osborne, M. (1962): "Periodic Structure in the Brownian Motion of Stock Prices". *Operations Research*, v.10, pp.345-379.

Patel, J.; Zeckhauser, R.; Hendricks, D. (1991): "Invisible-hand explanations". *American Economic Review*, May, v.84, iss.2, pp.314-318.

Patterson, D.W. (1996): *Artificial Neural Networks. Theory and Applications*. Prentice.

Perold, A.F. (2004): "The Capital Asset Pricing Model". *Journal of Economic Perspectives*, v.18, iss.3, pp.3-24.

Pinker, S. (1993): *The Language Instinct*. New York: Morrow.

Pinker, S. (1997): *How the mind works*. New York: W.W. Norton and Co.

Pontiff, J. (2006): "Costly arbitrage and the myth of idiosyncratic risk". *Journal of accounting and Economics*, v.42, pp.35-52.

Poterba, J.M.; Summers, L.H. (1988): "Mean reversion in stock prices: Evidence and implications". *Journal of Financial Economics*, n.22, October, pp.27-59.

Price, K. (1999): "An introduction to differential evolution" in Corne, D.; Dorigo, M.; Glover, F. (eds): *New Ideas in Optimization.* London: McGraw-Hill.

Reilly, F.; Wright, D. (1988): "A comparison of published betas". *Journal of Portfolio Management*, Spring 1988, pp. 64-69.

Rendleman, R. J.; Jones, C. P.; Latané, H.A. (1979): "Efficiency of the Treasury Bill Futures Market". *Journal of Finance*, September.

Richards, A. J. (1997): "Winner-loser reversals in national stock market indexes: can they be explained?". *Journal of Finance*, v.52, pp.2129-2144.

Richardson, M. (1993): "Temporary components of stock prices: A skeptic's view". *Journal of Business and Economic Statistics*, v.11, pp.199-207.

Richardson, M.; Stock, J. H. (1989): "Drawing inferences from statistics based on multiyear asset returns". *Journal of Financial Economics*, v.25, pp.323-348.

Roberts, H. (1959): "Stock-Market 'Patterns' and Financial Analysis: Methodological Suggestions". *Journal of Finance*, v.14, pp.1-10.

Roll, R.; Ross, S. (1980): "An Empirical Investigation of the Arbitrage Pricing Theory". *Journal of Finance*, December.

Roll, R.; Ross, S. (1984): "A critical reexamination of the Empirical Evidence on the Arbitrage Pricing Theory: A Reply". *Journal of Finance*, june.

Ross, S. (1976): "The Arbitrage Theory of Capital Asset pricing". *Journal of Economic Theory*, v.13, pp.341-360.

Rubinstein, M. (1976): "The valuation of Uncertain Income Streams and the Pricing of Options". *Bell Journal of Economics and Management Science*, v.7, pp.407-425.

Ruggiero, M. A. (1997): *Cybernetic Trading Strategies*. New York: Wiley.

Samuelson, P. (1965): "Proof that properly anticipated prices fluctuate randomly". *Industrial Management Review*, n.6, pp.41.

Scheinkman, J. A.; LeBaron, B. (1989): "Nonlinear dynamics and stock returns". *Journal of Business*, v.62, pp.311-338.

Schoenfeld, S. A. (2004): *Active Index Investing*. New Jersey:Wiley.

Schumpeter, J. (1939): *Business Cycles: A Theoretical, Historical and Statistical Analysis of the Capitalist Process* . New York: McGraw-Hill.

Sciubba, E. (2006): "The Evolution of Portfolio Rules and the Capital Asset Pricing Model". *Economic Theory*, v.29, iss.1, pp.123-150.

Seyhun, H. N. (1998): *Investment Intelligence From Insider Trading*. Cambridge: MIT Press.

Sharpe, W.F. (1963): "A simplified model of portfolio analysis." *Management Science*, v.9, pp.277-293.

Sharpe, W.F. (1964): "Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk". *Journal of Finance*, v.19, iss.3, pp. 425-442.

Sharpe, W.F. (1991): "The Arithmetic of Active Management". *Financial Analysts Journal*, (January/February 1991).

Shiller, R.J. (1984): "Stock prices and social dynamics". *Brookings Papers on Economic Activity*, v.2, pp. 457-510.

Shiller, R. J. (1989): *Market Volatility*. Cambridge, Mass: MIT Press.

Shiller, R. J. (2003): "From efficient markets theory to behavioural finance". *Journal of Economic Perspectives*, v.17, n.1 pp. 83-104.

Shleifer, A. (1986): "Do demand curves for stocks slope down?". *Journal of Finance*, v.41, pp.579-590.

Shleifer, A. (2000): *Inefficient markets: An introduction to behavioural finance*. Oxford: Oxford University Press.

Simon, H. (1955): "A Behavioral Model of Rational Choice". *Quarterly Journal of Economics*, v.69, pp.99-118.

Smith, R.L.; Smith, T.M. (2001): *Ecología*. Madrid:Pearson Addison Wesley.

Solnik, B. (1974): "An Equilibrium Model of the International Capital Market". *Journal of Economic Theory*, v.8, pp.500-524.

Solnik, B. (1993): "The Performance of International Asset Allocation Using Conditioning Information". *Journal of Empirical Finance*, n.1, iss.1, pp.33-55.

Soydemir, G.A. (2005): "Differences in the Price of Risk and the Resulting Response to Shocks: An Analysis of Asian Markets". *Journal of International Financial Markets, Institutions and Money*, v.15, iss.4, pp. 285-313.

Stambaugh, R. F. (1982): "On The Exclusion of Assets from Tests of the Two-Parameter Model: A Sensitivity Analysis". *Journal of Financial Economics*, v.10, iss.3, pp.237-268.

Stapleton, R.; Subrahymanyam, M. (1977): "Multi period equilibrium asset pricing model". *Econometrica*, v.46.

Storn, R.; Price, K. (1995): "Differential evolution-a simple and efficient adaptive scheme for global optimization over continuous spaces". *Technical Report TR-95-012: International Computer Science Institute*, Berkeley, 1995.

Storn, R.; Price, K. (1997): "Differential evolution-a simple and efficient heuristic for global optimization over continuous spaces". *Journal of Global Optimization*, v.11, pp.341-359.

Strong, N.; Xu, X. G. (1997): "Explaining the cross-section of UK expected stock returns". *British Accounting Review*, v.29, pp.1-24.

Stulz, R.M. (1981): "A Model of International Asset Pricing". *Journal of Financial Economics*, v.9, pp.383-406.

Summers, L.H. (1986): "Does the stock market rationally reflect fundamental values?". *Journal of Finance*, v.41, pp. 591-601.

Timmermann, A.; Granger, C.W.J. (2004): "Efficient market hypothesis and forecasting". *International Journal of Forecasting*, v.20, pp.15-27.

Tobin, J. (1958): "Liquidity preferences as Behavior Towards Risk". *Review of Economic Studies*, v.26, iss.2, pp. 65-86.

Torra, S. (2004): "Siniestralidad en seguros de consumo anual de las entidades de previsión social. La perspectiva probabilística y econométrica. Propuesta de un modelo econométrico neuronal para Cataluña". *PhD thesis*. Universidad de Barcelona, Barcelona 2004.

Trivers, R. (1985): *Social Evolution*. Menlo Park, NJ: Benhamin/Cummings.

Trivers, R. (1991): "Deceit and Self-Deception" in Robinson, R.; Tiger, L. (ed): *Man and Beast Revisited*. Washington, DC: Smithsonian Press.

Tse, Y. (1999): "Round-the-Clock Market Efficiency and Home Bias: Evidence from the International Japanese Government Bonds Futures Markets". *Journal of Banking and Finance*, v.23, iss.2, pp.1831-1860.

Urrutia, J.L. (1995): "Tests of random walk and market efficiency for Latin American emerging markets". *Journal of Financial Research*, v.18, pp.299-309.

Varetto, F. (1998): "Genetic algorithms in the analysis of insolvency risk". *Journal of Banking and Finance*, v.22, iss.10, pp.1421-1439.

Waldman, M. (1994): "Systematic Errors and the Theory of Natural Selection". *American Economic Review*, v.84, n.3, pp.482-497.

Wilson, E. (1975): *Socio-biology: The New Synthesis*. Cambridge, Mass.: Belknap Press of Harvard University Press.

Working, H. (1960): "Note on the correlation of first differences of averages in a random chain". *Econometrica*, v.28, pp.916-918.

Wright, J. H. (2000): "Alternative variance-ratio tests using ranks and signs". *Journal of Business and Economic Statistics*, v.18, pp.1-19.

Yadav, P. (1999): "Non-linear Dependence in Stock Returns: Does Trading Frequency Matter?". *Journal of Business Finance and Accounting*, v.26, iss.5-6, pp.651-679.

Yao, J.; Tan, C. (2000): "A case study on using neural networks to perform technical forecasting of forex". *Neurocomputing*, v.34, pp.79-98.

Zarowin, P. (1989): "Does the stock market overreact to corporate earnings information?". *Journal of Finance*, v.44, pp.1385-1399.

Zhang, Y-C. (1999): "Toward a theory of marginally efficient markets". *Physica* , v.269, pp.30-44.

Zhang, Z.; Lai, R. N. (2006): "Pricing Efficiency and Arbitrage: Hong Kong Derivatives markets Revisited". *Applied Financial Economics*, v.16, iss.16, pp.1185-1198.

Zirilli, J. (1997): *Financial Prediction Using Neural Networks*. London:Thomson Computer Press.

Zou, L. (2006): "The Best-Beta CAPM". *Applied Financial Economics Letters*, v.2, iss.2, pp. 131-137.

CHAPTER 2. MARKET INDEXES IN DETAIL: BIASES REVEALED

ABSTRACT

According to traditional financial theory, passive investing is an optimal strategy for the average investor. In practice, however, markets are not totally efficient, and indexes, even in a perfect market, are not always well designed. Passive investors must be aware that their investments are affected by five index biases: i)The sample bias, which is a statistical bias between population and sample, and is especially important in Hedge Fund, Fixed Income or Real Estate Indexes; ii) The construction bias, which appears when an index is not built using a capitalization-weighted Laspeyres methodology; iii)Tracking error, which is the difference between a portfolio and the index it tries to proxy; iv)The efficiency bias, which appears when indexes and market portfolios are not on the efficient frontier. In this case, mean-variance efficient portfolios can obtain better risk-adjusted returns; v) The active bias, which compares the market return with the best achievable return (opportunity cost). This article discusses index foundations, uses and problems and thereby aims to redress the surprising absence of deep index analysis in the literature.

1. INTRODUCTION

Traditional financial approaches are clear: passive investing is the optimal strategy a rational investor must follow if he/she invests in an efficient zero-sum game market where CAPM hypotheses are fulfilled. A passive strategy means an investor buys and holds the total market, but because this is impossible, a proxy is needed: a market index. In their search for new tools, practitioners and academics have developed hundreds of indexes, have divided markets into pieces and have proxied markets, strategies and assets, taking sometimes for granted all these innovations. All these indexes can cover certain commercial necessities or even offer investors solutions to specific needs, but they are sometimes quite distant from the main objectives of an index. Given the essential role of indexes in traditional and modern finance, it is surprising that no deep analysis has been made into how they are theoretically justified or should be constructed. In this article I aim to continue the discussion started in Andreu (2008) and reflect on the theoretical foundations of indexing, its traditional and modern uses, current index characteristics, and the problems regarding indexing.

The main objectives of the paper are twofold: to discuss index foundations and to pay special attention to market index biases. Passive investors following market indexes must be aware that their investment results are affected by the following five market index biases. The sample bias prevents the index from obtaining the market return because of a statistical bias between population and sample. How this bias is treated, especially in Hedge Fund, Fixed Income or Real Estate Indexes is of vital importance. The sample bias is sometimes created by 'subjectivity' in the application of index governance rules; I call this the selection bias. The construction of the index, the weights used by vendors, and the inclusions, exclusions or rebalancements are not trivial. One index can generate very different risk-adjusted returns depending on how it is calculated. Theoretically, an index must be constructed using a capitalization-weighted Laspeyres methodology, adjusting this for benefits not incorporated into prices and for inflation. However, in practice, other criteria are used which create the construction bias. Tracking error is the difference between a portfolio and the index it tries to proxy; this difference is caused by commissions and turnover costs. Although an exact replication is impossible, tracking error and its exogenous and endogenous sources must be controlled and reduced if possible. The last two biases focus on how inefficiencies in the market create opportunities for different investment strategies. The efficiency bias ensures that it is difficult to find indexes and market portfolios on the efficient frontier, meaning that mean-variance efficient portfolios can obtain better risk-adjusted returns. There are different ways to create efficient portfolios, but the traditional strategy is return maximization or risk minimization using traditional or new risk measures as VaR [Andreu and Torra (2008)]. Finally, the active bias compares the market return with the best achievable return (opportunity cost). It is absolutely essential in my opinion to also index active strategies to provide a reference point for evaluating passive investors' performance and active managers' activity.

The structure of the paper is as follows. In Section 2 I revisit the origins of indexing, its theoretical foundations and its functions and characteristics. This analysis is completed in Section 3 with a study of the aforementioned market index biases: the sample bias, the construction bias, tracking error, the efficiency and the active bias. Section 4 contains the conclusions, Section 5 suggests future research, Section 6 contains the appendix, and Section 7 the references.

2. REVIEW OF MARKET INDEXES

2.1. ORIGINS OF INDEXING

Both practitioners and academics have contributed to the development of indexing. Practitioners created the first equity indexes in 1884 (Dow Jones Average), 1896 (Industrial Average) and 1923 (S&P's first capitalization-weighted index). On the other hand, the advances in Modern Portfolio Theory (MPT), the CAPM and the expansion of the efficiency analysis [Markowitz (1952), Sharpe (1964), Kendall (1953) or Fama (1970)] provided theoretical bases for passive investment strategies. However, indexing was ignored during its initial years[1] because active

1 Only Wells Fargo, American National and the Vanguard Group launched indexed funds in the 1970s.

managers were thought to be better at selecting winning stocks. After the 1970s, disillusionment with active management started to become widespread and passive investment was seen as an alternative. Malkiel (1973) and Samuelson (1974) urged the introduction of mutual funds tied to indexes, and Ellis (1975)[2] showed that 85% of active managers failed to beat the S&P500. Indexing grew exponentially and was supported by lower management costs, simplified manager selection and evaluation, superior average performance, and the enthusiasm for EMH and CAPM. Empirical evidence is clear on this point: in the last 20 years indexed institutional assets have grown 40% annually, although indexing is still a small part of the investing universe[3].

Indexing has also spread to other assets, although it is yet to be consolidated. Fixed Income Indexes appeared in the 1970s, and nowadays there are indexes for analytical markets such as Treasury, Agency, or Mortgage-backed Securities (MBS). Although the basic principles of indexing can be applied to these assets, they have special characteristics that are difficult to manage such as maturity, price formation and diversity. Commodity indexes based on commodity futures have been provided by S&P-Goldman Sachs, the Dow Jones Company or Reuters-Jefferies-CBR since the 1950s, and Real Estate Indexes appeared in the 1960s. The NAREIT Index, the Morgan Stanley or Dow Jones Wilshire REITs Indexes are some of the most well known. Hedge Funds[4], with their diverse strategies and data problems have been indexed since the 1990s by CSFB-Tremont or MSCI, although Hedge Fund Indexes may be considered a coarse approach to indexing. Other asset classes such as Private Equity, Venture Capital or Art have nascent indexes, but these have yet to be fully developed. For a brief summary of the origins and literature on indexing see Table 1 and Table 2 in the appendix.

2 See Ellis (1995), reprinted from 1975.
3 Institutional indexed investments were 35% (USA) and 20% (UK) in 1999. Indexed equity mutual funds were only 5% (USA) in 1996 and equity indexed assets were only 10% (USA) in 1999 [Shoenfeld (2004), Hutchinson (1999), Malkiel (2000) or Malkiel and Radisich (2001)]. See Figure1 in the appendix for more information about the evolution of US institutional indexed assets.
4 For more information about Hedge Funds, see Brown *et al.*(1999), Lamm (1999), Purcell and Crowley (1999), Agarwal and Naik (2000) or Gregoriou *et al.* (2005).

2.2. THE THEORETICAL FOUNDATIONS OF MARKET INDEXES

In a *perfect market*,[5] a passive strategy (buying and holding the total market) is optimal. An investor should hold a capitalization weighted portfolio that replicates the whole market (commonly known as a market portfolio.) Indexes were developed to help passive investors realize this objective. Theoretically, a market index must proxy the total market using a capitalization approach thereby becoming a reference market portfolio (benchmark). Because indexing has basically been developed by practitioners, and for practical reasons, market indexes do not always replicate complete markets and are usually only 'samples'. Also, thousands of indexes try to slice and break down analytical markets in industries, sectors, subsectors, even investment strategies. Moreover, indexes are calculated using different methodologies that do not always follow theoretical justifications. It can said, therefore, that market indexes are affected by certain biases.[6] These are (i) the sample bias: this is the difference between the return of an index and the total market return it tries to proxy. It is a statistical bias between the population (universe) and the sample. Within this bias one special issue must be analysed: the selection bias. The selection bias is a source of sample bias generated during the selection of the index's components; (ii)The construction bias: this is the difference between the return of an index (calculated using its individual weighting criterion and methodology) and the total market return. Theoretically, an index must be capitalization weighted, meaning that the further we go from this weighting, the more we increase the construction bias; (iii)Tracking error: this is the difference between the index (calculated without fees or transaction costs) and the real profitability and risk obtained by a passive investor following the index.

Regarding the fourth and fifth biases, it must first be said that perfect markets do not exist. Markets are not zero-sum games, EMH is not always fulfilled and the CAPM hypotheses are not real assumptions. The conclusions of MPT in an imperfect market are also clear: the market portfolio does not provide the best mix of risky securities

5 A perfect market is defined as an efficient zero-sum game where the CAPM hypotheses are fulfilled.
6 See Andreu (2008) for a detailed analysis of these conclusions.

for everybody and a passive strategy is not optimal and not located in the efficient frontier [Markowitz (2005), Hsu (2006)]. In this situation, the two remaining biases must be taken into account. These are (iv) The efficiency bias: this is the difference between the return of the market portfolio and the return that the efficient frontier pays for the assumed risk; and (v) The active bias: this is the difference between the total market return and the best investment opportunity using the best investment strategy (opportunity cost).

This brief summary shows, therefore, that indexes have become more complex than they should be given their theoretical foundations. Today, dozens of index providers, local exchanges and financial firms calculate indexes by proxying markets, countries, sectors, asset classes or investment styles and by using lots of techniques and methodologies. There are always biases, and if these are really big for a specific index, a passive strategy following that index is sub-optimal, although such a strategy can also be understood as a hidden active strategy. Is it passive to follow a negotiation weighted sectorial market index of only 40 assets selected by a very subjective committee? I do not think so. With the theoretical analysis presented here I hope to conclude the discussion started in Estrada (2006) and Burr (2005).

Given the future we can expect for indexing, it is necessary here to continue the study started in Andreu (2008). Indexes' objectives, functions, and biases must be carefully observed and useful guidelines for active and passive investors must be established.

2.3. FUNCTIONS OF AN INDEX

As can be seen in Andreu (2008), it is possible to define two types of markets: i) an Analytical Financial System (Analytical Market), which is a financial system with a Physical or Electronic (POE) market where only one type of asset is negotiated; ii) A Global Financial System (Global Market), which is a theoretical financial system that includes all products, investors, financiers, POE markets, techniques and intermediaries and can be defined as a Global World-wide market without boundaries, or as countries or monetary regions within certain boundaries.

The main objective of an index in a perfect market is to proxy the return and risk in an Analytical or Global Market. I shall call this objective performance measurement. An analytical market index represents the performance of an asset class whereas a global market index represents the global performance of all existing assets. Thus, indexes become references for evaluating active and passive managers' results or for establishing asset allocation strategies.

Table 3
Major Index Providers.

Provider	Available Indexes											
	An	Gl	St	In	Rg	Ma	Cu	He	Fi	Re	Co	Al
Dow Jones/Wilshire	✓	✓	✓	✓	✓	✓	✓	✓	✗	✓	✓	✗
FTSE International	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗	✓
Morningstar Indexes	✓	✓	✓	✓	✗	✓	✓	✗	✓	✗	✓	✗
MSCI Barra	✓	✓	✓	✓	✓	✓	✓	✓	✗	✓	✗	✓
NASDAQ/OMX	✓	✗	✗	✓	✓	✓	✓	✗	✗	✗	✗	✗
RUSSELL	✓	✓	✓	✓	✓	✓	✓	✗	✗	✗	✗	✗
S&P	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Note: Available index families for major providers are presented in the table. A tick (cross) means the provider offers (doesn't offer) this type of index. An=analytical stock market indices; Gl=global stock market indices; St=style indices (growth/value, socially responsible or strategy indices); In=industry/sector indices; Rg=regional indices; Ma=market capitalization indices (large, mid, small); Cu=customized indices. He=Hedge Fund indices; Fi=Fixed Income indices; Re=Real Estate indices; Co=commodity indices. Al=alternative strategy indices.
Source: Author's own, developed using Fabozzi (2000), Schoenfeld (2004) and Index Provider's Websites from www.IndexUniverse.com. Data up to December 2008.

When looking for ways to measure performance, practitioners have developed hundreds of indexes by dividing a market into pieces. Table 3 shows a brief summary of the most important index providers' indices. As can be seen, typical divisions in stocks are style (growth vs value[7], Socially Responsible Investments[8], or even strategy benchmarks[9]), sector, industry, region (developed vs. emergent) or market capitalization (large-cap, mid-cap, and small-cap). In Fixed Income Indexes, common divisions are based on sector, structure, maturity or rating; whereas in Hedge Fund Indexes, style is one of the preferred criteria. There is no theoretical justification to break down a market unless each piece can be treated as a 'different' asset class (a different analytical market) with different risk and return characteristics. For example, if growth and value stocks are proven to have a different associated risk and return, we can create one index for the growth analytical market, and another for the value analytical market[10]. Despite the

7 Complete information in Platt *et al.* (2004).
8 The Social Investment Forum defines this as an investment process that considers the social and environmental consequences of investments within the context of rigorous financial analysis.
9 Strategy benchmarks are developed by proxying active investment strategies [Kuenzi (2003)].
10 Value stocks are attractive for their intrinsic value, whereas growth stocks show above-average earnings growth [see Fama and French (1993) for an introduction].

historical interest in these questions, no consensus exists. Empirical literature shows contradictory results, and the category definitions are not totally accepted for all the participants in the market. It should be noted that almost all index providers and academics use a different criterion to classify stocks, bonds, or Hedge Funds into groups.[11]

The second objective of an index, related to the discussion presented above, is exposure. This objective is clearly practical. Financial institutions might want to provide their clients with exposure to styles, market subsectors, countries or active strategies. It should be remembered here that the right exposure to an efficient market is total exposure to a global or analytical market (passive strategy). Other types of exposure can be understood as active strategies where we use a market index to provide them.

Although they are less important in my view, other practical uses for indexes are as economic indicators (reflecting economic cycles bubbles, crashes, or the hopes and fears of economic actors) and as the basis for investment vehicles such as index funds, derivatives or ETFs [Schoenfeld (2004)].

2.4. BUILDING MARKET INDEXES: CHARACTERISTICS OF A 'GOOD' INDEX

Before I start a detailed analysis of market indexes, I should describe some of the characteristics[12] that a market index must have. From a theoretical point of view, the main objective of an index is to measure total market performance. This means that an index must be a representative complete investable capitalization-weighted portfolio. A market index is representative if it proxies well an analytical or global market; complete if it reflects all existing assets, investable if it uses available assets and capitalization-weighted if it is weighted according to components' capitalization. If an index has these four characteristics, it 'is' the total market, if not, biases

11 See Fama and French (1993), Fung and Hsieh (2002), Schoenfeld (2004) or Andreu (2008) for some examples of this problem.
12 A similar approach can be found in Schoenfeld's (2004) seven key criterias.

appear. From a practical point of view, the following characteristics are also interesting: (i) clear rules (index component selection, rebalancements, treatment of coupons or cash flows in Fixed Income indexes, Hedge Funds inclusions or exclusions, etc.); (ii) easy access to data (still a key issue for Fixed Income or Hedge Fund indexes); (iii) availability of index based products to reduce costs and tracking errors; (iv) low turnover and transactional costs.

3. ANALYSING MARKET INDEX BIASES

This section contains a conceptual and theoretical analysis of market index biases. At the end of the article, in the appendix, Table 4 presents market index's biases in different world indices such as stocks, Fixed Income, Hedge Funds, Real Estate and commodities.

3.1. THE SAMPLE BIAS

The sample bias is the difference between the return of an index and the total market return. This difference depends on which sample is chosen for the index. It is a statistical bias between population (universe) and sample[13]. It is also known as *representation bias*[14]. If the sample bias is not correctly taken into account, we can find that different market indexes can have very different performance and risk characteristics whilst proxying the same analytical market [Amenc and Martellini (2003) or Reilly *et al.* (1992)]. I analyse this bias in detail in the following paragraphs.

Completeness and investability are two sides of the same coin, and the trade-off between them is clear. If an index is complete (i.e. if it contains all the existing assets at each point in time) the sample bias is zero, although the tracking error increases because of illiquidity, turnover and transactional costs and because it is impossible to buy all components. On the other hand, tracking error can be

13 Different sampling strategies are commented on in the tracking error section of this article.
14 See Liew (2003).

minimized by limiting completeness and representativeness and by putting more emphasis on investability (i.e. all index components must be liquid and available, and costs must be lower) when selecting the sample, although this increases the sample bias.

Float adjustment: A market index must proxy all the available assets in an analytical market. Some assets are not really part of the market because they are held by founders, directors, officers, family trusts, governmental bodies or venture capital firms; therefore, these must be eliminated from the calculation of outstanding assets[15]. This is the principal aim of float adjustment. It is very difficult to know which assets are not available because this information is not always public. Float adjustment is also a source of tracking error for index followers because it is not clearly defined (each index vendor uses a different methodology). It is also a source of selection and sample bias if the rules used to calculate float adjustments are not totally clear or not adjusted to representativeness guides.

Selection bias: The selection bias is a source of sample bias. In theory, an index must be built using representativeness as the key consideration. In practice, other rules of governance (which have nothing in common with representativeness) are used to select index's components, and to decide the moment and magnitude of the index's rebalancements[16]. Governance rules can generate sample bias through 'subjectivity' in their application and can disturb indexes in a more active way than would a passive tool. Again, a trade-off is clear. Continuous rebalancement or very strict governance rules can be useful for reducing selection and sample biases, but they also increase costs and consequently tracking error. The selection bias is more important in narrower stock indexes (those proxying sectors, countries or investment styles) because there is no agreement on what can be called 'growth' or on how many sectors can be defined in an economy. The selection bias is also very important in other assets such as bonds, Hedge Funds or REITs where constructing a good sample of the market is difficult and sometimes subjective. Although some voices call for less transparency in index construction rules [Gastineau (2002a,

15 The cases of Japan (1980s-90s), telecom privatization in Europe (1990s) and Yahoo's inclusion in the S&P500 are clear examples of the need for this adjustment.

16 Limited components, minimum ranges of capitalization, fluctuation bands or arbitrary rebalancements are some typical governance rules usually applied by index committees. Committees can be also affected by psychological biases (euphoria or panic) or by the agent problem.

2002b)], transparency and clear governance are the cornerstone of attempts to reduce selection bias and its related tracking error.

The sample bias in Hedge Fund Indices: As we have seen before, the sample bias appears when the index portfolio does not fully represent the “universe” or “population”. The sample bias becomes especially important in Hedge Funds because it biases performance upward and risk downward [Fung and Hsieh (2002) and Park *et al.* (1999)]. There are various issues with this. First, it is difficult to define the ‘universe’ when it grows constantly and is made up of a great variety of strategies [Amenc and Martinelli (2003)]. This heterogeneity means that more empirical studies are necessary to determine whether we can speak of one or more analytical markets inside the Hedge Fund family. Second, existing Hedge Fund Indexes are calculated with a small part of the population, sometimes with a fixed number of components, and using different inclusion and classification criteria (selection bias). Third, the sample bias is affected by other well known biases in the Hedge Fund literature: the survivorship bias (only surviving Hedge Funds are usually taken into account when building the indexes), the participation bias (available information is affected by the Hedge Fund manager’s ability to stop reporting information to the index vendor), and the instant history bias (index vendors backfill a fund’s historical returns when a fund is incorporated). Finally, not all Hedge Funds are available to investors, meaning that the trade-off between completeness and investability reaches a high degree of complexity. For all these reasons, some authors have proposed different ways of creating alternative Hedge Fund Indexes: Fung and Hsieh (2002) use FOFs (Fund of Funds), and Amenc and Martinelli (2003) use the Kalman filter, principal components and portfolio techniques.

The sample bias in Fixed Income Indexes: Again, the sample bias with Fixed Income becomes especially important. Some points must be taken into account. First of all, the universe is broader than that of the stock. Second, the great variety of Fixed Income assets (maturities, risks, etc.) makes it difficult to speak about one or more analytical markets and consequently to decide whether one or more Fixed Income Indexes are needed. Third, the lack of public information about prices is also a key issue. There is no widely accepted source of bond prices, although some efforts have been made in that direction. Four, the selection bias is always present because each index provider uses different rules to select the index’s constituents. Finally, turnover costs are high in Fixed Income Indexes, so sampling strategies such as Factor-Based Optimization or Stratified sampling are necessary. Although existing Fixed Income Indexes are calculated using very different rules, the statistical characteristics and high correlation between them suggest that they are measuring the same market [Schoenfeld (2004)].

The sample bias in Real Estate: the first problem with this type of asset is the magnitude of the universe. Most of the attempts to create a Real Estate Index use only traded Real Estate, so the sample bias is clear because REITs or Real Estate assets created by securitization are only a small part of the universe. Despite this general approach, some attempts try to capture a bigger part of the market [Englund *et al.* (1999)]. The second problem is heterogeneity: Equity REITs, Mortgage REITs or Hybrid REITs are only some of the huge number of Real Estate assets, and it is difficult to determine if we are dealing with one or more analytical markets. Finally, Real Estate Investment Trusts (REITs) are actively managed Real Estate portfolios. The creation of REITs Indexes as a proxy of the Real Estate analytical market must be done with care because hidden active strategies disturb the supposed passive indexes.

3.2. THE CONSTRUCTION BIAS

The construction bias is the difference between the return of an index and the real total market return and is caused by the use of different weighting criteria or methodologies in the construction of the index. The main objective of an index in a perfect market is to measure the total performance of an analytical or global market. As seen in Andreu (2008), the performance and risk of a passive strategy must coincide with the performance and risk supported by a hypothetical investor holding the total market. To get this result we must compare total market value at two points in time whilst taking into account all the benefits derived from holding the portfolio that are not incorporated into asset prices (dividends, offering rights, etc.) These are commonly known as total return indexes (as opposed to net return indexes[17]). The only theoretically based tool to achieve this aim is a Laspeyres index

17 Market Indexes measuring returns nets of dividends and rights offerings.

with capitalization weightings, which has been adjusted for inflation and for benefits not incorporated into prices. Despite this evidence, index vendors use different weighting criteria and construction rules to calculate market indexes. The use of different weightings is done to capture consciously or not the efficient and active bias. Equal-weightings, GDP weightings or alternative weightings are less affected by the incorporation of new assets to the market, but do not hold the macro-consistency criterion and can be interpreted as hidden active strategies. In the following paragraphs, I revisit the most common construction rules and study how they generate the construction bias, paying special attention to this bias in Hedge Funds, Fixed Income and Real Estate.

Available Data: The basic input needed for a good market index is information. Although most analytical markets provide enough trustworthy information, this is not the case for all markets. Emerging Stock, Fixed Income or Hedge Funds markets do not always provide all the necessary information: thin trading, bond price databases that are not widely accepted and even non-coinciding closing prices are part of this problem.

Weighting criteria

1) Market-Capitalization-Weighting (Value Weighting): This is the only theoretically based weighting criterion. It is consistent with a buy and hold strategy, with the Modern Portfolio Theory (MPT) and the macro-consistency criterion. It also has low turnover costs [Schoenfeld (2004), Cloyd *et al.* (2004) and Arnott *et al.* (2005)]. As was shown previously, the capitalization weighting must take into account float adjustments.

2) Equal-Weighting: Here the portfolio's components are held in the same proportion. This is possibly because for every defined period, the index is rebalanced by selling those assets that have risen and by buying those that have suffered a price decline. This weighting criterion can be seen as a pseudo-active strategy for two reasons. First, it is a contrarian strategy that tries to capture mean reversion by selling 'winners' and buying 'losers'. Second, it captures the small firm effect thus

giving more weighting to smaller assets. Although this weighting criterion could generate higher risk adjusted returns through the capture of part of the efficient and active bias [Hamza *et al.* (2006)], it suffers from higher transactional and illiquidity costs. Liquidity-Tiered Weighted indexes or Structured-Tiered Weighted indexes are sophisticated variants of equal-weighted indexes.

3) GDP-Weighting: In this instance, GDP is used to weight countries in the construction of worldwide market indexes or emergent market indexes. Because the weighting criterion is the GDP, indexes are rebalanced yearly. Although this weighting can produce superior returns, it has problems with the delay and reliability of the economic data and suffers from sample and construction bias.

4) Price-Weighting: Here price is the only underlying variable taken into account, which means a higher priced stock will have a higher weighting in the index. A price-weighted index measures the average price level of the market. The index value is obtained by dividing the sum of the current prices of the component stocks by a divisor. Price Indexes suffer from a higher tracking error because it is difficult to track this type of index and from a sample bias and construction bias.

5) Alternative weighting criteria: This allows us to have as many alternative weightings as we like. Lo (2001) even suggests personal indexes (with personalized weightings) could be available in the near future. Until personal indexes arrive, some indexes are weighted using price, negotiation criteria, volume of operations or fundamental rules[18].

Construction rules: Mathematical rules that are used to build market indexes are called construction rules. First of all, a market index should be a Laspeyres capitalization-weighted index[19], but index vendors construct Paasche, Modified Laspeyres, Marshall-Edgeworth or Arithmetic Price Mean indexes and these are followed by hundreds of passive investors. In the end, I have decided not to include a complete list of most common index construction formulae here, although these

18 Fundamental Indexes in Arnott *et al.* (2005), Estrada (2006) or Hsu (2006).
19 An arithmetic capitalization weighted mean is equivalent to a Laspeyres Index.

can be found in Broby (2007). Second, another important fact is how special events are treated. How cash/stock elections, dividends, mergers, spin-offs, splits or bankruptcies are treated is not a trivial matter. As a general rule, it should be remembered that all necessary adjustments must look for the correct measurement of an analytical market's performance. Finally, periodical rebalancements affect market and indexer's performances[20] and reduce or increase the *sample bias and tracking error.*

The construction bias in Hedge Fund Indexes: Almost all Hedge Fund indexes are equal-weighted. Only CSFB/Tremont uses a capitalization weighting approach[21] [Amenc and Martinelly (2003), Schoenfeld (2004) and Fung and Hsieh (2002)]. More sophisticated techniques such as the pure synthetic methodology have been suggested in the literature as means of solving data restriction problems [Amenc and Martinelli (2003)]. Another important issue in the construction of Hedge Fund indexes are fees. Managers' fees must be eliminated before performance is measured, and there is no consensus as to how this must be done.

The construction bias in Fixed Income Indexes: Although almost all index providers use value-weighting and monthly reconstitution (because only some equal-weighted indexes are available), almost nobody uses the same rules to treat intramonth payments and principal repayments[22].

The construction bias in Real Estate Indexes: As seen before, most Real Estate Indexes are built using REITs. These indices are float-adjusted value-weighted portfolios with monthly or quarterly reconstitution. Other approaches to building Real Estate Indexes are mean or median price indexes, Crude Regression Models or more complex EQR or WRS models[23] [Englund *et al*. (1999)].

20 Reactions to reconstitution activities were detected in the S&P500 and the Russell 3000 [Siegel (2003) and Madhavan (2002)].

21 The opacity of the Hedge Fund industry makes it easier to construct Equal-Weighted Indexes than Capitalization-Weighted Indexes.

22 Despite differences in construction rules and sampling strategies, empirical evidence concludes that existing Fixed Income Indexes share similar statistical characteristics and a high correlation [Reilly *et al.* (1992) and Schoenfeld (2004)].

23 Contrary to Fixed Income Indexes, construction and sample biases generate performance and risk differences in Real Estate Indexes. For example, Breidenback *et al.* (2006) shows how betas can differ when using several Real Estate Indexes as market proxies.

3.3. TRACKING ERROR

Tracking Error has been defined as the difference between portfolios' performance when proxying the index and the index at a point in time. This difference is caused by commissions and turnover costs. Despite the importance of risk in an investment, no emphasis has been placed on applying it to tracking error measures. If e_{pt} is the tracking error of the portfolio at a specific point in time (t), R_{pt} the portfolio's return and R_b the benchmark return, tracking error is calculated as:

$$e_{pt} = R_{pt} - R_b$$

Using Chiang (1998), Frino and Gallagher (2001, 2002) and Frino *et al.* (2004) as a reference, I classify the sources of tracking error into two main groups:

a) *Exogenous component:* This is derived from factors affecting the index maintenance and from market frictions. The exogenous component is formed by: i) Index Adjustments, which are additions, deletions, share issuances, share repurchases, spin-offs, mergers, takeovers, splits and periodical changes;[24] ii) Dividends and rights offering, which are basically related to time delay in the receipt and re-investment hypothesis; and iii) Market Frictions, which are explicit transactional costs (brokerage fees and commissions) and implicit transactional costs (bid-ask spreads, price impact and wealth erosion[25]).

b) *Endogenous component*: this arises from the portfolio replication of the underlying index (proxying the market), the benchmark volatility when the portfolio is not perfectly aligned with the index, and cash flow movements in the portfolio (flow-induced trading, for example, in funds).

24 Blume and Edelen (2004) show tracking error can be very sensitive to changes in portfolio or index's weightings.
25 Wealth erosion appears when an index change is announced and becomes known to all participants. Then, it is possible to stablish a winning strategy against those who follow the index.

As we have seen, one of the most important sources of tracking error is the portfolio's replication strategy. Understanding how the portfolio proxies the benchmark is absolutely essential to understanding tracking error. The following are four different strategies for achieving this:

1) *Full replication:* This requires holding all the index's components in the exact same proportion as the underlying benchmark. This strategy minimizes the sample bias and tracking error (if all the index construction rules are clear and public), but does not eliminate them. Full replication is difficult to implement due to money restrictions, transactional costs and illiquidity questions.

2) Sampling strategies: These are sampled portfolio proxy indexes that do not hold exactly the same components, or do not hold these components in the same proportion. Here we find the following traditional statistical sampling strategies and specific financial approaches[26]: i)Traditional Sampling Strategies; these include random sampling or unequal-probability sampling and are clearly not optimal; ii) Stratified Sampling; this divides Index components into groups using dimensions. Some assets are then selected within each group and weighted using market capitalization. This approach is common to proxy Stock and Fixed Income Indexes; iii) Factor Analysis, or Principal Components Analysis; this looks for factors defining an index's performance and risk and uses them to build proxy portfolios. Other techniques such as clustering or discriminant analysis can also help achieve this objective; iv) Sample quadratic optimization; this, along with other linear techniques, constructs minimum tracking error portfolios within a set of parameters. Least Square estimators or other more robust estimators (LAV, Huber M or Tukey redescending M-estimators) are used to define weightings in proxy portfolios [Hampel *et al.* (1986), Alderson and Zivney (1989), Duarte and Mendes (1998), Rudolf *et al.* (1999), Bamberg and Wagner (2000) and Huber (2004)]; and finally, v) Blended Approaches; these combine different sampling strategies to proxy different parts of a benchmark.

26 See Benedetto and Ferreira (2000) for a more complete analysis of modern sampling.

3) *Derivatives based strategies*: These provide another way of replicating a market index by using index based derivatives. This strategy eliminates common problems regarding sample strategies, rebalancements, component changes, etc. Tracking error is then only created by transactional costs, liquidity costs and counterpart and basis risk. Sample bias in this type of portfolio is also eliminated.

4) *Naïve Replication*: Nanda and Peters (2006) show that a very long naïve buy and hold strategy can generate almost the same performance (risk-adjusted returns) than a specific index. The strategy consists of buying the existing stocks in a market and not changing them for a long time. Although this strategy should be analysed in the context of other markets, it seems an interesting idea to exploit.

Once we have understood tracking error sources, we need to measure them. Some Tracking Error Measures have been defined in the literature by using performance differences, although these calculations do not incorporate risk [see Roll (1992), Pope and Yadav (1994), Rudolf *et al.* (1999), Frino and Gallagher (2001, 2002) and Frino *et al.* (2004) for further information]:

1) *Average Absolute Tracking Error:* $TE_{1,p} = \dfrac{\sum_{t=1}^{n} |e_{pt}|}{n}$

This measure calculates the average of the absolute differences between the portfolio and the benchmark's return.

2) *Standard Deviation of Tracking Error:* $TE_{2,p} = \sqrt{\dfrac{1}{n-1}\sum_{t=1}^{n}(e_{pt} - \overline{e}_p)^2}$

This formula measures how the difference in returns varies between the index portfolio and the benchmark's return.

3) *Standard Error of the Residuals of a Return Regression:*

$$R_{pt} = \alpha_p + \beta_p \cdot R_{bt} + \varepsilon_{pt}$$

If the return of the index portfolio p is regressed on the return of the benchmark index b, the standard error of the regression equation provides an estimate of the tracking error similar to $TE_{2,p}$. Pope and Yadav (1994) note that if beta is not exactly equal to one, then the regression residuals will differ from $TE_{2,p}$.

4)*Standarized measures*: $$TE_{3,p} = \frac{TE_{1,p,i} - \overline{TE}_{1,p}}{\sigma_{1,p}}$$

$$TE_{4,p} = \frac{TE_{2,p,i} - \overline{TE}_{2,p}}{\sigma_{2,p}}$$

If we have different indexed portfolios we can define standarized measures. We subtract the average tracking error of the group of indexed portfolios ($\overline{TE}_{1,p}$, $\overline{TE}_{2,p}$) from the $TE_{1,p}$ of each portfolio and divide it by the standard deviation of the funds' tracking error ($\sigma_{1,p}$, $\sigma_{2,p}$) [Frino *et al*. (2004)].

Once tracking error has been measured, two interesting questions need to be answered. Is tracking error constant over time or does it present seasonal patterns? Is the proxy portfolio constantly underperforming or outperforming the market? Frino and Gallagher (2002) and Frino *et al*. (2004) introduce different strategies to test these ideas while analysing indexed funds' returns.

Up to now I have said that obtaining the benchmark's return seems to be impossible. In some studies, even passive investment does not seem to be a superior alternative to active funds after costs are taken into account. Chiang (1998) and Frino and Gallagher (2001, 2002) show Australian indexed funds suffer tracking errors of between 0.03 and 0.242, whereas tracking error is a little lower with American indexed funds. Despite these clear results, Nanda and Peters (2006) reach some curious conclusions in their paper. The construction of a very naïve portfolio at a specific point in time and its maintenance during a long period without any change can provide investors with an almost costless strategy to proxy the market, without following an index and its rebalancements. Treating this last result cautiously, it seems we should focus empirical studies on tracking error sources and on strategies to reduce it. In empirical terms, the literature has focused on exogenous factors and tracking error minimization models and has paid less attention to endogenous components and what I call here alternative strategies. It is obvious that funds following a full-replication have less tracking error, but differences in tracking errors when using alternative replication strategies have been not analysed.

Regarding exogenous sources of tracking error, Frino and Gallagher (2001, 2002) and Frino *et al*. (2004) analyse the S&P500 and the Australian market. They conclude that revisions and share issuances have the greatest impact on tracking error, whereas spin-offs, dividends and transactional costs are less important. The cost of trading is positively related to tracking error, whereas share repurchases are found to be negatively related. In these studies, the authors also point out seasonalities in tracking error. These are basically created in March, June, September and December, and can be partly avoided by certain buying strategies [Blume and Edelen (2004)]. Finally, there is no evidence that index funds consistently outperform or underperform the benchmark.

Regarding endogenous sources of tracking error, Blume and Edelen (2004) conclude that large index funds have lower tracking error because they usually use full-replication techniques. Another interesting result is that tracking error is positively and significantly related to cash flows and index volatility.

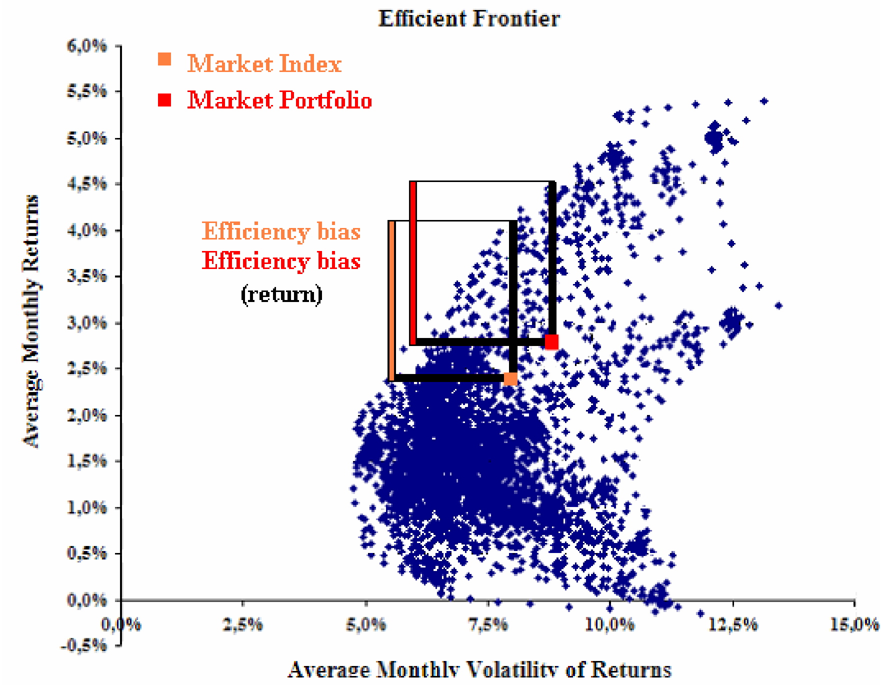
Most tracking error minimization models are linear and based on simple definitions of tracking error. From a practical point of view, quadratic objective functions are difficult to interpret so simple models have come out on top [Rudolf *et al*. (1999)].

Finally, I should make two additional comments. First, tracking error is impossible to eliminate in a portfolio that proxies the market, although it must be controlled carefully by passive investors. Tracking error is generated by exogenous or endogenous factors, but tracking error sometimes hides active strategies or active bets taken by managers. These bets are not genuine tracking error. To determine a manager's active bets, it is useful to use the Information Ratio. Second, alternative strategies are not necessarily related to a better replication of the market or a reduction in transactional costs. For example, some mutual funds and institutional investors use Internal or External Crossing, Alternative Trading Systems, Tax-Efficient Strategies or Security Lending to generate returns to compensate tracking error deviations [Schoendfeld (2004)].

3.4. THE EFFICIENCY BIAS

Efficiency bias: This is the difference between the return of the market portfolio and the return that the efficient frontier pays for the assumed risk (see Figure 2 for a graphic representation of this bias). The efficiency bias can also be defined in terms of risk. In this case it is the difference between the market portfolio's risk and the efficient frontier's risk for that return. The efficiency bias is so called because it reduces when efficiency in a financial market grows.

Figure 2
Efficiency bias in Market Indices



Note: Two hypothetical portfolios are presented in the figure. The efficiency bias has been defined from the return point of view.

Passive investors follow a specific market index. We have seen that market indices are biased, so it is easy to understand that they are not optimal portfolios but a proxy of the market portfolio (assumed to be mean-variance efficient). Roll (1992), Diacogiannis (1999) and Gómez and Zapatero (2003) show theoretically that the use of non-mean-variance efficient benchmarks generates suboptimal solutions to the investment problem. Imagine, then, the best market index you can construct, with the lowest sample bias and with correct building strategies and capitalization weightings. Although we can expect better results using this 'good' index, we are still unable to draw more optimistic conclusions regarding mean-variance efficiency. Haugen and Baker (1991) show the inefficiency of capitalization weighted portfolios, so our 'good' index clearly falls within this group. This result is confirmed in Hwang and Satchell (2002) who also showed that equally weighted portfolios are far from

Markowitz's efficient frontier. It is therefore obvious that even when they use the best index, passive investors are suffering from the efficiency bias.

Having established this, we must propose another strategy, namely, the replication of the market portfolio. In reality, financial markets are not perfect and even if the total market is held, the efficiency bias is inevitable. On balance, even if security market prices efficiently reflect the best estimation of future cash flows, and even if all investors rationally optimize the relationship between risk and expected return, the market portfolio is likely to be inefficient [Haugen and Baker (1991), Haugen (2001) or Markowitz (2005)].

Given that the efficiency bias is almost always present, it is strange that it has not been directly analysed by more authors. Indeed, some of the conclusions from well-known articles on anomalies, the CAPM utility, the calculation of betas or the validity of active investments should be clarified by taking into account the efficiency bias. The use of non mean-variance efficient portfolios leads to mistakes in portfolio performance measurement, asset allocation, risk measurement, as well as miscalculations in betas [Hwang and Satchell (2002)]. These examples are merely illustrative, but it is easy to understand the implications of the efficiency bias for every financial development.

All this suggests that the empirical analysis and measurement of the efficiency bias would provide a broad and interesting field of research for future articles. Before starting such an investigation, additional comments must be made. One way to reduce the efficiency bias is by working for a better market efficiency. The reduction of illiquidities, transactional costs or trade limitations improves market efficiency and smoothes the efficiency bias. These measures must be implemented by institutions, governmental bodies and regulatory organisations, although investors are also equipped with the tools for carrying this out. The only way an individual or institutional investor can reduce the efficiency bias is by changing their portfolio to reach a point nearer to the mean-variance efficient frontier. This should be given careful consideration by any passive investor holding the total market or worse, a market index proxying the market. How to reach better mean-variance efficient points is not a new problem. As Markowitz (1952) showed in his pioneering article, two basic strategies are available: the investor should return to benefit maximization or risk minimization. The great difference now is that this can be done with different and more powerful tools than those existing in the 1950s. The financial community can now manage risk with VaR, CVaR and Downside measures and can proxy volatilities using GARCH methodologies, tools which were not available in the early days of finance. The next question here is whether a return maximization or risk minimization strategy can be developed as a reference or benchmark once we have seen that traditional indexes are biased. Would it be beneficial to create a market index with the explicit intention of reaching points nearer to the efficient frontier? Would it be beneficial to create return maximization or risk minimization indexes? My answer is affirmative.

Andreu and Torra (2008) built Minimum Risk Indexes (MRI) to eliminate the efficiency bias by using parametric Value-at-Risk as a risk measure while constructing an ‘efficient index’ based on optimization. Although this study has received criticism from some quarters, it should be seen as a work in progress. The results in Andreu and Torra (2008) conclude that MRI are more mean-variance efficient than existing market indexes. Other studies have indirectly discovered strategies to reduce the efficiency bias in market indexes. Different weighting criteria (equal or GDP weightings) have been exploited to construct mean-variance efficient portfolios. Jobson and Korkie (1981) and Nanda and Peters (2006) found equally weighted portfolios can be efficient and can outperform a cap-weighted index; Hamza *et al.* (2006) shows that GDP-weighted portfolios produce higher returns than cap-weighted indexes and the same author concludes that equal weighted portfolios dominate GDP and cap-weighted indexes in terms of adjusted risk-returns. Finally, Arnott *et al.* (2005) show that fundamental indexes[27] provide better mean-variance returns than cap weighted indexes, these results being robust to time and cycles and to bear and bull markets.

To conclude, because the market portfolio is not on the efficient frontier, each investor must consider the best way to optimize his portfolio. Passive investors must reconsider their objectives by taking into account efficiency and index biases. The buy and hold strategy may be substituted by a pseudo-active one (e.g. fundamental

27 Fundamental indexes are built using fundamental criteria to weight index components.

or minimum risk indices) in some cases. Consequently, frontiers between active and passive investment become blurred when we penetrate the heart of finance.

3.5. THE ACTIVE BIAS

Active bias: This is the difference between the market return and the best investment opportunity using the best investment strategy (opportunity cost).

Active management has been one of the most analysed items in finance. Traditional financial theory concludes that an average investor outperforms an efficient market only by chance. Things are quite different when the market is inefficient [Etzioni (1992)] or when investors follow a market index full of construction biases. Marginal efficiency permits active investors to outperform the market using anomalies. Some anomalies disappear when they are made public because active investors trade against them recovering efficiency [Timmermann and Granger (2004), Dimson and Marsh (1999)] whereas other pockets of inefficiency are impossible to eliminate. Empirically, the usefulness of active strategies is not clear. Some studies [Sharpe (1966), Jensen (1968), Ippolito and Turner (1987), Malkiel (2000, 2003), Lakanishok *et al.* (1992)] evaluate the risk-adjusted performance of professional managers and provide evidence that average mutual funds do not outperform a buy and hold strategy, whereas Williamson (1972) and McDonald (1974) suggest mutual funds generate excess returns. Kon and Jen (1979) show investor's market timing capacities and Bauman *et al.* (2005) conclude that advisors help to generate a better investment performance. Other studies analyse predictability in financial markets using hundreds of financial variables with mixed results [see Andreu (2008) for a complete discussion]. Some conclude that predictability is possible whereas others deny this. In my opinion, some of the existing contradictions regarding the utility of active investment come from the fact that market index biases have never been taken into account. Empirical studies use market indexes without paying attention to their differences or biases or to the way they have been constructed. Most results are shown not to be robust to changes in indexes because market index biases play an important role that has never been deeply analysed.

Active investment is complementary rather than contrary to passive investment. The active bias can be solved using active or pseudo-active strategies. These strategies have been widely analysed [Grinold (1989)] but have not been indexed as a group of benchmarks that reflect opportunity costs. In my personal opinion, active benchmarks need to be developed for two reasons: 1) to measure active biases, and 2) to evaluate active managers using the correct tools. An active manager can not be evaluated using the CAPM or a cap-weighted index because alpha generation, betas and performance are not correctly calculated using traditional market indexes. Active benchmarks or strategy benchmarks [Kuenzi (2003)] seem to be the solution.

4. CONCLUSIONS

Market indexes have been used since the beginning of the 1900s. Belief in market efficiency and Modern Portfolio Theory meant that passive investment was considered optimal and that indexes quickly spread to cover all existing assets, sometimes too fast. These indexes are affected by five biases because the market is not totally efficient and because practioners and academics do not always correctly develop market indexes.

Given that an index must proxy a market, it is essential to correctly define the term 'market'. Market divisions using capitalization ranges, sectors, industries or investment styles cannot be justified if each piece cannot be treated as a different asset class (or analytical market) with different risk and return characteristics. A 'good' market index must be a *representative complete investable capitalization-weighted portfolio*: it should proxy a market that reflects all existing available assets and weights them using capitalization.

This article first looked at the sample bias (i.e. statistical bias between population and sample). If an index is complete, the sample bias is zero but tracking error increases because of illiquidity, turnover and transactional costs. On the other hand, limiting completeness minimizes tracking error but increases the sample bias. The selection bias is an important source of sample bias. In practice, indexes' rules of governance generate sample bias because they are 'subjectively' applied. The sample bias is especially important in Hedge Funds, Fixed Income assets and Real Estate. The construction bias (i.e. the differences between indexes' and markets' returns due to construction criteria) is specially important in Hedge Fund Indexes. Theoretically, a market index must be built using a capitalization-weighted Laspeyres methodology which is adjusted for benefits that are not incorporated into prices and for inflation; however, in practice other criteria are used. Altering the weights or the construction methodology when building an index has clear performance and risk implications. Tracking error, the third bias, can be generated by exogenous (i.e. index adjustments, dividends, rights offerings and market frictions) and endogenous sources (replication strategy). Although it is impossible to eliminate tracking error, it must be reduced and be seen as a way of determining active bets taken by portfolio managers. The efficiency bias (i.e. the difference between the market's return and the efficient frontier's return for an assumed risk) emphasizes the fact that neither market indexes nor market portfolios are located in the efficient frontier. To reduce the efficiency bias, it is necessary to invest in mean-variance efficient portfolios, building them with, for example, return maximization or risk minimization processes. This idea is partly developed with Minimum Risk Indices (MRI) by Andreu and Torra (2008). Although the approach is simple, the technique seems to produce market indexes with less risk and higher return. Finally, the active bias (i.e. the difference between a market's return and the best investment opportunity) can be reduced by using active or pseudo-active strategies. This paper has not tried to analyse active strategies although it should be noted that active strategies should be benchmarked as references of active bias and opportunity costs.

5. FUTURE RESEARCH

In our opinion, a theoretical analysis of indexes' foundations, objectives and characteristics and an in depth study of index biases were necessary. This is because we have to understand why passive investors get specific results, or why an index performs in a given way before we embark on an empirical study of indexing. There was nothing in the literature that showed how proposed biases affect active and passive strategies or indexes' performance. Future research should aim to complete this theoretical approach with empirical evidence regarding statistical index properties and index bias measurement.

6. APPENDIX

Table 1
Time Line of Index-Based Vehicles and Strategies

Period	Major Developments
1890-1920s	First Stock Market Indexes (Dow Jones and S&P's pioneer indexes).
1950-60s	Modern Portfolio Theory, Efficiency and CAPM studies. Pioneering Commodity and REITs Indexes (Commodity Research Bureau Index, 1957; NAREIT Index, 1969)
1970s	First Institutional (1971/1973) and retail index fund (1976). First International Equity Index Fund (1979). Malkiel (1973) and Samuelson (1974) spread indexing. First Fixed Income Indexes.
1980s	Stock Index Futures and Options (US,1982;UK,1984; Japan,1986). Expansion of index funds. First Fixed-Income Index Fund.
1990s	Growth of index fund assets. Indexing grows in Japan and Europe. Launch of ETFs in Canada and US (TIPs, SuperShares, SPDR) (1991-1993). Investable Emerging Markets Indexes (EMI) and first EMI funds (1991-1994). Spread of Commodity and REITs Indexes (S&P-Goldman Sachs Commodity Index, 1991; Dow Jones-AIG, 1998; Dow Jones Wilshire REIT Index, 1991; Morgan Stanley REIT Index, 1994) First Hedge Fund Indexes.
2000s	Global ETF explosion (equity and fixed income). Float-adjustment. Options and SSFs on ETFs. Launch of commodity and currency ETFs. Development of alternative indexed assets (Art, Venture Capital, etc.).

Source: Author's own using Schoenfeld (2004), p.12 and p.627, and Bloomberg.

Table 2
Indexing Literature by asset class: some guidelines

Asset	Literature
Stocks	Bamberg and Wagner (2000), Frino and Gallagher (2001), Madhavan (2002), Gastineau (2002a, 2002b), Kuenzi (2003), Schoenfeld (2004), Blume and Edelen (2004) and Arnott *et al.* (2005) among others.
Fixed Income	Reilly *et al*. (1992) and Schoenfeld (2004) among others.
Hedge Funds	Park *et al.* (1999), Fung and Hsieh (2001, 2002), Amenc and Martinelli (2003) and Schoenfeld (2004) among others.
Real Estate	Englund *et al.* (1999), Schoendfeld (2004) and Breidenbach *et al.* (2006) among others.
Emerging Markets	Tyandela and Biekpe (2001), Frino and Gallagher (2002), Fernandes (2004) or Hamza *et al*. (2006) among others

Source: Author's own.

Table 4
Biases in well-know Market Indices

Stocks	Sample Bias			Construction Bias			
	Sample Bias	Float adjustment	Selection bias	Weighting	Formula	Special Events	Rebalancement
Historical							
Dow Jones Averages	✓	×	✓	Pri	A.M.	Div, Spl, Spi, Mer, Tak, Rig	Rare
FTSE 100	✓	✓	×	Cap	A.W.M.	Div, Spl, Spi, Mer, Tak, Rig	Quarterly
NASDAQ Composite Index	✓ (83.7%)	✓	×	Cap	A.W.M	Div, Spl, Spi, Mer, Tak, Rig	Ongoing
NYSE Composite Index	✓	✓	×	Cap	L.	Div, Spl, Spi, Mer, Tak, Rig	Ongoing
S&P500	✓ (75%)	✓	✓	Cap	A.W.M	Div, Spl, Spi, Mer, Tak, Rig	Ongoing
Total Market-USA							
DJ Willshire 5000	×	✓	×	Cap	A.W.M	Div, Spl, Spi, Mer, Tak, Rig	Quarterly
Morningstar US Market	× (97%)	✓	×	Cap	A.W.M	Div, Spl, Spi, Mer, Tak, Rig	Quarterly/Semi-annual
MSCI US Broad Market	× (99.5%)	✓	✓	Cap	A.W.M	Div, Spl, Spi, Mer, Tak, Rig	Quarterly/Semi-annual
Russell 3000	× (99%)	✓	×	Cap	A.W.M	Div, Spl, Spi, Mer, Tak, Rig	Annual
S&P1500	✓ (85%)	✓	✓	Cap	A.W.M	Div, Spl, Spi, Mer, Tak, Rig	Ongoing
Total Market-World							
DJ Wilshire Global	✓	✓	×	Cap	A.W.M	Div, Spl, Spi, Mer, Tak, Rig	Monthly
DJ Global Titans 50	✓	✓	✓	Cap	A.W.M	Div, Spl, Spi, Mer, Tak, Rig	Annual
FTSE All World	× (98%)	✓	✓	Cap	A.W.M	Div, Spl, Spi, Mer, Tak, Rig	Annual
MSCI ACWI	×	✓	✓	Cap	A.W.M	Div, Spl, Spi, Mer, Tak, Rig	Quarterly/Semi-annual
Russell Global	× (98%)	✓	×	Cap	A.W.M	Div, Spl, Spi, Mer, Tak, Rig	Annual
S&P Global 1200	✓(70%)	✓	×	Cap	A.W.M	Div, Spl, Spi, Mer, Tak, Rig	Ongoing
S&P/Citigroup BMI	✓	✓	×	Cap	A.W.M	Div, Spl, Spi, Mer, Tak, Rig	Annual
Emerging Markets							
DJ Wilshire Emerging M.	✓	✓	×	Cap	A.W.M	Div, Spl, Spi, Mer, Tak, Rig	Monthly
FTSE Emerging M.	✓	✓	×	Cap	A.W.M	Div, Spl, Spi, Mer, Tak, Rig	Annual
MSCI Emerging M.	✓	✓	×	Cap	A.W.M	Div, Spl, Spi, Mer, Tak, Rig	Quarterly/Semi-annual
S&P Emerging M.	✓	✓	×	Cap	A.W.M	Div, Spl, Spi, Mer, Tak, Rig	Quarterly/Semi-annual

Note: The table does not try to include all existing market indices. *Sample and Selection bias*=a tick (cross) means the index presents a high (low) bias. %= market percentage captured by the index. *Float adjustment*= a tick (cross) means the index has (does not have) free float adjustment. *Weighting*: Pri=Price, Cap=Capitalization. *Formula*: A.M.=Arithmetic Mean, A.W.M.=Arithmetic Weighted Mean, E.W.A.=Equal Weighted Average. L=Laspeyres (L=AWM if weighting=capitalization). *Special Events (Adjustments)*: Div=Dividend, Spl=Split, Spi=Spin-off, Mer=Merger, Tak=Takeover, Rig=Rights offering. *Rebalancements*=rebalancement periodicity.
Source: Author's own using Index Providers' Websites from www.IndexUniverse.com. Data up to December 2008.

Table 4 (continuation 1)
Biases in well-know Market Indices

Stocks and Fixed Income	Sample Bias			Construction Bias			
	Sample Bias	Float adjustment	Selection bias	Weighting	Formula	Special Events	Rebalancement
Style Indexes							
DJ Wilshire Style	✓	✓	✓	Cap	A.W.M	Div, Spl, Spi, Mer, Tak, Rig	Quarterly/Semi-annual
DJ Sustainability World	✓(10%)	✓	✓	Cap	A.W.M	Div, Spl, Spi, Mer, Tak, Rig	Annual
FTSE4Good	✓	✓	✓	Cap	A.W.M	Div, Spl, Spi, Mer, Tak, Rig	Semi-annual
FTSE RAFI	✓	✓	✓	Fac	¿?	¿?	Annual
Morningstar Style	✓	✓	×	Cap	A.W.M	Div, Spl, Spi, Mer, Tak, Rig	Semi-annual
Morningstar Wide Moat	✓	✓	✓	Equ	L.	Div, Spl, Spi, Mer, Tak, Rig	Quarterly
MSCI Global Value&Growth	✓	✓	✓	Cap	¿¿??	Div, Spl, Spi, Mer, Tak, Rig	Semi-annual
MSCI Equal Weigthed	✓	✓	✓	Equ	E.W.A.	Div, Spl, Spi, Mer, Tak, Rig	Quarterly/Semi-annual
Russell Value&Growth	✓	✓	✓	Cap	A.W.M	Div, Spl, Spi, Mer, Tak, Rig	Annual
S&P Shariah	✓(60%)	✓	✓	Cap	A.W.M	Div, Spl, Spi, Mer, Tak, Rig	Monthly
Jantzi Social Index	✓	✓	✓	Cap	A.W.M	Div, Spl, Spi, Mer, Tak, Rig	Ongoing
Ethibel Sustainability Index	✓	✓	✓	Cap	A.W.M	Div, Spl, Spi, Mer, Tak, Rig	Semi-annual
Domini 400 Social Index	✓	✓	✓	Cap	A.W.M	Div, Spl, Spi, Mer, Tak, Rig	Ongoing
Fixed Income							
FTSE Global Bond	✓	✓	✓	Mak	A.W.M	Cou, Acc, Rep, Rei	Monthly
Morningstar MCBI	✓	✓	✓	Mak	A.W.M	Cou, Acc, Rep, Rei	Monthly
S&P NMBI	✓	✓	✓	Mak	A.W.M	Cou, Acc, Rep, Rei	Monthly
LehmanBrothers F.Income	✓	✓	✓	Mak	¿?	Cou, Acc, Rep, Rei	Monthly
Merrill-Lynch Global Bond	✓	✓	✓	Mak	¿?	Cou, Acc, Rep, Rei	Monthly
JP Morgan Fixed Income	✓	✓	✓	Mak	¿?	Cou, Acc, Rep, Rei	Monthly
EuroMTS	✓	✓	✓	Mak	A.W.M	Cou, Acc, Rep, Rei	Monthly

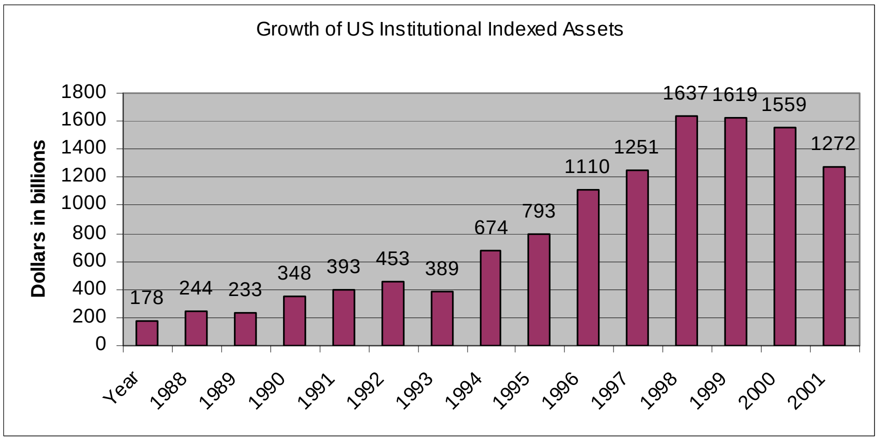
Note: The table does not try to include all existing market indices. *Sample and Selection bias*=a tick (cross) means the index presents a high (low) bias. %=market percentage captured by the index. *Float adjustment*= a tick (cross) means the index has (does not have) free float adjustment. *Weighting*: Pri=Price, Cap=Capitalization, Fac= Factors defined by the index provider, Equ=Equal, Mak=Market Value. *Formula*: A.M.=Arithmetic Mean, A.W.M.=Arithmetic Weighted Mean, E.W.A.=Equal Weighted Average. L=Laspeyres (L=AWM if weighting=capitalization). *Special Events (Adjustments)*: Div=Dividend, Spl=Split, Spi=Spin-off, Mer=Merger, Tak=Takeover, Rig=Rights offering, Cou=Coupons, Acc=Accrued interest, Rep=Principal Repayments, Rei=Reinvestment hypothesis. *Rebalancements*=rebalancement periodicity.
Source: Author's own using Index Providers' Websites from www.IndexUniverse.com. Data up to December 2008.

Table 4 (continuation 2)
Biases in well-know Market Indices

	Sample Bias						Construction Bias		
	Sample (Sa)	Selection (Se)	Float (Fl)	Survivorship (Su)	Participation (Pa)	Instant (In)	Weighting	Formula	Rebalancement
Hedge Funds									
Dow Jones Hedge Fund	✓	✓	✓	✓	✓	✓	Cap	A.M.	Monthly
FTSE Hedge Fund	✓	✓	✓	✓	✓	✓	Inv	S.I.	Annual
S&P Hedge Fund	✓	✓	✓	✓	✓	✓	Equal	E.W.A	Monthly
CASAM CISDM Hedge Fund	✓	✓	✓	✓	✓	✓	Equal	A.M.	Monthly
CSFB/Tremont	✓	✓	✓	✓	✓	✓	Cap	¿?	Quarterly
Hennesse Group	✓	✓	✓	✓	✓	✓	Equal	E.W.A	Annual
Hedge Fund Research	✓	✓	✓	✓	✓	✓	Equal	E.W.A	Monthly
MSCI Barra Hedge	✓	✓	✓	✓	✓	✓	Med	¿?	Quarterly
Van Hedge	✓	✓	✓	✓	✓	✓	Equal	E.W.A	Monthly
	Sample Bias						Construction Bias		
	Sample Bias (Sa)			Selection Bias (Se)	Float Adjustment (Fl)		Weighting	Formula	Rebalancement
Real Estate (REIT)*									
DJ Wilshire Real Estate	✓			✓	✓		Cap	A.W.M	Quarterly
FTSE EPRA/NAREIT	✓			✓	✓		Cap	P.	Quarterly
MSCI US REIT	✓			✓	✓		Cap	L.	Semi-annual
S&P/Citi Global Property	✓			✓	✓		Cap	A.W.M	Quarterly
Commodities									
Dow Jones AIG CI	✓			✓	×		Liq&Pro	¿?	Annual
Morningstar CI♠	✓			✓	×		Mag&Mom	¿?	Annual
S&P GSCI	✓			✓	×		Pro	¿?	Monthly
Goldman Sachs CI	✓			✓	×		Pro	¿?	¿?
Reuters-CBR	✓			✓	×		Equal	A.M/G.M.	¿?

Note: The table does not try to include all existing market indices. *Sample (Sa) and Selection bias (Se)* =a tick (cross) means the index presents a high (low) bias. *Float adjustment (Fl)*= a tick (cross) means the index has (does not have) free float adjustment. *Survivorship (Su), Participation (Pa) and Instant history (In) biases* = a tick (cross) means the index presents a high (low) bias. *Weighting*: Inv=Investability, Equ=Equal, Cap=Capitalization, Med=Median Asset, Liq=Future's Liquidity, Pro=Commodity Production, Mag=Magnitude, Mom=Momentum. *Formula*: S.I.=Simple Index, A.M.=Arithmetic Mean, A.W.M.=Arithmetic Weighted Mean, E.W.A.=Equal Weighted Average, G.M.=Geometric Mean, L=Laspeyres (L=AWM if weighting=capitalization), P=Paasche. *Rebalancements*=rebalancement periodicity. **All Real Estate Indexes present dividend, split, spin-offs, mergers, takeover and rights offering adjustments. ♠ Morningstar Commodity Indexes are really indexed active strategies.*
Source: Author's own using Index Providers' Websites from www.IndexUniverse.com. Data up to June 2008.

Figure 1
Growth of US Institutional Indexed Assets



Note: Following Schoenfeld (2004), p.33. Source: *Pensions and Investments*, annual May Surveys.

7. REFERENCES

Agarwal, V.; Naik, N. (2000): "On Taking the Alternative Route: Risks, Rewards, and Performance Persistence of Hedge Funds", *Journal of Alternative Investments*, v.2, iss.4, pp.6-23.

Alderson, M.J.; Zivney, T.L. (1989): "Optimal Cross-Hedge Portfolios for Hedging Stock Index Options", *Journal of Futures Markets*, v.9, pp.67-75.

Amenc, N.; Martinelli, L. (2003): "The Brave New World of Hedge Fund Indices", *Working paper*, Risk and asset management research center; Edhec, Business School, Nice.

Andreu, J. (2008): "Some traditional financial ideas revisited: the market, efficiency and indexing under new financial approaches", *Working paper*.

Andreu, J.; Torra, S. (2008): "Optimal Market Indices using VaR. A first empirical approach for three stock markets", *Working paper*.

Arnott, R.; Hsu, J.; Moore, P. (2005): "Fundamental Indexation*", Financial Analysts Journal*, v.61, iss.2, p.83-99.

Bamberg, G.; Wagner, N. (2000): "Equity Index Replication with Standard and Robust Regression Estimators", *OR Spektrum*, v.22, iss.4, pp.525-543.

Bauman, W. S.; Miller, R. E.; Veit, E. T. (2005): "Managing Portfolio Turnover: An Empirical Study", *Quarterly Journal of Business and Economics*, v.44, iss.3-4, pp.15-32.

Benedetto, J.; Ferreira, PJSG. (2000): *Modern Sampling Theory: Mathematics and Applications*, Cambridge.

Blume, M. E.; Edelen, R. M. (2004): “S&P500 Indexers, Tracking Error, and Liquidity”, *Journal of Portfolio Management*, v.30, iss.3, pp.37-46.

Breidenbach, M.; Mueller, G. R.; Schulte, K.-W. (2006): “Determining Real Estate Betas for Markets and Property Types to Set Better Investment Hurdle Rates”, *Journal of Real Estate Portfolio Management*, v.12, iss.1, pp. 73-80.

Brown, J.; Goetzmann, W.; Ibbotson, R. (1999): “Offshore Hedge Funds, Survival and Performance, 1989-1995”, *Journal of Business*, v.72.

Burr, B. (2005): “Fundamental Index Concept Gets Institutional Investors' Attention”, *Pensions & Investments*, v.33, iss.19, p.45.

Chiang, W. (1998): “Optimizing Performance”, in A. Neubert, Ed. (1998): *Indexing for Maximum Investment Results*. Chicago: GPCo Publishers.

Cloyd, P.; Siegel, L.; Schoenfeld, S.A. (2004): “Market Uncertainty and the Role of Indexing”, in Shoenfeld, S.A. (2004): *Active Index Investing*. New Jersey: Wiley.

Diacogiannis, G. P. (1999): “A three-dimensional risk-return relationship based upon the inefficiency of a portfolio: derivation and implications”, *The European Journal of Finance*, v.5, iss.3, pp.225-235.

Dimson, E.; Marsh, P. (1999): "Murphy’s law and market anomalies". *Journal of Portfolio Management*, v.25, iss.2, pp.53-69.

Duarte, A. M.; Mendes, B.V.M. (1998): “Robust Hedging Using Futures Contracts with an Application to Emerging Markets”, *Journal of Derivatives*, v.6, pp.75-95.

Ellis, C. D. (1995): “The Loser’s Game”, *Financial Analysts Journal*, v. 95.

Englund, P.; Quigley, J.; Redfearn, C.L. (1999): “The Choice of Methodology for Computing Housing Price Indexes”, *Journal of Real Estate Finance and Economics*, v.19, iss.2, pp.91-112.

Estrada, J. (2006): “Fundamental Indexation and International Diversification”, *Working paper.*

Etzioni, E. (1992): “Indexing can be beat”, *Journal of Portfolio Management*, v.19, iss.1, pp.24-26.

Fabozzi, F. J., ed. (2000): *Perspectives on Equity Indexing*. Pennsylvania: Frank J. Fabozzi Associates.

Fama, E. (1970): "Efficient capital markets: A review of theory and empirical work". *Journal of finance*, v.25, pp.383-417.

Fama, E.; French, K. R. (1993): “Common risk factors in the returns on stocks and bonds.”, *Journal of Financial Economics*, v.33, pp.3-56.

Fernandes, N. (2004): “Portfolio disaggregation in emerging market investments”, *Journal of Portfolio Management*, Winter 2004, v.31, iss.2, pp.41-49.

Frino, A.; Gallagher, D. (2001): “Tracking S&P 500 Index Funds”, *Journal of Portfolio Management*, v.28, iss.1, pp.44-55.

Frino, A.; Gallagher, D.R. (2002): “Is Index Performance Achievable? An Analysis of Australian Equity Index Funds”, *Abacus*, v.38, iss.2, pp.200-214.

Frino, A.; Gallagher, D. R.; Neubert, A. S.; Oetomo, T. N. (2004): ”Index Design and Implications for Index Tracking”, *Journal of Portfolio Management*, v.30, iss.2, pp.89-95.

Fung, W.; Hsieh, D. (2001): “The risk in hedge fund strategies: theory and evidence from trend followers”, *Review of Financial Studies*, v.14, iss.2, pp.313-341.

Fung, W.; Hsieh, D. (2002): “Hedge-Fund Benchmarks: Information Content and Biases”, *Financial Analysts Journal*, v.58, iss.1, pp.22-34.

Gastineau, Gary, L. (2002a): “Equity index funds Have lost their way” *Journal of Portfolio Management*, winter 2002, v.28, iss.2, pp.55-64.

Gastineau, Gary, L. (2002b): “Silence is Golden-The importance of Stealth in Pursuit of the Perfect Fund Index”, *Journal of indexes*, second quarter 2002.

Gomez, J-P.; Zapatero, F. (2003): “Asset Pricing Implications of Benchmarking: A Two-Factor CAPM”, *European Journal of Finance*, v.9, iss.4, pp.343-357.

Gregoriou, G.; Hubner, G.; Papageorgio, N; Rouah, F. (2005): *Hedge Funds, Insights in Performance Measurement, Risk Analysis, and Portfolio Allocation*,.New Jersey: Wiley.

Grinold, R. C. (1989): “The Fundamental Law of Active Management”, The *Journal of Portfolio Management*, v.15, n.3, pp.30-37.

Hampel, F. R.; Ronchetti, E. M.; Rousseeuw, P. J.; Stahel, W. A. (1986): *Robust Statistics: The Approach Based on Influence Functions.* New York: Wiley.

Hamza, O.; Kortas, M.; L'Her, J-F.; Roberge, M. (2006): “International Equity Portfolios: Selecting the Right Benchmark for Emerging Markets”, *Emerging Markets Review*, v.7, iss.2, pp.111-128.

Haugen, R. (2001): *Modern Investment Theory*. New Jersey: Prentice Hall.

Haugen, R. A.; Baker, H. L. (1991): "The Efficient market inefficiency of Capitalization-Weighted Stock Portfolios". *Journal of Portfolio Management*, Spring 1991, v.17, iss.3, pp.35-40.

Hsu, J. (2006): “Cap-Weighted Portfolios are Sub-Optimal Portfolios”, *Journal of Investment Management*, v.4, iss.3, p.1-10.

Huber, P. J. (2004): *Robust Statistics*. Wiley-Interscience.

Hutchinson, H. (1999): *Portfolio indexing. Theory and Practice*. New York: Wiley.

Hwang, S.; Satchell, S.E. (2002): “Calculating the misspecification in beta from using a proxy for the market portfolio”, *Applied Financial Economics*, v.12, iss.11, pp.771-781.

Ippolito, R. A.; Turner, J. A. (1987): "Turnover, fees and pension plan performance". *Financial Analysts Journal*, v.43, pp.16-36.

Jensen, M.C. (1968): “The Performance of Mutual Funds in the Period 1945-1964”, *Journal of Finance*, v.23, iss.2, pp.389-416.

Jobson, J. D.; Korkie, B. (1981): “Putting Markowitz Theory to Work”, *Journal of Portfolio Management*, v.7, iss.4, pp.70-74.

Kendall, M. G. (1953): "The analysis of Economic Time-Series, Part-I Prices". *The Journal of the Royal Statistical Society*, v.96, pp.11-25.

Kon, S.; Jen, F. (1979): “The investment performance of mutual funds: an empirical investigation of timing, selectivity and market efficiency”, *Journal of Business*, v.52, pp.263-289.

Kuenzi, D. E. (2003): “Strategy Benchmarks”, *Journal of Portfolio Management*, v.29, iss.2, pp.46-56.

Lakanishok, J.; Shleifer, A.; Vishny, R. (1992): “The impact of institutional trading on stock prices”, *Journal of Financial Economics*, v.32, pp.23-43.

Lamm, R. M. (1999): “Portfolios of Alternative Assets: why not 100% Hedge Funds?”, *Journal of Investing*, v.8, iss.4 pp.87-97.

Liew, J. (2003): “Hedge Fund Index Investing Examined”, *Journal of Portfolio Management,* Winter 2003, v.29, iss.2, pp.113-123.

Lo, A.W. (2001): “Personal Indexes”, *Journal of indexes*, q.3, pp.26-35.

Madhavan, A. (2002): ”Index Reconstitution and Equity Returns”, (www.itginc.com/research/whitepapers/madhavan/RusselStudy.pdf).

Malkiel, B. G. (1973): *A Random Walk down Wall Street*. New York: W.W. Norton & Company.

Malkiel, B. G. (2000): “Indexing Continues to be a winning strategy”, in Fabozzi, F. (2000): *Perspectives on Equity Indexing*. New Hope, Pennsylvania: Fabozzi Associates.

Malkiel, B. G. (2003): ”Passive Investment Strategies and Efficient Markets”, *European Financial Management*, v.9, n.1, pp.1-10.

Malkiel, B. G.; Radisich, A. (2001): “The Growth of Index Funds and the Pricing of Equity Securities”, *Journal of Portfolio Management*, v.27, iss.2, pp.9-21.

Markowitz, H. M. (1952): “Portfolio selection”. *Journal of Finance*, 25(1), pp.71-79.

Markowitz, H. M. (2005): "Market efficiency: a theoretical distinction and so what?". *Financial Analysts Journal*, v. 61, iss. 5, pp.17-30.

McDonald, J. G. (1974): “Objectives and Performance of Mutual Funds, 1960-1969”, *The Journal of Financial and Quantitative Analysis*, v.9, iss.3, pp.311-333.

Nanda, S.; Peters, D. J. (2006): “A Very Long-Term Buy-and-Hold Portfolio”, *Journal of Portfolio Management*, v.32, iss.3, pp.65-73.

Park, J.; Brown, S.; Goetzmann, W.; (1999): “Performance benchmarks and survivorship bias for hedge funds and commodity trading advisors”, *Hedge Fund News*, August 1999.

Platt, G.; Pope, B.; Rakvin, C. (2004): “US Equity Benchmarks-Broad-cap, size, and style indexes”, in Shoenfeld, S.A. (2004): *Active Index Investing*,. New Jersey: Wiley.

Pope, P.; Yadav, P. (1994): “Discovering Errors in Tracking Error”, *The Journal of Portfolio Management*, v.20, iss.2,pp.27-32.

Purcell, D.; Crowley, P. (1999): “The Reality of Hedge Funds”, *Journal of Investing*, Fall 1999, pp. 26-44.

Reilly, F.; Kao, G.; Wright, D. (1992): “Alternative bond market indices”, *Financial Analysts Journal*, v.48, iss.3, pp,44-58.

Roll, R. (1992): “A Mean/Variance Analysis of Tracking Error”, *The Journal of Portfolio Management*, v.18, iss.4, pp.13-22.

Rudolf, M.; Wolter, H. J.; Zimmermann, H. (1999): “A linear model for tracking error minimization”, *Journal of Banking and Finance*, v.23, iss.1, pp.85-103.

Samuelson, P. (1974): “Challenge to Judgment”, *Journal of Portfolio Management*, v.1, iss.1, p.17-19.

Sharpe, W. G. (1964): “Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk”. *Journal of Finance*, 19, pp. 425-442.

Sharpe, W. (1966): “Mutual Fund Performance”, *The Journal of Business*, v.39, iss.1, pp-119-138.

Schoenfeld, S. A. (2004): *Active Index Investing*. Hooboken, New Jersey: Wiley.

Siegel, L. (2003): “The impact of Benchmarking on Markets and Institutions”, in Siegel, L. (2003): *Benchmarks and Investment Management*, Charlottestille, VA: AIMR Research Foundation.

Timmermann, A.; Granger, C. W. J. (2004): "Efficient market hypothesis and forecasting". *International Journal of Forecasting*, v.20, pp.15-27.

Tyandela, L.; Biekpe, N. (2001): “Construction of All SADC Stock Markets Indices”, *African Finance Journal*, v.3, iss.1, pp.59-76.

Williamson, P. (1972): “Measurement and forecasting of mutual fund performance: choosing and investment strategy”, *Financial Analysts Journal*, v.28, pp.78-84.

CHAPTER 3. REFUNDING MARKET INDEXES USING VaR

ABSTRACT

Is it possible to create new Market Indices that are less risky than current ones? I propose a methodological approach to deal with this question using Value-at-Risk Minimization on the parametric VaR method. With this approach I can obtain the optimal weights each share must have in the Index to minimize Risk measured by VaR. The author apply the method to three different stock markets and estimate Covariance matrices by different length moving averages. I would like to point out two innovations in the paper. First, an error dimension has been included in the backtesting and, second, the Sharpe Ratio has been used to select the 'best' model from all models presented. Although the estimation methods used are very simple, results seem very interesting. All indices are less risky than the Spanish IBEX 35® and the Argentinian Merval (current Market Index) and, surprisingly, more profitable; this does not happen in the American DowJones℠ .This highlights two points. First, my indices could manage market risk without the problems of current risk measures [Basak and Shapiro (2001)]. Second, similar investment strategies could beat the market in some cases, thus questioning the Efficient Market Hypothesis. The possible applications of Minimum Risk Indices are clear: they could reduce the risk assumed by institutional and mutual funds that nowadays follow current Market Indices. They could also be used as a benchmark for risky assets or as a basis for developing derivatives.

A short version of this article is going to appear in Applied Financial Economics, coauthored by Salvador Torra.

1. INTRODUCTION

Can we create new Market Indices that are less risky than current ones? Could these Indices beat market yield? The results of this paper, which deals with these questions, seem to imply that the answer to both questions could be affirmative in some cases.

The market, to the contrary of what is defended in traditional finance, is sometimes a non-zero-sum game where EMH is not fulfilled. Investors' irrationality, correlation in irrational trades and limitations in arbitrage create anomalies and other inefficiencies. In that situation, the market portfolio is not the proper mix among risky securities for everybody, so a passive strategy is not optimal and is not located in the efficient frontier. Index followers could be supporting a suboptimal strategy if they do not pay attention to index biases. Market Indexes present biases depending how they are constructed, selected, and replicated (*sample bias, construction bias and tracking error*). In addition, other biases appear when the market is really inefficient (e.g. *efficiency and active biases*) [Andreu (2008)]. In this paper, I present the Minimum Risk Strategy as a way to deal with the *efficiency bias*.

There are three important reasons for creating new Market Indices. First, there is a huge interest in market risk analysis and management. This is clear from the Basel agreements and other documents[1], or from the concern regarding a bearish market context and financial bankruptcies such as the Long Term Capital Management case, or the more recent Bear Stearns or Merrill Lynch disasters. Because of their special characteristics, such interest is perhaps greater when we speak about emergent markets. Second, because of the traditional Fama idea of efficient markets and models such as CAPM, which uses market yield as an essencial parameter, a lot of money is invested by following the market. Third, the efficiency bias shows us that a capitalization-weighted Index is not always located in the efficient frontier, therefore, there are other portfolios able to beat the market with lower assumed risk. The results of the paper seem to point towards

1 See the Group of Thirty (1993), the documents of the Basel Committee on Banking Supervision or England (1997).

that possibility, at least in the Spanish and Argentinian markets. For example, if Spanish mutual funds had followed the investment strategy presented here during 2000-2004, they would have increased their profits by €58.5 billion[2].

Financial risk has historically been analyzed by multiple measures[3] and models[4]. However, the increasing volatility in financial markets, derivatives and technological advances force academics to now treat market risk from another perspective [Simons (1996) and Hendricks (1996)]. One emerging perspective is Value-at-Risk [Riskmetrics (1995)]—a measure of risk that has been rapidly and widely accepted since it was introduced in 1995. There are four VaR calculation methods: (1) the Parametric Method, with the original Riskmetrics (1995) approach or similar approximations such as Jorion's [Jorion (2001)]; (2) Historical Simulation and its evolutions [Boudoukh *et al.* (1998) and Hull and White (1998)]; (3) Stochastic Simulation (also known as the MonteCarlo Method) and its evolutions[5]; (4) Hybrid Methods such as Weighted Historical Simulation or mixed Stochastic and Historical Simulation [Boudoukh *et al.* (1998)]. Each VaR calculation method has its pros and cons and provides quite different VaR measures [Hendricks (1996)]. Also, after a method has been applied it is necessary to check its reliability using a backtesting process. All the methods discussed above have several problems regarding leptokurtosis or skewness, non-linear positions or extreme returns. Other complementary techniques have therefore been developed to study extreme returns: Stress testing [Robinson (1996)], Conditional VaR or Extreme Value Theory with the Expected Shortfall method[6]. VaR has several problems, but if we can determine a controlled scenario with some interesting conditions, traditional VaR methods are reliable enough [Danielsson *et al.* (1998)] and easier to calculate than extreme value methods.

VaR and complementary methods have grown rapidly over the last few years. However, no attention has been paid to VaR as a possible tool for market risk

2 Around €196,000 milion invested in Spanish mutual funds in 2003 according to FEFSI (Federación Europea de Fondos y sociedades de Inversión). Turnover and transaction costs are not included in the calculations.
3 See Stone (1973), Pedersen and Satchell (1998) or Nawrocki (1999).
4 See Markowitz (1952), Sharpe (1964) or Merton (1973).
5 See Pearson and Smithson (2002), Frye (1996), or empirical studies in Pritsker (1997), Abken (2000), or Gibson and Pritsker (2000).
6 See Pearson and Smithson (2002) or Neftci (2000).

management[7] or portfolio optimization[8] [Froot *et al.* (1994)]. This study is therefore involved in these areas because, in my opinion, portfolio optimization using VaR or Conditional VaR is a natural evolution of Modern Portfolio Management Theory. On the other hand, the idea of active indexing [Schoenfeld (2004)] has recently appeared as a framework to unify the traditional antagonic perspective of indexing and active investment. In this paper I propose the creation of new Market Indices through VaR minimization as another step in the direction of active indexing. These new Indices (created using VaR minimization) may be interesting for controlling market risk because better and more optimal instruments could then be used by institutions to reduce market risk. Moreover, the new indices could be used to estimate more stable Betas in the CAPM model, or as a reference for active investment strategies. With this methodology, therefore, mutual funds willing to beat the current Market Indices would have a clearer and more transparent active management benchmark to be compared with because active management performance measurement is known to be difficult today and is much criticized for its opacity.

I first discuss the theoretical framework of Minimum Risk Indices and then apply it to the Spanish, American and Argentinian Stock Markets using weekly logarithmic returns to determine the optimal weight each share must have within the index to minimize risk. Using historical data from the period 1999–2004, where we can find bearish and bullish markets, I reconstruct the performance of our Minimum Risk Indices for the 2000-2004 periods. Despite the simple methods used to solve the problem, which can be easily improved by more complex econometric methods such as GARCH, my main objective has been achieved. Minimum Risk Indices have less risk and, in the Spanish and Argentinian markets, present higher returns than current Market Indices.

This paper is organized as follows. In Part 2 I describe the theoretical framework of Minimum Risk Market Indices using VaR minimization. In part 3 I discuss the results for the three Stock Markets taken as applications of our methodology. In

7 Only a few examples can be found in Garman (1996), (1997a,c) or in Aragall (2002).
8 Only a few authors, such as Sentana (2001) or Rockafellar and Uryasev (2000), have dealt with this question.

part 4 the paper draws conclusions. In part 5 the author outlines future research. Finally, in part 6 and 7 the Appendix and References are provided.

2. THEORETICAL FRAMEWORK FOR MINIMUM RISK MARKET INDICES USING VaR MINIMIZATION. A METHODOLOGICAL PROPOSAL

If we want to use VaR as a risk management tool, we have to find a method that institutions and investors find easy to follow. These characteristics mean that parametric VaR is the most suitable method for our objectives. However, it is necessary to take into account the weaknesses of the parametric approach, which are basically related to its principal hypothesis of normality. If this Hypothesis is not fulfilled, the VaR measure will not be coherent [Artzner *et al*. (1999)]. In this study, I should be optimistic because the Central Limit Theorem should make Market Index returns similar to a Gaussian distribution if the number of shares forming the Index is high enough[9]. Indeed, as in the portfolio, I do not include non-linear positions, and I use weekly data at a 5% significance level to calculate the VaR, the parametric Gaussian approach is considered reasonably good [Hendricks (1996), Danielsson *et al*. (1998)].

The problem to solve can be written as follows:

$$\begin{aligned} & Min\left\{ \; Z_\alpha \sqrt{x'\Sigma x} \; \right\} \\ & s.a. \\ & \qquad x_i \geq 0;\; i = 1..N \\ & \qquad \sum_{i=1}^{N} x_i = 1 \end{aligned} \tag{1}$$

where Z_α is the Normal distribution value at the desired significance level, the x vector contains the weight of each share within the alternative Market Index we are trying to build, N is the number of shares forming the Index, and Σ is the

9 See Appendix for further information.

logarithmic return Covariance matrix, which is assumed to be a Multivariate Normal. This problem is easy to simplify [Peña (2002)] using a Lagrangian optimization (see Appendix). The optimization result will provide the optimal weight each share must have within the Index to minimize risk using the estimated Covariance matrix. The literature contains several methods for estimating the Covariance matrix:

a) *The Historical Volatility Method*: empirical studies show that this method is not very good because it pays no attention to time-varying volatility.

b) *The Moving Average Method:* this method provides better estimations but also has some problems [Alexander (1996)].

c) *The Exponential Weighted Moving Average Method* [Riskmetrics (1995)]. With this method, the last observations in the data receive a higher weight, which solves the problems of the Moving Average Method. The decay factor election is critical [Hendricks (1996)].

d) *The GARCH and E-GARCH Methods*[10]. With these models [Engle (1982), Bollerslev (1986) and Nelson (1991)] it is possible to deal with heteroskedastic time dependent variance. GARCH Covariance estimations are clearly better than those from simple methods, but GARCH is not very used in the professional world.

e) *Variance and covariance estimations by Implied Volatility* and other less well-known methods such as those using the expectations of experts in the financial field.

After the Covariance matrix has been estimated using one of these methods, the minimization method can be used to obtain the optimal weights each share must

10 See extended models in Johnson (2001).

have within the Index to minimize the Index's market risk. With the historical data available, I can reconstruct the performance and evolution of Minimum Risk Indices to compare returns and risks between Minimum Risk Indices and current Market Indices.

Once the reconstruction of Minimum Risk Indices is available for a certain market, the validity of each approximation must be checked by a backtesting process. This process will establish how well the model applied to the data fits the real market. The more the main objective is to analyse and control risk by VaR techniques, the more backtesting is needed to study market risk not controlled by the VaR model and extreme losses within the significance level.

Finally, there is more than one Minimum Risk Index. With each Covariance matrix estimation method, I can build a different Minimum Risk Index. There are therefore a great number of these Indices, depending on the Covariance matrix estimation method and different parameters used by each method (data availability, moving average length or decay factor). Here we can see the importance of selecting the 'best' Index from all the approximations. Not too much work is available here and I only need to point to the paper written by Sarma *et al*. (2003). In my opinion the 'best' model should be selected in accordance with two key ideas:

(1) The model's capacity to explain reality or, in other words, the model's capacity to be accepted by a periodic backtesting process. Risk not controlled by the model appears in extremely negative returns, but so far only the frequency of these extreme returns has usually been used for backtesting[11]. In this study I present a very simple way of incorporating the size of error in the traditional backtesting using the *Excess Total Loss (ETL) measure*. ETL is an ex-post measure that gives Total Losses beyond the VaR.

11 Few authors have tried to incorporate error size in the backtesting process [Blanco and Oks (2004) or Lopez (1999)]

(2) It is important to establish a relationship between return and risk in VaR measures [Dembo and Freeman (1998)]. To do so, I use Sharpe's Ratio as a first approach and leave more appropriate tools such as the Reward-to-VaR Ratio [Alexander and Baptista (2003)] for future research.

Selecting the 'best' model is a very complex process. Here I only provide some simple orientative approximations to the problem. Further research on this issue should be developed soon.

3. MINIMUM RISK INDICES IN REAL MARKETS. SOME EXAMPLES

Using the theoretical framework developed in part 2, I am able to generate Minimum Risk Indices for each Stock Market I chose. As an example of how our methodology reacts to different Market Indices, in this section I apply it to the Spanish Stock Market (developing some Minimum Risk Indices for the IBEX35®), to the American Stock Market (developing some Minimum Risk Indices for the Dow Jones Industrial Average℠), and to the Argentinian Stock Market (developing some Minimum Risk Indices for the MERVAL). These examples have two objectives. First, it is interesting to test how Minimum Risk Indices work in Stock Markets with different volatility and efficiency. Second, each of the Indices represents a different way to build a Market Index[12], and using them in my approximation is a first step to determining the importance of different weightings and construction rules in the calculation of a market index. As it can be seen in Andreu (2008), the sample and construction bias must be taken into account when analysing the performance of an index. In the Argentinian and American markets, the MERVAL and the DowJones show important construction and sample biases; first because they are weighted using negotiation and price, and second because they are not calculated using a Laspeyres capitalization approach.

12 See Appendix for more information.

The aim is to create Minimum Risk Indices based on the historical composition of the IBEX35®, the Dow Jones Industrial Average[SM] and the MERVAL for the 2000-2004 period. To put it more simply: Minimum Risk Indices would be developed by taking into account only the shares contained in each Index in each period. In this way it is possible to determine whether a different weighting in the components of the actual Indices using a VaR Minimization criterion can reduce risk and to analyse how this affects the profitability of Market Indices. I call my Indices IndexVaR35 (IVaR35 for the Spanish Market, IndexVaR30 (IVaR30) for the American Market—35 and 30 are the numbers of traditional components of the IBEX35® and Dow Jones Industrial Average[SM] (DJIA[SM])—and IndexVaRM (IVaRM) for the Argentinian Market. As I have mentioned, there is not just one Minimum Risk Index for each market, because with each estimation criterion we can create a Minimum Risk Index. The Covariance matrix was estimated in all the markets by the simplest estimation methods (the Historical method[13] and the Moving Average Method using lengths of 4, 8, 10, 15, 25, 30, 40, 52, 60, 70, 78, 85 and 100 weeks) in order to explain the method's potential benefits, although the author knows these estimates can be improved by more complex methods. In the end I decided to present IVaR35, IVaR30 and IVaRM Indices calculated only by some of these Moving Averages as being representative of the short, medium and long terms.

Covariance matrices estimated with a few data (4-30 weeks) are problematical because the minimization process is difficult or rather unstable in some cases. Short-length Moving Averages change quickly in response to financial data but they consistently underestimate the VaR value and cause problems inside the minimization process because the positive and semi-defined Variance-Covariance Matrix condition is sometimes not fullfilled, which has an impact on the backtesting process. Medium-length Moving Averages (30-52 weeks) are more stable and VaR measures closer to real values. Finally, long-length Moving Averages (e.g. 60-100 weeks) are the most stable but are less able to adapt to

13 Results in the 2002-2004 period were poor, similar to other empirical studies, so I decided not to generate a IVaR35 using this estimation method.

volatile short-term changes. Despite the limited prediction capacity of Moving Averages[14], results with these approximations are quite interesting.

3.1 VOLATILITY ANALYSIS

The basic objective of the study is, by VaR minimization, to create Minimum Risk Market Indices that are less risky than current ones. Table 1 shows clearly how my Indices are less risky than the current ones in each market because overall volatility is lower. From the data, it is easy to see how the reduction in volatitily is greater in the Spanish Market than in the American or Argentinian Markets. It also shows that, in general, the longer the moving average, the less volatile, which means that risk is reduced. This seems not to be true in all the cases with the longest moving averages (52 and 78 in the IVaR35, 78-100 in the IVaR30 and 52-78 in the IVaRM) for which volatility is more or less the same or increases slightly. As with longer lengths, it is more difficult to estimate short changes in volatility, which could mean that there is an optimal moving average length beyond which it is impossible to reduce risk using the moving average method. Improved Moving Averages (in the IVaR35 and in IVaRM) are a little more risky than those with no improvements. This result is rational because, firstly, multiple-step estimation is applied to avoid underestimating the risk and, secondly, the 0.01% weighting restricts one asset so that the portfolio can be less diversified and risk rises.

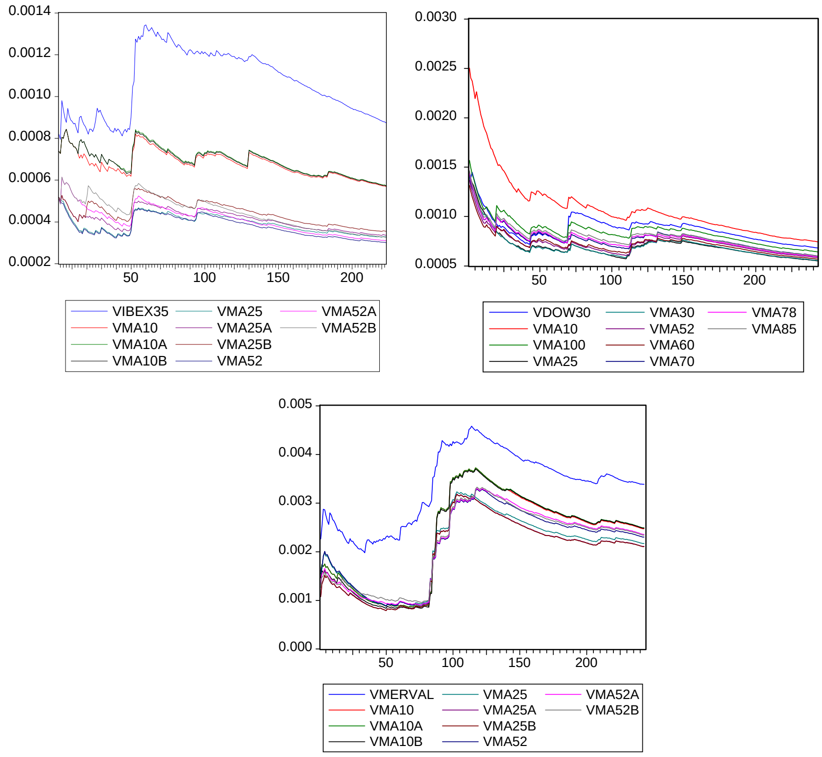
14 Hopper (1996) shows that more complex estimation methods such as GARCH do not provide better results than simple *estimation methods with long-term data.*

Table 1
Market Indices' Standard Deviation

Approximation		Standard Deviation		Standard Deviation		Standard Deviation
Current Market Index	IBEX35®	0.02958	DJIA^{SM}	0.02610	MERVAL	0.05817
Minimum Risk Index	IVaR35		IVaR30		IVaRM	
MA10	MA10	0.02387		0.02727		0.04970
	MA10a	0.02397				0.04989
	MA10b	0.02393				0.04983
MA25	MA25	0.01811		0.02347		0.04655
	MA25a	0.01835				0.04591
	MA25b	0.01884				0.04587
MA30				0.02354		
MA52	MA52	0.01736		0.02415		0.04796
	MA52a	0.01760				0.04855
	MA52b	0.01807				0.04838
MA60				0.02383		
MA70				0.02452		
MA78	MA78	0.01742		0.02433		0.04805
	MA78a	0.01835				0.04822
	MA78b	0.01867				0.04971
MA85				0.02451		
MA100				0.02533		

Volatility reduction is clearer in Figure 1, which shows cumulative volatilities. The first picture shows the Spanish Market. The first line represents the riskiest Index, which in our case is the current Market Index (IBEX 35®). The second group of lines is made up of the MA10, MA10a and MA10b approximations. The third group, with half the IBEX 35® risk, is made up of the MA25, MA52, MA78 approximations and all their modifications. The most stable approximations are the modifications a, especially MA52a, which is less risky than MA25a and MA78a. This again indicates the existence of an optimal length for moving averages beyond which it is impossible to better estimate the covariance matrix and reduce risk with moving average methods. The second picture shows the American Market. The first line again represents the riskiest Index, which is now the MA10 approximation due to problems with the positive and semi-defined covariance matrix condition. Then, and before observation 75, the second riskiest Index is the current Dow Jones Industrial Average[SM]. Below the current Market Index, and with less risk, we find all the other MA approximations. The least risky approximations are MA25 and MA30 and the more data is used to build the MA, the riskier the Moving Average approximation seems to be, which supports the idea of the optimal length for moving averages. Finally, the third picture represents the Argentinian Market. The first line (the riskiest Index) is again the current Market Index (MERVAL). Below we can see the MA10,a,b as the second group of riskiest approximations, and below this group, with less risk, the other MAs. Again we can establish the idea of optimal length for moving averages because MA25 and MA30 have less risk than MA52, MA60, MA70, MA78, MA85 and MA100.

Figure 1
Market Indices' cumulative volatility (IVaR35, IVaR30, IVaRM)



Note: here, volatility is variance. In the first, second and third graphs, VIBEX35, VDOW30 and VMERVAL are the cumulative volatilities of the IBEX35®, the DowJones Industrial Average[SM] and the MERVAL, respectively, and VMA are the cumulative volatilities of each moving average approximation used for each market.

3.2. ANALYSIS OF EXTREME LOSSES IN VaR MINIMIZATION.

Basak and Shapiro (2001) and Larsen *et al.* (2002) show that not allowing agents to assume more risk than a certain VaR value or to develop VaR minimizations can increase extreme losses, especially when return distributions are very different from Normal distributions. Following these ideas, Basak and Shapiro (2001) do not think it is useful to set limits on VaR values in institutions to control risk and Larsen *et al.* (2002) propose Conditional VaR minimizations. These results appear basically when distributions are heavily skewed or have long fat tails. In our case, the problems noticed by the above authors are not excessively important (see Table 2).

Table 2
Extreme Losses

Highest Extreme Loss (%)					
IBEX35®	11.1	DJIA℠	15.4	MERVAL	15.3
IVaR35		IVaR30		IVaRM	
MA10	11.2		10.8		12.1
MA10a	11.2				12.1
MA10b	11.2				12.1
MA25	6.6		9.2		11.4
MA25a	6.6				9.8
MA25b	6.6				9.8
MA30			8.4		
MA52	6.3		10.4		15.3
MA52a	6.8				15.3
MA52b	6.8				15.3
MA60			9.2		
MA70			10.6		
MA78	6.4		10.5		11.4
MA78a	7.6				10.9
MA78b	7.6				10.0
MA85			10.7		
MA100			10.9		

For the shortest moving average, extreme losses are similar to those of IBEX 35® and lower in the American and Argentinean market and decrease when we increased the moving average length. There is a certain moving average length when extreme losses start to rise again (Ma78 in IVaR35, MA70 in IVaR30 and more difficult to define in IVaRM), which again supports the existence of an optimal length moving average.

3.3. VaR ANALYSIS

Each approximation has a different VaR measure that evolves over time. Figure 2 shows how great changes in volatility that are common in moving average approximations are greater in short length moving averages than in long length ones. On the other hand, as these types of averages do not attach different weights to more recent data than to older data, moving averages are indicators of ‘past’ volatility, regarding inappropriate Covariance estimations when price trends change[15]. This problem decreases when the lengths are longer[16]. Finally, I should point out that short moving averages usually underestimate VaR, so the losses beyond the VaR will be more frequent in those cases. The longer the moving average length, the lesser the underestimation of the VaR measure.

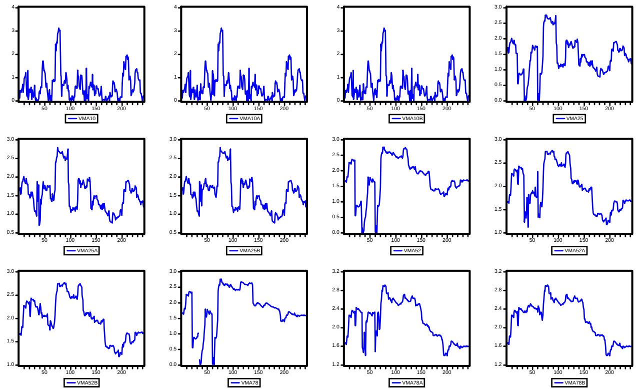
15 Beginning of 2000 and 2002-2003.

16 Short moving averages collect short-term volatility better but face tendency changes more often; long moving averages collect long-term tendencies and only have problems when facing long-term changes.

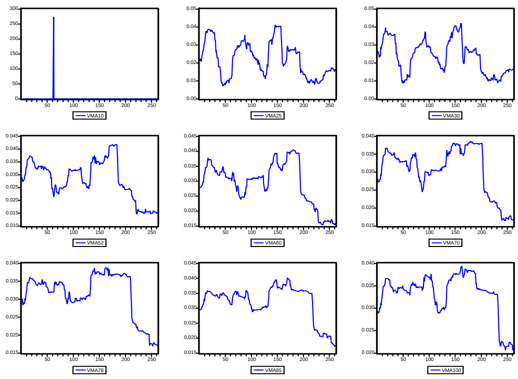
Figure 2

VAR

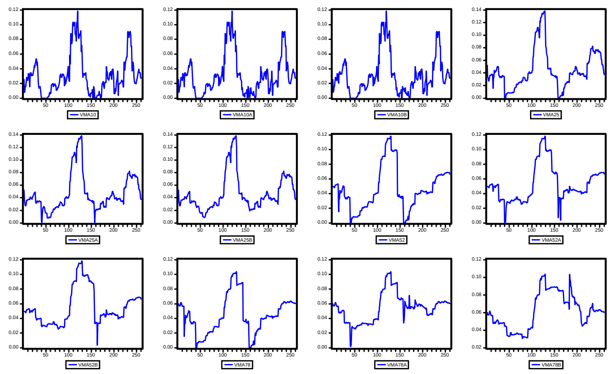
IVaR35



IVaR30



IvaRM

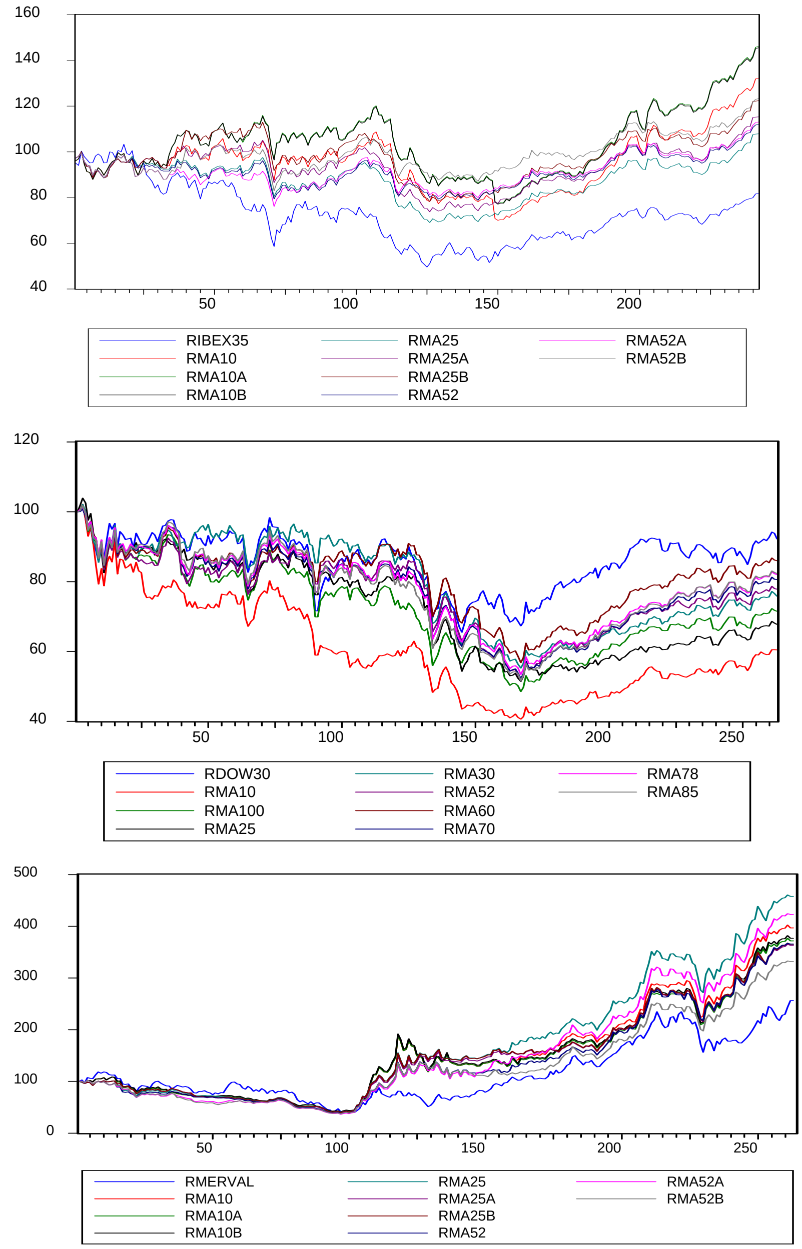


Note: The figures in IVaR35, IvaR30 and IvaRM, (left to right and top to bottom) are the VaR evolution for each approximation in each market.

3.4. RETURN ANALYSIS

Figure 3 shows all the returns of Moving Average approximations. In the Spanish market, all our Indices have higher returns than the IBEX 35®. These results are surprising but not unique. Other authors have constructed portfolios able to beat the market [e.g. Alexander and Dimitriu (2004, 2005) and Carosa (2005)]. There are two reasons for these data. First, the Spanish Stock Market is suffering efficiency bias and Minimum Risk Indices are harvesting part of it. Second, the way the IBEX35® is constructed generates sample and construction bias, even they should be relatively small comparing with these biases in other markets [Andreu (2008)].

Figure 3
Evolution of the indices (IVaR35, IVaR30, IVaRM)



Note: 100 based. In the first, second and third graph RIBEX35, RDOW30 and RMERVAL are the evolution of IBEX35®, DJIA[SM] and MERVAL, and RMA are the evolutions of each moving average approximation in each market.

In contrast, in the American Stock Market, no Minimum Risk Index beats the market. In the IVaR30, the approximations with the worst returns are MA10. The other approximations performed quite well during the bearish market, being near the actual index or beating it in some periods, but they performed worse than the current index during and after the Iraq war in the bullish market. In the end, the best approximation in terms of profitability is the MA60, with a 20% lower return than the Dow Jones Industrial Average[SM]. The same two reasons could be put forward in this case. The efficiency bias in the American market is clearly lower than in the Argentinian or Spanish case, but the sample and construction biases in the Dow Jones are theoretically important. How these biases seem to compensate among them avoiding the outperformance of our indices is an interesting idea to be analysed in the future. Finally, in the Argentinian Market, all approximations (except 78b) are able to beat the market. The same reasons put forward for the IBEX 35® are valid here. The efficiency biases seem to be especially important knowing a little about the Argentinian market. The construction and sample biases are very important taking into account the sample is not totally representative of the market and the index is constructed using a negotiation weighting.

It is necessary to take into account some considerations. Firstly, it is essential to discover how efficiency affects these conclusions. This would mean calculating efficiency tests for each market and comparing results, but I must leave this for further research. Secondly, I must deeply analyse how other index biases affect risk and return. Finally, it is essential to analyse how results could improve using more powerful techniques to estimate variances and covariances.

3.5. IVaR35, IVaR30 AND IVaRM COMPOSITION

Each Minimum Risk Index has a different optimal composition. With short-term Moving Averages, the optimal composition changes frequently over the weeks, whereas it is more stable with long-term Moving Averages. This is important if we bear in mind that my index approximation needs a weekly adjustment, which means higher turnover and transaction costs with shorter moving averages. As optimal compositions are calculated using a risk measure and the Covariance matrix estimation, they are not like current compositions of Market Indices.

3.6. NORMALITY ANALYSIS AND BACKTESTING

Normality analysis of logarithmic returns is not very positive, as it can be seen in Tables 3, 4 and 5. In all cases, the Normality Hypothesis has been rejected except in the case of Merval. In the case of IVaR35, the distributions have leptokurtosis and are slightly negatively skewed. In the case of IVaR30, the distributions are more leptokurtical, and negative skewness is especially important, affecting, as I have said, the probitability results. Finally, in the case of IVaRM, distributions are extremelly leptokutical and skewness is positive because of the evolution of Merval during the period analysed.

Table 3
Backtesting process in the IVaR35

	Normality		Backtesting			
	Jarque-Bera	Probability	Mean VaR(%)	Errors	% Errors	Mean Error(%)
IBEX35®	12.22	0.002				
MA10	102.4	0.000	0.654	75	31	1.95
MA10a	101.4	0.000	0.671	75	31	1.89
MA10b	103.3	0.000	0.665	75	31	1.89
MA25	32.9	0.000	1.446	43	18	1.35
MA25a	40.86	0.000	1.557	37	15	1.35
MA25b	48.3	0.000	1.582	36	15	1.37
MA52	43.3	0.000	1.746	33	13	1.41
MA52a	65.8	0.000	1.943	28	11	1.34
MA52b	56.9	0.000	1.996	24	10	1.52
MA78	45.6	0.000	1.843	33	13	1.41
MA78a	116.2	0.000	2.136	25	10	1.46
MA78b	124.95	0.000	2.192	22	9	1.63

Note: VaR value calculated at 5% significance level using data available from 242 weeks.
Errors: losses worse than the VaR value. % Errors: ‘real’ significance level. Mean Error (%) shows the mean loss exceeding the VaR value.

Table 4
Backtesting process in the IVaR30

	Normality		Backtesting			
	Jarque-Bera	Probability	Mean VaR(%)	Errors	% Errors	Mean Error(%)
DJIA[SM]	271.4	0.000				
MA10	85.85	0.000	1.208	67	25.5	2.23
MA25	62.09	0.000	2.221	38	14.5	2.03
MA30	49.04	0.000	2.365	33	12.6	2.05
MA52	124.2	0.000	2.825	22	8.3	2.78
MA60	75.17	0.000	2.944	23	8.7	2.39
MA70	123.32	0.000	3.062	20	7.6	2.80
MA78	117.74	0.000	3.141	19	7.2	2.81
MA85	130.40	0.000	3.205	22	8.3	2.41
MA100	150.14	0.000	3.329	24	9.1	2.45

Note: VaR value calculated at 5% significance level using data available from 262 weeks.
Errors: losses worse than the VaR value. % Errors: ‘real’ significance level. Mean Error (%) shows the mean loss exceeding the VaR value.

Table 5
Backtesting process in the IVaRM

	Normality		Backtesting			
	Jarque-Bera	Probability	Mean VaR(%)	Errors	% Errors	Mean Error(%)
MERVAL	4.28	0.111				
MA10	664.02	0.000	3.08	61	23.4	1.98
MA10a	648.56	0.000	3.10	60	22.9	2.07
MA10b	652.00	0.000	3.12	60	22.9	2.03
MA25	534.82	0.000	4.47	38	14,5	2.02
MA25a	619.22	0.000	4.64	33	12.64	2.04
MA25b	623.62	0.000	4.69	31	11.87	2.02
MA52	292.01	0.000	4.57	39	14.94	1.92
MA52a	243.82	0.000	5.24	29	11.11	1.98
MA52b	264.28	0.000	5.48	26	9.96	1.95
MA78	276.73	0.000	4.38	42	16.09	1.85
MA78a	207.70	0.000	5.59	29	11.11	1.90
MA78b	177.41	0.000	6.15	26	9.96	1.99

Note: VaR value calculated at 5% significance level using data available from 261 weeks.
Errors: losses worse than the VaR value. % Errors: ‘real’ significance level. Mean Error (%) shows the mean loss exceeding the VaR value.

If we look at the backtesting results, though real errors are more frequent than the 5% significance level expected, they are not very large (around 2% higher than the VaR value in the IVaR35, 2.80% higher in the IVaR30, and 2% higher in the IVaRM). Errors are more controlled in terms of frequency in the case of IVaR30 than for the IVaR35 or IVaRM, but they are less controlled in terms of magnitude (the mean error in the American and Argentinian cases is higher than in the Spanish case). Basak and Shapiro (2001) observed that setting VaR limits on institutions could lead to higher extreme losses than when these limits are not set. We can see from results, however, that this theoretical result is not clear here.

3.7. MODEL SELECTION. THE BEST IVaR35, IVaR30 AND IVaRM

We have seen how parametric VaR minimization could create Minimum Risk Indices with less risk and, in the Spanish and Argentinian case, with greater profitability than current market indices. In this paper I construct 12 approximations using Moving Averages of different lengths for the Spanish and Argentinian market and 9 approximations for the American market. It is necessary now to decide which is the ‘best’ approximation to use in each market from all the approximations presented. I think this selection should be done on the basis of two ideas:

(1) The model’s capacity to explain reality or, in other words, its capacity to be accepted by the backtesting process. After determining the number of returns lower than the VaR value (classic backtesting), it is important to also measure the error magnitude. This type of backtesting has not yet been developed and here I only propose a very simple method that deals with error magnitude using the Excess Total Loss (ETL) measure, which is defined as the total sum of all returns lower than the VaR value over the studied period. I will choose those approximations with the lowest ETL in order to take into account the risk ‘out of the model’. It is then necessary to select those approximations with less VaR or with less risk ‘within the model’.

Table 6
Model Selection in the Spanish Stock Market

	Backtesting			Mean VaR (%)	Sharpe's Ratio
	Errors	Mean Error (%)	ETL in 2000-2004		
MA10	75	1.95	146.25	0.654	11.71
MA10a	75	1.89	141.75	0.671	15.83
MA10b	75	1.89	141.75	0.665	15.69
MA25	43	1.35	58.05	1.446	4.18
MA25a	37	1.35	49.95	1.557	7.75
MA25b	36	1.37	49.32	1.582	10.70
MA52	33	1.41	46.53	1.746	6.52
MA52a	28	1.34	37.52	1.943	6.96
MA52b	24	1.52	36.48	1.996	11.56
MA78	33	1.41	46.53	1.843	8.03
MA78a	25	1.46	36.5	2.136	6.25
MA78b	22	1.63	35.86	2.192	10.54

Note: in Sharpe's ratio non-risk return has been considered equal to zero.

Table 7
Model Selection in the American Stock Market

	Backtesting			Mean VaR (%)	Sharpe's Ratio
	Errors	Mean Error(%)	ETL in 2000-2004		
MA10	67	2.23	149.42	1.208	-18.46
MA25	38	2.03	77.14	2.221	-16.58
MA30	33	2.05	67.80	2.365	-11.75
MA52	22	2.78	61.21	2.825	-10.39
MA60	23	2.39	55.03	2.944	-6.34
MA70	20	2.80	56.18	3.062	-8.80
MA78	19	2.81	53.56	3.141	-8.17
MA85	22	2.41	53.14	3.205	-7.94
MA100	24	2.45	59.02	3.329	-13.22

Note: in Sharpe's ratio non-risk return has been considered equal to zero.

As we can see in Table 6, in the Spanish market, using ETL the best approximations are MA52a,b and MA78a,b. Moreover, studying the 'controlled' risk within the model, I conclude that the MA52a,b approximations are the least risky.

In the American Market, (see Table 7), the best approximations using the ETL are MA60, MA70, MA78, MA85 and MA100. Using the 'controlled' risk, I conclude that the best approximations are MA60, MA78 and MA85.

Table 8
Model Selection in the Argentinian Stock Market

	Backtesting			Mean VaR (%)	Sharpe's Ratio
	Errors	Mean Error(%)	ETL in 2000-2004		
MA10	61	1.98	121.1	3.08	27.73
MA10a	60	2.07	124.2	3.10	26.31
MA10b	60	2.03	122.2	3.12	26.62
MA25	38	2.02	77.1	4.47	32.67
MA25a	33	2.04	67.4	4.64	28.08
MA25b	31	2.02	62.7	4.69	28.22
MA52	39	1.92	75.2	4.57	27.03
MA52a	29	1.98	57.7	5.24	29.69
MA52b	26	1.95	50.8	5.48	24.78
MA78	42	1.85	77.8	4.38	27.02
MA78a	29	1.90	55.2	5.59	27.67
MA78b	26	1.99	51.7	6.15	17.92

Note: in Sharpe's ratio non-risk return has been considered equal to zero.

Finally, in the Argentinian market (Table 8), the best approximations by ETL are MA52a,b and MA78a,b, and, after using the 'controlled risk' measured by the VaR, I can conclude that the best approximations are MA52a and MA52b.

(2) The relationship between return and risk, since Dembo and Freeman (1998) criticize not attaching importance to that point in VaR calculations. Here the author uses Sharpe's ratio to analyse this relationship.

In the Spanish market, Sharpe's ratio in the MA52b approximation is bigger than in the MA52a approximation so I can conclude that MA52b is the best approximation with which to construct the Spanish Minimum Risk Index. In the American market, Sharpe's ratio in the MA60 approximation is the lowest of the selected MAs, so MA60 is the best approximation with which to construct the American Minimum Risk Index. Finally, in the Argentinian market, Sharpe's ratio in MA52a is higher than in MA52b, so it is reasonable to conclude that MA52a is the best approximation to construct the Argentinian Minimum Risk Index.

4. CONCLUSIONS

In this article I propose using the VaR as an active risk measure to construct Minimum Risk Market Indices. I have used the parametric VaR approach to construct a very simple minimization problem in which the Covariance matrix among asset returns has to be estimated. Covariance matrix estimation can be done using many methods. There are, therefore, many ways of constructing a Minimum Risk Index—one for each way of estimating the Covariance matrix—so a method of selecting the best model is needed. This selection method must be based on the model's capacity to be accepted by the backtesting process, (taking into account error frequency and error magnitude) and the return-risk relationship.

I apply this method to the Spanish, American and Argentinian markets to create different Minimum Risk Indices (IVaR35, IVaR30 and IVaRM) for the 2000-2004 period. Using the simplest Covariance matrix estimation methods, I achieve interesting results: my indices are less risky than the current ones (half the risk in the Spanish Market). Also, thanks to their optimal portfolio characteristics, the Spanish and Argentinian cases achieved bigger returns than those of the current market Indices, contrary to what is expected from the Efficient Market Hypothesis. These results show that both markets suffer from efficiency biases, and that Minimum Risk Indices could partially solve this. Part of the results, in the Argentinian case, can be due to the existence of an important sample and construction bias created by how the Argentinian index is built. This highlights an interesting discussion that needs to be dealt with care in future research and which must be based on the following ideas: (i) the ability to moving averages to estimate future covariance matrices and the possibility of obtaining better results with more complex estimation methods; (ii) the influence of the weighting process and other construction rules on market indices (the sample and construction bias); (iii) the influence of market index biases in the performance and risk of indexes and how biases are additive or can be compensated among them; (iv) the Minimum Risk Index aproximation in order to prove the efficiency of a market and to solve the efficiency bias; and (v) whether it is possible to obtain better results by not limiting our Minimum Risk Index shares to the current Market Index components and to the particular and 'legal' timing of changes in components.

The potential uses of Minimum Risk Indices are clear. Firstly, they are less risky and in some cases more profitable than current ones, which makes them a suitable benchmark of risky assets for mutual funds that currently follow market indices or a suitable base for derivatives. Secondly, Minimum Risk Indices may generate more stable Betas in the CAPM model, which is a possibility that must be developed in the future.

5. FUTURE LINES OF RESEARCH

The results achieved by very simple methods in the examples presented are interesting but it must also be said that there is still a lot to do. First it is necessary to determine whether better Covariance estimations using EWMA or GARCH or HAC methods [Newey and West (1987, 1994)] can achieve better results in terms of risk and profitability. Perasan and Zaffaroni (2006) indicate EWMA could be even more powerful than more complex models as O-GARCH or Factor GARCH. Taking into account the complexity of applying complicated models in an empirical analysis I would like to start with EWMA in the near future. I also need to determine how the sample and construction biases affect eficiency tests, market index performance and the possibility of beating it. Finally, methods for selecting the 'best' model must be further developed since here I have only provided some general guidelines.

I would like to thank Maxim Borrell and Daniel Liviano for their interesting comments and help.

6. APPENDIX

6.1. MINIMUM VaR INDICES.

The problem to solve can be written as follows:

$$\begin{aligned} & Min\left\{ \; Z_\alpha \sqrt{x'\Sigma x} \; \right\} \\ & s.a. \\ & \qquad x_i \geq 0; \; i = 1..N \\ & \qquad \sum_{i=1}^{N} x_i = 1 \end{aligned} \tag{1}$$

Where Z_α is the Normal distribution value at the desired significance level. The x vector contains the weights of each share within the alternative Market Index I am trying to build, N is the number of shares that make up the Index, and Σ is the logarithmic return Covariance matrix assumed to be a Multivariate Normal. This problem is easy to simplify [Peña (2003)]:

$$\begin{aligned} & Min\, x'\Sigma x \\ & s.a. \\ & x'\vec{1} = 1 \end{aligned} \tag{2}$$

Where $\vec{1}$ is a Nx1 vector all of whose values are equal to one.

Using a Lagrangian optimization (3) in which the optimal weights are the objective of our study and λ is a positive constant:

$$MínL = x'\Sigma x + \lambda(1 - x'\vec{1}) \qquad (3)$$

As the Covariance matrix is positive and semi-defined, if the number of observations is bigger than the number of assets, first order conditions are enough for a minimum[17].

$$\frac{\partial L}{\partial x} = \Sigma x - \lambda\vec{1} = \vec{0} \qquad (4)$$

$$\frac{\partial L}{\partial \lambda} = 1 - x'\vec{1} = 0 \qquad (5)$$

To solve the solution:

$$\begin{aligned} x^* &= \lambda(\Sigma^{-1}\vec{1}) \\ \lambda &= (\vec{1}'\Sigma^{-1}\vec{1})^{-1} \end{aligned} \qquad (6)$$

The solution gives the optimal weight each share must have within the index to minimize market risk.

In the empirical case for the Spanish, American and Argentinian Stock Markets, I use an iterative algorithm based on Newton's method to make the minimization process and the command of the quadratic programming problem in Gauss with similar results.

17 Second-order conditions are also necessary in the other case.

6.2. THE NORMALITY HYPOTHESIS IN THE PARAMETRIC VaR AND THE 'BEST' MODEL SELECTION PROCESS

The return distributions of the indices are similar to those of a Normal distribution by the Central Limit Theorem, weekly data and VaR at a 5% significance level. However, when Normality is rejected, this affects extreme losses and the 'best' model selection process.

If return distributions are Normal, error risk is equal to the significance level. In the best model selection, it should be enough to choose the approximation with the biggest mean and the lowest volatility or the one with an interesting combination of mean and volatility to allow this approximation to suffer fewer losses.

If return distributions are not normal, extreme losses will not be controlled by the model. In this case, the main problem is negative skewness because leptokurtosis makes small gains and losses more likely. The importance of negative skewness should be determined and backtesting of the model should be performed in order to detect extreme losses that are not controlled by the model. In the 'best' model selection, error frequency, error magnitude and the return-risk relationship must be taken into account.

6.3. MARKET INDICES AND THEIR MINIMUM RISK INDICES

In this section I will briefly explain how the Spanish IBEX35®, the American Dow Jones Industrial Average℠ and the Argentinian MERVAL are built. It is also important to explain certain characteristics and problems I found and solved by creating the Minimum Risk Index for each Market.

The Spanish IBEX35®:

The IBEX35® is built using the 35 largest companies in the Spanish Stock Market in terms of market capitalization and liquidity. Every six months the components of the Index are checked, some shares are included or excluded but the total number of assets is maintained. The Index is calculated using a market capitalization weighting criterion.

The Minimum Risk Indices I created for this market were named IvaR35, and comprise the 35 shares of the IBEX 35® at each moment with the optimal weight established by the VaR minimization process. I must point out one problem with the IBEX 35® Spanish Market Index. In the six-month revision of the composition of the IBEX 35®, it is normal to include shares and companies with very little history on the Stock Exchange because it is relatively easy to be both new and one of the biggest 35 companies in the Spanish Market. During the period of our analysis I sometimes encountered this problem—especially in 1999-2000 because of the Internet and .com companies that grew quickly at that time. This makes it difficult to obtain complete data for all the IBEX 35® components in some periods and has important consequences in Covariance matrix estimation. After April 2000 I solved this problem with the following techniques:

a) *Covariance matrix estimation using a multiple-step method*: when I did not have complete data on the 35 shares, Covariance matrix estimation was done using a multiple-step method, estimating each individual value in the covariance matrix with all the available data.

b) *0.01% Weighting*: the above solution improved the results, but shares with short historical data tended to underestimate risk and therefore received high weights because of their 'artificial' low risk. With this approximation I forced these shares to have the minimum weight accepted for our study.

The approximations I finally developed are shown in Table 9:

Table 9
Approximations used for Covariance matrix estimation in the IVaR35

Approximation	Method	Length	Improvements Applied.
MA10	M.A.	10 weeks	None
MA10a	M.A.	10 weeks	Multiple-step Method
MA10b	M.A.	10 weeks	Multiple-step Method and 0.01% Weighting
MA25	M.A.	25 weeks	None
MA25a	M.A.	25 weeks	Multiple-step Method
MA25b	M.A.	25 weeks	Multiple-step Method and 0.01% Weighting
MA52	M.A.	52 weeks	None
MA52a	M.A.	52 weeks	Multiple-step Method
MA52b	M.A.	52 weeks	Multiple-step Method and 0.01% Weighting
MA78	M.A.	78 weeks	None
MA78a	M.A.	78 weeks	Multiple-step Method
MA78b	M.A.	78 weeks	Multiple-step Method and 0.01% Weighting

Note: M.A.: Moving Average.

The American Dow Jones Industrial Average^{SM}:

The DJIA^{SM} is built using the 30 biggest companies in the American Stock Market and, for the sake of continuity, composition changes are rare. Inclusions and exclusions of shares are therefore rare and basically related to corporate acquisitions or dramatic business events. The Index is calculated using a price-weighting criterion.

Using available data for the DJIA^{SM} I did not need to apply improvements to the Covariance matrix estimation. The good quality of these data means that I used the methodology with a greater number of moving average lengths. The Minimum Risk Index I created to this market was named IVaR30.

The approximations I developed are shown in Table 10.

Table 10
Approximations used for Covariance matrix estimation in the IVaR30

Approximation	Method	Length	Improvements Applied
MA10	M.A	10 weeks	None
MA25	M.A	25 weeks	None
MA30	M.A	30 weeks	None
MA52	M.A	52 weeks	None
MA60	M.A	60 weeks	None
MA70	M.A	70 weeks	None
MA78	M.A	78 weeks	None
MA85	M.A	85 weeks	None
MA100	M.A	100 weeks	None

Note: M.A.: Moving Average.

The Argentinian MERVAL®:

The MERVAL is built using the most traded companies in the Argentinian Stock Market. The weights of each share in the index are calculated using the number of transactions of these shares in the Stock Market and the Volume of these transactions, so the Index is calculated using a negotiation weighting criterion.

The Minimum Risk Indices I created for this market were named IVaRM and comprise the shares of the Merval at each moment with the optimal weight established by the VaR minimization process. Every three months, the Merval composition is changed, and it is possible, as in the IBEX 35®, to find companies with very little historical data. Calculations must then be improved using the same techniques as for the the IBEX 35®. Table 11 shows the approximations I used in this paper. It is important to point out that approximation b is especially influenced in the Merval by the fact that there are a lot of stocks with a short or no history when they enter the Index, this affects the performance and backtesting of the approximation.

Table 11
Approximations used for Covariance matrix estimation in the IVaRM

Approximation	Method	Length	Improvements Applied.
MA10	M.A	10 weeks	None
MA10a	M.A	10 weeks	Multiple-step Method
MA10b	M.A	10 weeks	Multiple-step Method and 0.01% Weighting
MA25	M.A	25 weeks	None
MA25a	M.A	25 weeks	Multiple-step Method
MA25b	M.A	25 weeks	Multiple-step Method and 0.01% Weighting
MA52	M.A	52 weeks	None
MA52a	M.A	52 weeks	Multiple-step Method
MA52b	M.A	52 weeks	Multiple-step Method and 0.01% Weighting
MA78	M.A	78 weeks	None
MA78a	M.A	78 weeks	Multiple-step Method
MA78b	M.A	78 weeks	Multiple-step Method and 0.01% Weighting

Note: M.A.: Moving Average.

7. REFERENCES

Abken, P. A. (2000); "An empirical evaluation of value at risk by scenario simulation". *Journal of Derivatives*, pp.12-30

Ackerman, C.; McEnally, R.; Ravenscraft, D. (1999): "The Performance of Hedge Funds: Risk, Return and Incentives", *Journal of Finance,* v.54, iss.3, pp.833-874.

Alexander, C. (1996): "Volatility and Correlation Forecasting", *Handbook of Risk Management and Analysis*, John Wiley & Sons.

Alexander, J.; Baptista, M. (2003): "Portfolio Performance Evaluation Using Value at Risk: The Reward-to-VaR Ratio", *Journal of Portfolio Management*, v. 29, iss. 4, pp. 93-102.

Alexander, C.; Dimitriu, A. (2004): "Indexing and Statistical Arbitrage", *The Journal of Portfolio Management*, v.31, iss.2, pp.50-63.

Alexander, C.; Dimitriu, A. (2005): "Indexing, cointegration and equity market regimes", *International Journal of Finance and Economics,* v.10, pp.213-231.

Andreu, J. (2008): "Some traditional financial ideas revisited: the market, efficiency and indexing under new financial approaches", *Working paper.*

Aragall, E. (2002): “Valor en risc. Mètodes de càlcul. Crítiques i alternatives: Introducció a la Teoria del Valor Extreme”. *Estudis sobre el mercat de Valors*, n.30, Servei d'Estudis, Borsa de Barcelona.

Artzner, P.; Delbaen, F.; Eber, J.M.; Heath, D. (1999): “Coherent Measures of Risk”. *Mathematical Finance*, v.9, n.3, pp.203-228, July.

Basak, S.; Shapiro, A. (2001): “Value-at-risk based risk management: optimal policies and asset prices”. *The Review of Financial Studies*, v.14 (2), pp. 371-405.

Blanco, C.; Oks, M. (2004): “Backtesting VaR models: Quantitative and Qualitative Tests”. *Risk Desk*, v. IV, n. I.

Bollerslev, T. (1986): “Generalized Autoregressive Conditional Heteroskedasticity*”. Journal of Econometrics*, v. 31, pp.307-327.

Boudoukh, J.; Richardson, M.; Whitelaw, R. (1998): “The Best of Both Worlds”. *Risk*, v.11 (5), pp.64-66, May.

Carosa, C. (2005): “Passive investing: The Emperor Exposed?”, *Journal of Financial Planning*, v.18. iss.10, pp.54-62.

Danielsson, J.; De Vries, C. G.; y Jorgensen, B. N. (1998): “The value of Value at Risk: Statistical, Financial, and Regulatory Considerations”. *FRBNY Economic Policy Review*, Oct.

Dembo, R.; Freeman, A. (1998): *Seeing Tomorrow: rewriting the Rules of Risk*. Wiley.

England, C. (1997): *Identify Critical Gap Between Risk Management Objectives and Practice*. Barra y RogersCasey. PR Newswire.

Engle, R. F. (1982): “Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of the United Kingdom Inflation”. *Econometrica*, v.50, n.4, pp. 987-1008, Jul.

Fama, E. (1970): “Efficient capital markets: A review of theory and empirical work”. *Journal of finance*, v.25, pp.383-417.

Froot, K.; Scharfstein, D.; Stein, J. (1994): “A Framework for Risk Management”. *Harvard Business Review*, v.72, pp. 59-71.

Frye, J. (1998): “Monte Carlo by day”, *Risk*, v.11 (11), pp.66-71.

Garman, M. (1996): “Improving on VaR”, *Risk,* v.9, iss.5, pp.61-63.

Garman, M. (1997a): “Ending the Search for Component VaR”, *Financial Engineering Associates, Inc*.

Garman, M. (1997b): *Value at Risk: Evolving Topics*, (Conferencia) Organizada por Financial Engineering Associates, Inc, Berkeley, 3-4 Nov.

Garman, M. B. (1997c): “Taking VaR to pieces”. *Risk*, v.10, iss.10, pp.70-71. Oct.

Gibson, M. S.; Pritsker, M. (2000): “Improving grid-based methods for estimating value at risk fixed-income portfolios”. *Federal Reserve Board Finance and Economics Discussion Series*, 2000-25.

Group of Thirty (1993): *Derivatives: Practices and Principles*. Global Derivatives Study Group. Washington, D. C.

Hendricks, D. (1996): “Evaluation of VAR models using historical data”. *Federal Reserve Bank of New York Economic Policy Review*, v.2, iss.1, pp.39-69.

Hopper, G. P. (1996): “Value at risk: a new methodology for measuring portfolio risk”. *Federal Reserve Bank of Philadelphia Business Review,* v.0, iss.0, pp.19-39.

Hull, J. C.; White, A. W. (1998): “Incorporating volatility up-dating into the Historical Simulation method for value at risk.” *Journal of Risk,* v.1 (1), pp.5-19, Fall.

Johnson, C. A. (2001): “Value at Risk: Teoria y aplicaciones”. *Estudios de Economía*a, v.28, n.2, pp.217-247. Dec.

Jorion, P. (2001): *Value at risk*. New York: McGraw-Hill.

Larsen, N.; Mausser, H.; Uryasev, S. (2002): “Algorithms for Optimization of Value-at-Risk”. *Financial Engineering, e-Commerce and SupplyChain*. Kluwer Academic Publishers, pp.129-157.

López, J. (1999): “Methods for evaluating Value at Risk estimates”. *Federal Reserva Bank of San Francisco Economic Review*, v.2, pp. 3-17.

Markowitz, H. M. (1952): “Portfolio selection”. *Journal of Finance*, 25(1), pp. 71-79.

Merton, R. (1973): “An Intemporal Capital Asset Pricing Model”. *Econometrica*, 41, pp. 767-887.

Nawrocki, D. N. (1999): “A Brief History of Downside Risk Measures”. *Journal of Investing*, v.8, n.3, pp.9-26, Fall.

Neftci, S. N. (2000): “Value at Risk calculations, extreme events, and tail estimation”. *Journal of Derivatives*, v.7, iss.3, pp.23-37.

Nelson, D. B. (1991): “Conditional heteroskedasticity in asset returns: A new approach”. *Econometrica*, n.59, pp.347-370.

Newey, W.K.; West, K.D. (1987): “A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix.” *Econometrica*, v.55, pp.703–708.

Newey, W.K.; West, K.D. (1994): "Automatic Lag Selection in Covariance Matrix Estimation". *Review of Economic Studies*, v.61, pp.631–653

Pearson, N. D.; Smithson, C. (2002): "VaR, the state of play". *Review of Financial Economics*, v.11, pp.175-189.

Pedersen, C. S.; Satchell, S. E. (1998): "An Extended Family of Financial-Risk Measures". *Geneva Papers on Risk and Insurance Theory*, v.23, iss.2, pp. 89-117.

Peña, J. I. (2002): *La gestión de riesgos financieros*. Madrid: Financial Times-Prentice Hall, Pearson Educación.

Perasan, M.H.; Zaffaroni, P. (2006): "Model Averaging and Value-at-Risk based Evaluation of Large Multi Asset Volatility Models for Risk Management", *Working paper.*

Pritsker, M. (1997): "Evaluating value at risk methodologies: accuracy versus computational time". *Journal of Financial Services Research*, v.12 (2/3), pp. 201-242.

Riskmetrics (1995): *Technical Document*. New York: Morgan Guarantee Trust Company, Global Research.

Robinson, G. (1996): "More Haste, Less Precision". *VaR: Understanding and applying Value-at-risk.* KPMG.

Rockafellar, R. T. ; Uryasev, S. (2000): "Optimization of Conditional Value at Risk". *The Journal of Risk*, v.2, n.3, pp.21-41.

Sarma, M.; Thomas, S.; Shah, A. (2003): "Selection of Value-at-Risk Models". *Journal of Forecasting*, v.22, iss.4, pp.337-358.

Schoenfeld, S. A. (2004): *Active Index Investing*. Hooboken, New Jersey: Wiley.

Sentana, E. (2001): "Mean-Variance Portfolio Allocation with a Value at Risk Constraint". *Working paper*, n.105, CEMFI.

Sharpe, W. G. (1964): "Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk". *Journal of Finance*, 19, pp. 425-442.

Simons, K. (1996): "VaR-New approaches to Risk Management". *New England Economic Review*, v. 0, iss. 0, pp. 3-13, Sept.-Oct.

Stone, B. K. (1973): "A General Class of Three-Parameter Risk Measures". *The Journal of Finance*, v.28, iss.3, pp.675-685.

CONCLUSIONS, CONTRIBUTIONS AND FUTURE RESEARCH

It is always complicated to sum up conclusions in some lines. The task is even more difficult if conclusions derived from the thesis are so different and heterogeneous. My main objective was to contribute to the literature with a complete theoretical and empirical analysis of market indexes. An exhaustive theoretical study is provided in the first and second article, and an empirical application to deal with the efficiency bias is developed in the third one. Unfortunately, a complete empirical analysis regarding statistical characteristics of market indexes or bias measurement is still not complete, and was not included at the end.

Perhaps one way to understand how this PhD thesis contributes to the literature is to know what questions it does not answer. Limitations of my study are various. They are commented in each chapter, but I decided to sum up the most important:

1. Regarding the first article, my approximation is theoretical and qualitative. A quantitative analysis should be developed in the future to measure and analyse market biases, completing in that way the work presented here. On the other hand, the ecological approximation is also totally qualitative. It is necessary to apply mathematical ecological models or quantitative analysis to prove, use and exploit all the ideas presented in this part.

2. Regarding the second article, one of the objectives was to sum up complete information about market indexes, information that is not available anywhere. Historical background, market index's foundations, characteristics and biases are provided, but trying to cover that objective I left quantitative analysis again. Measures, solutions and empirical analysis of market index's biases should be developed in the future.

3. Regarding the third article, limitations are also clear. Minimum Risk Indices (MRI) are presented as alternative market indices that solve partially the efficiency bias using a minimization risk approach. The risk measure is easily improvable (CVaR or Downside Risk are better measures if normality is not fulfilled) and also the covariance matrix estimation (using EWMA or GARCH). Another important limitation is robustness. Even results seem hopeful the Minimum Risk Index approach should be used in out of sample periods and in other markets (only the 2000-2004 period and three markets were used in calculations). Finally, the selection of the best index must be improved, defining clearer and structured processes.

Once important limitations of the PhD thesis have been presented, it is time to pay attention to conclusions of this work. Conclusions, contributions and future research are presented in connection with the chapter that generated it but relating them with the rest of the thesis and with my future research.

Conclusions in the first article, 'Some Traditional Financial Ideas Revisited: the Market, Efficiency and Indexing under new Financial Approaches' are commented in the next lines. Traditional financial theory concluded an investor could not consistently beat the market thus meaning that a passive investment strategy was optimal. The zero-sum game idea, CAPM conclusions and the EMH rule out the possibility of trading systems with expected returns in excess of what can be obtained following the market. According to traditional finance, the optimal strategy for an investor was to get the market return, to become passive. A deep analysis of these three bases brings us to different conclusions. Market is a minus-sum game, efficiency failures are more common that expected and the CAPM hypotheses are not fulfilled in the real world. These conclusions are not trivial. In this new scenario, the market portfolio is not optimal and not located in the efficient frontier, so passive investment strategies are suboptimal and biased.

One of the innovations of the paper is that it classifies market index biases in five groups (sample bias, construction bias, tracking error, efficiency bias, and active bias). The second innovation here is to analyse a Financial Market as an evolving ecosystem. Taking into account ideas from ecology, I developed a market structure along with its rules of evolution inspired by other pioneering studies. An alternative organism classification is presented and market participants are classified in organizers, producers, providers and predators. Finally, the article discusses species' relationships (competence, predation and mutualism) as determinants of the financial market's status and evolution and 13 rules guiding the ecosystem evolution are provided. In this new framework, not full rational individuals act in their own self-interest, making mistakes, learning, adapting and changing their investment strategies. Every investor is unique, holds a different portfolio, and learns and adapts with a personal non-constant survival probability. Organisms interact, guided by ecological factors, and interaction drives to adaptation, innovation and survival. Finally, surviving individuals transmit their knowledge to their offspring, meanwhile non surviving individuals disappear.

The first article generates hundreds of lines of research. First, a complete empirical analysis of market indices is not available in the literature. A study of Normality, Skewness, Kurtosis, efficiency or co integration with a complete and unique data base of indexes is one of our first objectives for the future. Second, the ecological approach is a huge source of investigation. Biologically Inspired Algorithms are perfect tools for modelling financial markets. Intraspecies and interspecies ecological models, territorial and temporal approaches, predation models or epidemiology studies should help academics in their study of financial markets. An enormous body of ecological literature is available to match ecology and finance. Analysis of survival probability using ecological factors or personal characteristics can be also interesting lines of research. Third, Behavioural Finance becomes a key issue in a market where individuals fail, learn and adapt. Finally, models covering and explaining the dynamics of the ecosystem are required and must be developed using the 13 rules presented in the paper.

Regarding some of the conclusions provided by the first article, the second one, 'Market Indexes in detail. Biases revealed'' is developed. The first contribution of the article is it supplies a complete historical index analysis and definitions of index's objectives and desirable characteristics. From a theoretical point of view, an index must be a representative complete investable capitalization-weighted portfolio that measures total market performance. The second contribution is a detailed study of market index biases, sources and strategies to solve or minimize them:

i) *The sample bias* is a statistical bias between universe and sample with a clear trade off. If the sample bias is reduced building a complete index, tracking error increases, and vice versa. This bias is quite big in Hedge Fund, Fixed Income and Real Estate Indices because these assets have a huge and heterogeneous universe that is difficult to sample. Float adjustment is absolutely essential, but not all indexes have it, and clear rules for its application do not exist. The selection bias is generated by subjectivity in rules of governance.

ii) *The construction bias* is generated by the application of different weighting criterions and construction methodologies instead of a Laspeyres capitalization index. Most commercial indexes are affected by this bias because index vendors use different weighting criterions and construction rules to build their indexes. In the construction process, it is also important how special events are managed. How dividends, mergers, spin-offs or splits are treated is not trivial. In some cases, the construction bias is so big, that indexes can be seen as hidden active strategies. The construction bias is important with Hedge Funds, Fixed Income and Real Estate. Almost all Hedge Fund Indices are equal-weighted, each Fixed Income index provider uses different construction rules and Real Estate Indices are built using median price indexes, EQR or WRS models.

iii) *Tracking error* is the difference between a portfolio and the index it tries to proxy. The article pays attention to sources (exogenous vs endogenous), replication strategies (Full replication, Sampling strategies, Derivatives based strategies and Naïve Replication) and common tracking error measures used in the literature. It is important to conclude some ideas about tracking error: (i) it is impossible to eliminate; (ii) different sources have different impact on tracking error; (iii) it is clearly seasonal; (iv) it can be used to discover hidden active strategies or active bets taken by managers using the Information Ratio; (v) alternative strategies as Internal Crossing, Tax-Efficient Strategies or Security Lending can be used to generate returns to compensate tracking error deviations.

iv) *The efficiency bias* is the gap between the market portfolio's return and efficient frontier's return for the assumed risk. A market portfolio is almost never optimal, so, even an index has no sample or construction bias, and even tracking error is zero, an indexer suffers from efficiency bias. According to that conclusion, it seems rare that not many authors have analysed it directly. Conclusions about anomalies, the CAPM utility, calculation of betas or the validity of active investments should be revisited taking into account the efficiency bias. This bias can be reduced increasing market efficiency, or achieving more efficient portfolios using different investment strategies. Benefit maximization or risk minimization strategies (with traditional and new risk measures), fundamental indices or pseudo-active strategies should be developed to build more efficient portfolios and indexes. The efficient bias from an empirical point of view opens a wide field of research. How this bias is detected, measured and reduced, how more mean-variance efficient indexes are created and replicated are extraordinary questions to answer.

v) *The active bias* is the difference between the market return and the best investment opportunity (opportunity cost). Active management has been widely analysed in finance but, contrary to what is defended by traditional finance, it has a place in the financial arena. Markets' inefficiencies and index's biases make active strategies theoretically useful. Empirically it is quite different. The utility of active investments and price predictability is not clear in the literature. In my opinion, contradicting results are generated by the fact market index biases have never been taken into account. Empirical studies should rethink the utility of active strategies paying attention to biases presented in this thesis. With this adjustment, I am quite sure conclusions should be slightly different. More over, active strategies should be indexed, becoming benchmarks, to measure the active bias and to evaluate managers with correct tools.

The detection of market index's biases show interesting fields to explore. How construction rules, sampling strategies, index committee decisions or other sources of market index biases affects the performance and risk of a market index have not been analysed in detail. Because market indices are used as proxies of the market, if the proxy is biased, the market is not well measured, and conclusions about the usefulness of active or passive investments, the CAPM or out performance or underperformance of the market must be redefined. Mathematical measures to compute sample, construction, efficiency and active biases must be developed. Tracking error measures should include risk in the future, and efficiency tests and mean-variance efficient portfolio analysis must be developed and compared to market index biases' presence or absence. Finally, a distribution analysis of market index biases can also be very interesting.

Finally, the third article of the thesis 'Refunding Market Indexes using VaR' answers two main questions presented in the second article: if market indexes suffer from the efficiency bias, and if it can be solved using any strategy. Starting from the point most market indices are not in the efficient frontier and suffer from index biases, I try to develop more efficient indexes using a very simple strategy. Minimum Risk Indices are presented as alternative benchmarks. The process to create Minimum Risk Indices is deliberately easy. First of all, the selected risk measure is parametric VaR, a simple but reliable measure in our controlled scenario (non-linear positions, weekly data and 5% significance level), easy to be followed by all investors. Second, the Covariance Matrix estimation is developed using moving average methods, one of the most undemanding methodologies. Why moving averages are selected is clear. Results can be easily improved using more powerful methodologies. Third, once the covariance matrix is estimated, a minimization problem is solved to obtain the optimal composition of Minimum Risk Indexes. Four, it is possible to reconstruct index's performance using historical data, and this performance is compared with current Market Indices. Five, a back testing process is used to determine how models fit real markets. Finally, different Minimum Risk Indices are provided for each market and the 'best' index is selected using back testing and risk-adjusted performance.

Using the theoretical framework presented above, Minimum Risk Indices are generated for the Spanish, American and Argentinean Stock Markets. The aim is to create Minimum Risk Indices based on the historical composition of the IBEX35®, the Dow Jones Industrial AverageSM and the MERVAL for the 2000-2004 period. The Covariance matrix is estimated by the Moving Average Method using different lengths (4 to 100 weeks). Conclusions of the empirical application can be summed up in six: (1) Covariance matrices must be estimated using medium-length Moving Averages to provide stability to VaR. (2) Minimum Risk Indices show less risk than current indexes. The reduction of volatility is especially important in the Spanish case. (3) Extreme losses are similar to those of IBEX 35® and lower in the American and Argentinean market. (4) Our Minimum Risk Indices outperform Spanish and Argentinean indices, revealing that they suffer from the efficiency bias and from the sample and construction bias in the Argentinean case. In contrast, no Minimum Risk Index outperforms the American market even the construction and the sample bias is also theoretically present in the Dow Jones Index. The underlying idea that should be developed in the future is if the American market shows such efficiency that it solves in part market index biases as the construction and sample biases. (5) Normality is rejected in the Spanish and American markets. In the back testing process, real errors are more frequent than the significance level, but they are only 3% higher than the VaR value. (6) The selection of the 'best' approximation is done using: (a) a back testing process that included frequency and magnitude of errors. The innovation is to define the Excess Total Loss (ETL) as measure of error magnitude. (b) Minimum Risk Index performance using Sharpe's ratio. The paper concludes that the MA52b, the MA60 and the MA52a are the best approximations in the Spanish, American, and Argentinean markets respectively.

In this third article, a very simple methodology to build more efficient indexes was presented (Minimum Risk Indices). The construction methodology is deliberately simple because results can be easily improved using more powerful risk measures or estimation procedures. Approaches using CVaR, Downside risk, EWMA, GARCH or HAC estimations should be used to develop more efficient indices. On the other hand, results must be checked for robustness, applying the commented methodology in different markets and time periods. More over, Minimum Risk Indices should be analysed as a possible base for derivatives, as alternative benchmarks, or to calculate more stable betas.

It is possible to conclude, after some years of dedication to Market Indices, my main objective has been achieved, but future lines of research are overwhelming. When I started the PhD, my intention was to clarify some questions about indexing. Coming to the end of this trip, I may say, not only with hope but also with fear, that doors opened in every article and chapter, and there are more open doors now than at the beginning.