High-frequency trading and the efficient market hypothesis

Gianluca Virgilio
Social Sciences, Arts and Humanities Research Institute
University of Hertfordshire, United Kingdom

Keywords
HFT, EMH, efficiency, abnormal, consistent, profit

Abstract
The Efficient Market Hypothesis (EMH) was established by Fama and Samuelson in the 1960s. According to the Hypothesis, prices encompass market information and it is therefore impossible to consistently make abnormal profits, above the ones achievable with a buy-and-hold strategy. Whenever inefficiency in securities pricing arises, arbitraging will immediately sweep it out, re-aligning prices to their fundamental value. The EMH still holds in presence of arbitraging because the abnormal profit would not be consistent, being it shared by the very many market players.

However, a relatively new phenomenon arisen in the last ten years or so has put the EMH under severe questioning - that phenomenon is High-Frequency Trading (HFT).

Although price discrepancies between markets, related instruments or related securities have always been observed and exploited by arbitrageurs, HFT allows a limited number of traders to beat the rest of market operators by exploiting its higher speed, due to superior technology, faster networks and co-location. If the abnormal profits caused by arbitrage are shared by a small number of High-Frequency traders, the assumption that nobody can consistently beat the market does no longer hold and the Efficient Market Hypothesis should therefore be deemed as false.

I have simulated a simple model representing two markets trading the same securities, with a small number of High-Frequency traders and a large number of slower ones, to show how the former traders do make consistent, abnormal, risk-free profit at the expenses of the latter ones. The profit HF traders make is statistically significant.

1. Introduction
The original concept of efficient market dates back to the XIX century, when the French broker Jules Regnault hypothesised the random behaviour of the stock markets, later supported by his fellow countryman Louis Bachelier, at the beginning of the 1900. Recognition of the random walk was further developed into the so-called Efficient Market Hypothesis (EMH or "the Hypothesis") by Eugene Fama in two seminal papers, Fama (1965) and Fama (1970). There exist three forms of the EMH [Fama (1970)]: the Strong form states that the price of securities includes all the past and present information, whether publicly available or otherwise; the Semi-Strong forms claims that prices only reflect all publicly available information, whereas the Weak form further restricts the range of information contributing to price formation to the past.

There is a wide consensus among economists, Grossman and Stiglitz (1980) among others, according to whom the Strong form of the Hypothesis does not hold, if not else because insider trading (that is, trading on private information) is an illegal practice in many countries with financially advanced institutions. The debate, instead, is rather fierce on whether the Semi-Strong and the Weak forms have validity at all. Among the advocates of the non-validity of the Hypothesis, there are the supporters of the Technical Analysis (TA). In the Western World, practitioners have used TA, although to different degrees of depth, for nearly one and a half century, since Charles Dow first published some simple stock indexes on the "Customer's Afternoon Newsletter", although according to Kirkpatrick and Dahlquist (2007) the origin of TA

can, to some extent, be dated as back as the Sixteenth century in Japan. Despite its history, the academic world has often dismissed TA, at best, as some sort of self-fulfilling prophecy. In the following I shall describe various types of arbitraging and the impact ultra-fast transactions have on the theory of EMH. Transaction costs shall also be discussed because of their impact in making markets more efficient. Before drawing a conclusion, I believe it is appropriate to provide some quantitative arguments to support my thesis. In order to do so, I have developed a computer-based simulation of a simplified two-market environment in which a small number of HF traders interact with a much larger number of slow traders face to a significant price difference for the same security in one of the venues with respect to the other, giving rise to an arbitrage opportunity.

2. Arbitrage

It must be said that HF traders make use of several techniques and TA, in one way or another, is certainly one of those [Leshik and Cralle (2011)]. However, as far as arbitrage is concerned, the main focus of interest from the investors' point of view is the classical statement of the Hypothesis. There exist several wording for the Hypothesis but it can be summarised as stating that, assuming participants' rational expectations, in an efficient market all the prices do adjust *instantaneously* to *any new information* [Aldridge (2010)], preventing any market operator to *consistently* make *abnormal* profits. And, as argued by Fama (1970), who makes use of a few theoretical models ("Fair Game", Submartingale, and Random Walk models) as well as of Dow Jones Industrial Average (DJIA) data for the 1957-62 period, should abnormal profits in practice occur, they would disappear in presence of even minimum trading costs. There are a few points I would like to stress about the EMH definition above. First: timing is critical. Then, the type of information also has importance. Lastly, by definition consistent abnormal profit clashes with the very same concept of market efficiency. Let us look at each of those points in some more details.

2.1 Timing

In the context of market arbitraging, timing is critical: if prices do adjust *instantaneously*, then it seems impossible to make abnormal profits in a consistent way from past (weak form) or even from current (semi-strong form) publicly available information. On the other side, if prices adjust slowly, there seems to be plenty of time for a participant, or a group of participants, to exploit publicly available information, to make abnormal profit by trading on such information. It is interesting to notice that the delay prices need to adjust to the newly arrived information is a function of the environment and it comes as no surprise that earlier authors considered compliant with the Hypothesis' requirements delays ["prices of individual securities adjust *very rapidly* to new information", Van Horne (1995), p.52] that more recent studies no longer do. This is a sign of technological advance, the rationale simply being that something *very rapid* in 1995 did not allowed the time to exploit arbitrage opportunities, whereas the same very short delay turned out to be no longer short enough fifteen years later, as in Aldridge (2010). The very concept of *rapidity* has changed its scope a few orders of magnitude over the same period.

2.2 Type of information

Information in the context of arbitraging could be macroeconomic news, like unemployment data released periodically by government agencies, or change of interest rates set by monetary authorities. Other kinds of information could deal with the price of a stock which has broken a support or a resistance, or difference in price at the same time for the same security on different markets, or a temporary misalignment between stock and its related derivatives.

2.3 Consistent and abnormal profits

According to the EMH [Fama (1965)], it is certainly possible to make profits by trading on the market but it is impossible to make *abnormal* profits (i.e. profits above the market average, or average profit running lesser risk) *consistently* without falsifying the Hypothesis. As a matter of principle, no trading strategy should be able to gain more profits than a naïve buy-and-hold strategy, at least in the long term. Fama (1965) takes as an example the net returns of 39 US funds over the 1951-1960 decade and notices that "[t]he most impressive feature […] is the *inconsistency* [italics in the text] in the rankings of year-by-year return for any given fund" (p.92). Therefore, Fama (1965) highlights that "consistently is the crucial word here" (p.40). He also goes as far as to state that should a superior return be made, it would not necessarily be an evidence of superior knowledge, since it could be a "chance result" and "it is only when a fund *consistently* [italics in the text] does better than the market that there is any reason to feel that its higher than average returns may not be the work of lady luck" (p.92). Another dimension of the analysis involves trading strategies, those strategies that, according to the Hypothesis, should never be able of consistently making abnormal profits. Without going into much detail (which would otherwise require several books on its own right), there are three main classes of trading strategies. They are based on Price, News and Arbitrage. Trading strategies based on price are typical of technical analysis, those based on news are closer to fundamental analysis whereas arbitrage exploits misalignment in price and can, on its turn, be subdivided into several different types: a) different markets arbitrage; b) instruments arbitrage; c) related securities arbitrage.

2.4 Price-based strategies

Price-based strategies are somehow related to the broad concept of Technical Analysis. TA assumes, contrary to the EMH, that prices do have "memory", as shown for example in case of mean reversion after an exogenous shock, and therefore they do not follow a random walk [Kirkpatrick and Dahlquist (2007)]. Therefore, future securities prices are, at some extent, predictable. There are numerous signals provided by TA that can be exploited to predict future stock price movements.

2.5 News-based strategies

A news-based strategy is akin to Fundamental Analysis, although the latter tends to have a longer time-span whereas the former exploits not only stock-related information (like periodic earning announcements) but also a wide range of more or less indirectly linked news.

2.6 Markets arbitrage is a strategy conceived well before HFT entered the scene but it does require fast networks to retrieve timely information about prices for the same security from different markets, and as such, it is particularly suitable to HFT speed. If, and when, securities price discrepancy on different markets is noticed, then the standard operations of selling the security in the highly priced market and buying the same security in the other market applies. If both legs of the trading are executed with sufficient low latency to beat on time other investors, then a risk-free profit will be made. The simulation described later is based on this principle.

2.7 Instrument arbitrage

Many securities have closely related siblings in the same or in other markets. Typical examples are stock and the derivatives for the same stock. Up to a certain extent, and taking into considerations all the differences between different instruments, the prices of such securities tend to be correlated. If the spot price of a security drops, under certain conditions it is likely that the price of the related futures follows closely. If the reason for the fall is perceived to have long-term effect, then the futures price must adapt swiftly. Otherwise, if the consensus is for a

temporary effect, the operators being long on such security will hold firmly, simply waiting for the consequences of the effect to pass over. But at that point, the price will have had no reason to drop in the first place, or an arbitrage opportunity would have been created, this time on the stock rather than on the futures.

2.8 Related securities arbitrage is based upon observation and discovery of pairs, or groups, of securities that historically show similar patterns. The rationale behind such similar behaviours is to be deeply investigated to ensure that a rationale does exist, rather than being pure fortuitous coincidence. However, once it is proved as statistically significant, then any divergence of one security with respect to the other(s) is to be acted upon swiftly. Going long in the under-priced security and short in the overpriced one, is a standard arbitrage technique already observed in all other types of arbitraging activities. In addition, assuming that the historical similarity tends to be complied upon again, the investors who have got there first will, as usual, reap all the benefits. However, it must be highlighted that this type of arbitrage is not completely risk-free, as there is no guarantee that the relationship between the two securities will reaffirm itself in the future.

3. Impact of HFT on arbitrage

The most common opinion among academics is that HFT is, in general, beneficial as far as market efficiency is concerned. Aitken et al. (2012) provide evidence of improved efficiency in the wake of increases in HFT, even if the same paper admits that "despite being in the spotlight for some time, our understanding of high-frequency trading and its implications for market quality are at best, moderately informed" (p.3). However, the authors take the cost of trading as a proxy for efficiency, although the term "efficiency" in this context is more akin to the meaning of "efficient allocation of resources" rather than to the one used by Nobel laureate Eugene Fama in his original explanation of the Efficiency Market Hypothesis. Aitken et al. (2012) also find that most academic papers agree on a "predominantly positive overall impact" (p.8) of HFT on market efficiency. However, they also quote a 2012 study by McInish and Upson (2012), who developed a theoretical framework modelled on the US equity market that demonstrates how, under certain conditions, HF traders can profit from their superior knowledge of the current state of the market, thanks to both faster data analytical capabilities and quicker access to the order books, against slower competitors. Price discovery is another proxy often used for market efficiency, as it is closely linked to arbitraging. If the price is not aligned with the true value of a security, then arbitrage opportunities arise, and HFT operators are in the best position for assessing and exploiting them, wiping out such opportunities, as expected by EMH.

However, there is another aspect that, paradoxically, highlights market inefficiencies even in presence of all EMH pre-conditions. Let us assume that the markets are efficient in the semi-strong form. It means that, as shown by Fama (1965) and Fama (1970), as soon as an arbitrage opportunity arises, it will be *instantaneously* discounted by the market. As discussed above, the word *instantaneously* is relative to the technological environment, the common requirement being that any arbitrage opportunities are incorporated quickly enough into prices by the market. But an efficient market is made up of a large number of participants - and it is the community of the market participants that adjust prices according to the newly created situation. According to the EMH, *some* market operators (called arbitrageurs) actually do make abnormal profits when they re-adjust the price exploiting an arbitrage opportunity. Malkiel (2003) puts it stating that "any truly repetitive and exploitable pattern that can be discovered in the stock market and can be arbitraged away will self-destruct" (p.72), and Wilson and Marashdeh (2007) admit that arbitrage is a sort of short-term inefficiency that ensures long-term

efficiency. An apparent paradox seems to arise at this point: for the markets to be efficient arbitrageurs must exist to bring arbitrage opportunities down by exploiting them, but if they do so, abnormal risk-less profit would arise, making the markets inefficient. The EMH had to deal with arbitrageurs somehow and, if their existence could not be denied, the Hypothesis had at least to show some special characteristics of theirs.

The way it solves the paradox is by stating that nobody is able to make abnormal profits in a *consistent* way, by exploiting arbitrage opportunities because all rational investors (that is, all investors, except the most naïve or casual ones) will compete to do so. In an efficient market abnormal profits are therefore split between all (or at least between very many) operators, which sometimes reap the abnormal profit through arbitrage and sometimes, finding themselves on the wrong side of the arbitrage, suffer a loss [Fama (1965)]. With a sufficiently large number of observations, the abnormal profit for each arbitrageur averages to zero [Van Horne (1995)]. At the end of the day (or the month, or the year) the cumulative abnormal profits and losses for each operator ought to tend balancing each other out, reaffirming the market efficiency principle. Yet, there is a big difference between *never making any* abnormal profit and *averagely making no* abnormal profit. Again, if one participant, or a small group (relative to the total number) of participants (like HF traders) is *consistently* able to arrive first to exploiting arbitrage opportunities, then the average does no longer hold for them, although it may still hold well for the market as a whole, or the whole arbitrage community. HFT is an issue of modern technology that does not enable all participants equal access to the news, because ultra-fast access is not equal for all - quite the opposite, it is the main reason for technology-led market inefficiency. Obviously, the inefficiency only persists as long as the fast arbitrageurs are a small minority of the investor community. If they become the majority (or ideally, the totality) of the market participants, then the Efficiency Market Hypothesis will hold again, as the average abnormal profit *for each participant* (and not just for the totality of them) over the long-term would revert to the mean expected by the EMH: zero.

4. Arbitrage and transaction costs

As first devised by Fama (1970), transaction costs are a tool leading toward market efficiency. If small discrepancies are found, they may not be exploitable because of transaction costs, which would make the trade unprofitable. Under these conditions, the discrepancy is immaterial and does not exist from any practical point of view. This is true for all market participants and for HF traders too, but some considerations are to be taken into account. Impact analysis of the Markets in Financial Instruments Directive (MiFID) in some sense leads Aitken et al. (2012) to turn the table by stating that the regulatory changes are the cause of both implicit and explicit trading cost falling, from which HFT benefited, rather than the latter being the cause of the former. Sornette and von der Becke (2011), despite generally criticising HFT practice, recognise that HFT does improve liquidity, and that higher liquidity and higher volumes traded (usually linked to HFT activity) tend to lower transaction costs. According to Jain (2005), computerized trading systems lower spreads, fees, brokerage, and commission costs but Friederich and Payne (2011) suggest that the originator of such market efficiency improvements may not be HFT activity. Brogaard et al. (2012) go so far as to argue that HFT may also be the cause of an increase in transaction costs. They also highlight that most academic literature supporting the view according to which HFT lowers transaction costs, does only take into account execution costs, leaving to discuss other components of transaction costs, as commissions and technology costs.

5. Is arbitrage a real issue?

After having shown that HF traders can exploit arbitrage opportunities and make consistent abnormal profits, it is important to understand whether or not arbitrage opportunities are a real thing, and how often they occur in real life. In one of his papers on the EMH, Fama (1970) deals with this issue, concluding that departures from the independent price assumption, whereas they may deny the random walk model, are not incompatible with the Efficient Market Hypothesis. The matter is settled by recognising a role of arbitrage in absorbing minor market inefficiencies - and avoiding them to accumulate up to an intolerable level. The beneficial role of arbitrage is also appreciated by Shleifer and Vishny (1997), because of its effect to bring prices to fundamental values and to keep markets efficient. However, the critical point of involving the investor community at large, and having them to share the profits brought in by arbitraging cannot hold the analysis carried out by Shleifer and Vishny (1997):"the millions of little traders are typically not the ones who have the knowledge and information to engage in arbitrage. More commonly, arbitrage is conducted by relatively few professional, highly specialized investors who combine their knowledge with resources of outside investors to take large positions" (p.2).

5.1 News arbitrage

When news reach the market, they sometimes have unclear consequences but at some other times consequences are crystal clear. Iraq's invasion of Kuwait back in 1990 was unanimously interpreted by the oil markets as bad news - and prices raised accordingly. With automatic news feed interpretation, those who enjoy highest speed can exploit some unambiguous news to their own advantage. According to the classic Efficient Market Hypothesis, markets would *instantaneously* discount newly arrived information where the focus here is on the word "instantaneously", which may carry different quantitative meaning for different market participants. The casual trader may become aware of the information during the evening news television programme, whereas a professional trader may know it via its Bloomberg screen, and a HF trader equipped with automatic news feed interpretation may, by then, have already closed all its tradings. This is undeniably an arbitrage opportunity.

5.2 Latency arbitrage

In the HFT world, the focus is mostly on latency - and according to Arnuk and Saluzzi (2009), arbitrage is no exception: "We believe Latency Arbitrage is more than a simple case of technological evolution, but raises serious questions about the fairness and equal access of US equity markets" (p.1). They compare HFT to having access to the relevant news contained in the Wall Street Journal five microseconds into the future, and they report that thanks to this information, HFTs are able to achieve "(almost) risk free arbitrage opportunities" (p.2). The example they show is instructive.
1. The book for stock ABC shows $25.53 bid / $25.54 ask.
2. Due to Latency Arbitrage, a HFT computer realises that there is an incoming order that in a fraction of a second (but still a long time for HFT's standards) will move the NBBO quote higher, to $25.54 bid /offered at $25.56. How a HFT computer may get this information is explained in detail later.
3. The HFT hurries up, buying all available ABC shares at $25.54 and cheaper.
4. The institutional algorithm does not have its bid order at $25.54 executed as there is no longer stock available at this price and the market moves up to $25.54 bid / $25.56 ask as anticipated by the HFT after inspecting the order books.
5. The HFT then turns around and offers ABC at $25.55 or $25.56.

6. Because it is following a volume driven formula, the institutional algorithm is forced to buy available shares from the HFT at $25.55 or $25.56.

7. The HFT makes $0.01-$0.02 per share at the expense of the institutional investor.

With this tiny profit per share and per execution, Arnuk and Saluzzi (2009) calculate a daily profit ranging between $6 and $12 million, which multiplied by 250 trading days per year translates into $1.5-$3 billion profit generated at the expenses of retail and institutional investors. Obviously, there is no guarantee that on every trading day an institutional investor will launch a large buying or selling programme but the idea underlying the example is clear. According to a HFT software firm Advent's white paper, Advent (2012), "[i]n most instances, high-frequency
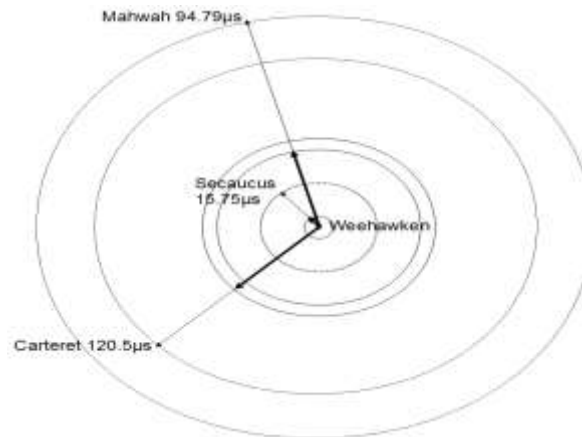
traders don't bet on the value of a company, currency or commodity - or the future outlook - but are simply looking to arbitrage price discrepancies in securities that are trading simultaneously on different exchanges or trading platforms" (p.7). Practical evidence of how latency arbitrage works is described in "Flash Boys" [Lewis (2014)].

Figure 1 In the microseconds it takes a HF trader — depicted with thin line — to reach the various stock exchanges housed in these New Jersey towns, the conventional trader's order, theoretically, makes it only as far as the bold line. Adapted from Lewis (2014)

Everything started when in early 2007 the trader team led by Brad Katsuyama realised that when they tried to complete a trade it became nearly impossible to do it properly, as the shares on offer suddenly vanished. Initially they thought of a computer or connection problem but an investigation showed that the market seemed to specifically respond to their action by removing liquidity as and when they were about to take it. As it turned out, it was not their exclusive problem - it was a market-wide issue. Katsuyama's team launched a deep study of the problem on 13 stock exchanges scattered over three different sites run by NYSE, Nasdaq, BATS and Direct Edge (Secaucus, Carteret, and Mahwah, all in New Jersey).They discovered that their orders were properly executed when sent to an individual exchange but did not when launched to all of them at the same time. One exception was noticed: no matter how many exchanges were their order sent to, they always achieved 100% of their orders sent to the exchange located in Secaucus. Further investigation revealed that orders to Secaucus were always 100% successful because it was located nearest to their premises in Weehawken. Further away exchanges did show phantom liquidity. Indeed, when a modification to their trading platform arranged to have all orders to reach all the exchanges at the same time, it turned out they achieved 100% successful execution. The issue might be simplified by modelling a trader T and two exchanges, E1 and E2, located at different distances from T. Let us suppose trader T reads the price P for stock S, which it considers appealing and then it sends a relatively large market order for a quantity Q of shares to the market, which is routed to E1. Yet, E1, the nearest exchange to T,

cannot trade all Q shares as requested by T, because it can only provide for nQ of them, with n < 1. So, E1 accepts the order for nQ shares, executes it and routes the order for the remaining (1-n)Q shares to E2. The discovery made by Katsuyama and his team, as described in Lewis (2014) is that the second order would never be executed at the originally intended price because the shares on the book would be bought (or sold) before T's order arrived there. Indeed, a HF trader, noticing the order filled on E1 at price P, would immediately (exploiting its superior technology and faster connections) front run T's order to E2, trading all shares at the current price P and posting a limit order at price P +Δp (with Δp greater or less than zero according to whether T intended to buy or to sell, but still better than the next limit order for the same stock S). When T's market order eventually arrives at E2 the price it is being executed at is worse than originally intended, to the benefit of the HF trader, somehow "frontrunning" it. What was happening to Katsuyama's team (and many others) was exactly that. Since there was little doubt they were not the first ones to realise what was going on, the only reasonable explanation was that those who understood the issue were making money off it.

This practice may be a legal form of frontrunning, but still leading to abnormal, consistent, risk-free profit. It is a form of arbitrage. Practitioners are generally rather suspicious about HFT. ZeroHedge blog by Tyler Durden is a very popular website among professionals; it publishes articles and receives many comments on a wide range of financial matters and not surprisingly, HFT is a hotly debated issue. About market efficiency, Durden (2014) states "High frequency traders use ultra-high speed connections with trading venues and sophisticated trading algorithms to exploit inefficiencies created by the new market structure and to identify patterns in 3rd parties' trading that they can use to their own advantage". A critical piece of information is the investment pattern followed by algorithms (usually more predictable than humans). "Information is leaked when electronic algorithms reveal patterns in their trading activity. These patterns can be detected by HFTs who then make trades that profit from them" (Durden 2014), leading to the sharp conclusion "HFT is legal frontrunning".

6. An arbitrage simulation

In order to provide quantitative evidence to what discussed above, I have developed a computer-based simulation to test the hypothesis that it is impossible to claim consistent abnormal profits.

6.1 HFT on EMH simulation

The simulation is made up of two markets trading the same set of securities. At each cycle a discrepancy arises in one of the markets by forcing the price change of one randomly-chosen security. All participants immediately notice that price difference but since the purpose of the simulation is to provide evidence of the principle, not working out the profit made through arbitrage, for sake of simplicity and without loss of generality, I set that one and only one trader is able to grasp the opportunity and that just one security is traded at every cycle. One randomly chosen trader buys the security in the market where the price is lower and sells it in the other market.

The user-defined parameters read at the beginning of each run are now described. The number in brackets is the values I used for the corresponding parameter in the simulation.

Number of traditional traders (11,859)

Number of HF traders (15)

Number of securities traded in both markets (500, e.g. securities in S&P 500 index)

Number of cycles to run the simulation for (100)

Maximum allowed price variation in cash units (0.1)

Trading ratio between HF and traditional traders (381)

Transaction costs, including bid-ask spread and fees (3.5 basis points)

The Launch subroutine repeatedly calls a certain number of times the Arbitrage subroutine, which implements the simulation. The structure of the algorithm is as follows.

The *Launch* subroutine reads some configuration parameters and then, for the specified number of times, it calls the *Arbitrage* subroutine, it works out the total of the cash account plus the securities account for each trader and then it calls the *ComputeAverage* subroutine.

At the end of the loop, it computes average and stdev of all repetitions

The Arbitrage Subroutine runs 100 times the following steps: it randomly selects one security, it randomly select the price variation applied to that security (within +/- the maximum allowed value) and randomly selects the market on which the price variation occurs (market A or market B). It then applies the price variation to the selected security only in the selected market. At this stage, one trader notices the arbitrage opportunity and the trader fastest to arbitrage according to the HF-to-non-HF-trader ratio parameter is randomly selected. Next, the price variation is added to the arbitrageur's cash account and the transaction costs are subtracted. The securities accounts of the two traders holding the security object of the transactions is updated accordingly (one security added to the buyer's account and subtracted from the seller's account).

The ComputeAverage loops through all the HF traders adding their cash accounts together ant then divides the result by the number of HF traders. The reason for having several repetitions of the Arbitrage subroutine is to provide a suitably high number of data that allows using the standard deviation of the sample as replacement of the standard deviation for the entire population. The reason for having several iterations (100 cycles in this case) within the Arbitrage subroutine is to simulate a suitable number of arbitrage opportunities occurred within the observation period. A quantitative estimation of transaction costs depends on various factors, including the venue, the fee structure, the bid-ask spread (which on its turn depends on market conditions), and possibly others. I have used 3.5 basis points as a rough estimate, bearing in mind that, according to many academic studies, higher HFT activity contributed to lowering transaction costs but also that too low an estimate of transaction costs would falsely increase arbitrage profits. For technical reasons, every computer-based random number generation needs a lower and an upper limit. For initialising the security price, I set those limits to 20 and 80 cash units (being GBP, USD, EUR or any others, or even fractions thereof). Whereas this choice is arbitrary, it is always possible to write arbitraging algorithms that act only upon securities whose price change falls within a predetermined range, so my choice does not drive the result in a way or another. HF traders are set to hold no securities throughout the simulation period, to sterilise their final account of any market gain or loss; the difference between their final and initial account only entails cash variations due to exploitation of arbitrage opportunities.

The numbers of HF and non-HF traders are taken from CFTC-SEC (2010), where at page 29 it displays summary statistics whence it is possible to extract the number of HF and non-HF (split among Intermediary, Buyer, Seller, Opportunistic and Noise) traders active in the E-mini contracts market on May 6, 2010 as well on the previous three days. I took the figures from the latter statistics in order to avoid any reference to an exceptional situation as the day of the Flash Crash. The same statistics reports the percentage of trades carried out by the different categories of traders and it is therefore possible to work out the ratio between the HF versus non-HF trades on those three days. According to CFTC-SEC (2010) the 15 HF traders carried out 32.56% of all trades, whereas the other categories (11,859 trader's altogether) shared the remaining 67.44% trades. Therefore, the average HF trader took 2.17% of all trades whereas the average traditional

trader only closed less than 0.0057% of the trades. The ratio between the two figures yields 381.7, which I took freedom to round down to 381 HF trades for every 1 non-HF trade. The run of the algorithm yields an average profit for each HF trader of 0.21 cash units (against the null hypothesis of zero, as expected by the EMH) and a standard deviation of 0.02 cash units. Bearing in mind the considerations above about transaction costs, I also launched a run of the simulation with a figure twice as high (7 basis points), which yielded an average profit for HF traders of 0.07 cash units and a standard deviation of 0.021. The result is extraordinarily strong for both levels of transaction costs, at virtually *any* significance level, against the hypothesis of no consistent risk-free gain. Therefore, this simulation leads to reject the null hypothesis - and to reject the Efficient Market Hypothesis *in presence of HF traders* with it.

6.2 Discussion
The result achieved raises three types of considerations:
a) the significance of the distorting role of HFT to market efficiency in the simulation is so strong that it is legitimate to suspect that the result was far too trivial;
b) should other investors worry about the impact HFT seems to have on market efficiency ?;
c) is the result something that should attract the regulators' attention or is it just another chapter of the old story according to which the smart guys have always made more money than (and to the expenses of) the dumb ones ?

6.2.1 Is the Test Far Too Trivial ?
The result is so strong that suspicions of a pre-determined test whose outcome is implicit in its starting assumptions cannot be easily dismissed. Yet, several authors, Aldridge (2010) among others, share the opinion that "[w]hoever detects the mispricing and gets his order posted on the exchange first is likely to generate the most profit" (p.245). According to Jarrow and Protter (2011) "with high frequency traders, we can show that there exist no arbitrage opportunities for ordinary traders" (p.2), which means that all arbitrage opportunities would be exploited by HF traders. The initialisation parameters have all been taken from the real world. The only exception is the number of arbitrage opportunities (called iterations in the algorithm), which I arbitrarily set to 100. The reason for using a relatively large number is to smooth out the randomness of the individual arbitrage opportunity. Moreover, in no one iteration the profit of the trader exploiting the arbitrage opportunity was negative. Therefore, the number of iterations does not distort the total amount of the arbitrage profit. The number of iterations balances out the unrealistic simplification of trading only one security per each arbitrage opportunity. The general idea is rather obvious: if a small group of traders has a definitive speed advantage over the large remaining majority, it is intuitive that as soon as an arbitrage opportunity arises, it shall be picked up by the fastest traders. A question may indeed arise about what will happen when the number of HF traders increases. In the simulation, all HF traders have the same probability to exploit the opportunity and if their number increases, they will all enjoy a lesser share of the arbitrage profits. But the principle does not change: as long as the value of the opportunity overcomes transaction costs, the ones who exploit the opportunity will make a risk-free profit at expenses of slower participants.

6.2.2 Is the Result Worrying for Other Investors ?
There is widespread worry among (non-HFT) practitioners about HF practice but, as seen above, academics are much more cautious. The latter opinion would obviously be of little help should slow practitioners get convinced they are playing an unfair game and thus decide to withdraw the industry altogether. However, there is no sign of such a trend. Financial markets are busy as ever and poor results are usually interpreted as long-wave effect of the 2008 crisis

rather than caused by HFT activities. I must stress the fact that in this research I did not purposely take into account any market manipulation factor. The purpose here is to evaluate the impact of clean, straight and transparent high-frequency trading practices have on market efficiency and not whether, as suggested by some practitioners and academics alike, HFT is suitable to be used for market manipulation goals. HFT-led manipulative practices may well be the scope of further research. Lewis (2014) lucidly explains some of the non-manipulative (as well as some manipulative) effects of HFT and in particular the slow-market arbitrage activity described there is essentially the one I have simulated in this paragraph. On one side, it is true that any successful innovation worries those who cannot replicate or satisfactorily tackle it. On the other side, this situation has occurred countless times in history – and financial market's history is no exception. Yet, financial markets continued to survive and to provide their services to the world economy.

6.2.3 Should the Result Alert Regulators ?

The role of regulators is obviously different from practitioners' but everybody's opinions must be (and usually are) taken into account ("*Es gibt noch Richter in Berlin*"). The playing field must not only be perceived as even, it must be even under any respect and it is beyond doubt that HFT practices do pose a question to regulators. On the other side, the search for speed, and in general the effort to ensure a more favourable position with respect to others, is a recurring theme in financial markets. The floor traders nearer to the broker did actually enjoy a favourable position. Remote traders that used telegraph instead of manually dispatched orders also enjoyed a favourable position, and so did telephone users with respect to the telegraph ones. A half-legend with a lot of truth in it (J.Kay, Financial Times, 28/5/2013), tells that when in 1815 Napoleon was defeated at Waterloo, Nathan Rotschild was the first to know, thanks to his carrier pigeons that allowed him to make a fortune by buying British government bonds and shorting French ones. Speed is often a definite advantage, and not only in Formula 1, but is this a valid criterion for requesting a stricter regulation of HFT practices? In many cases, some kind of advantage is accepted. Nobody would accuse a trader firm of unfair advantage just because of the smartness of its software, nor because of the terabytes of its disk or the gigaflops allowed by its computer's CPU (speed internal to CPU seems to be tolerated). These advantages are rather regarded as assets to replicate than as swindles to blame. Yet, networking speed and co-location present different characteristics from other computer-related innovations. Software, CPU and disks would not turn efficient markets into inefficient ones. No matter how smart software is, it will never nullify trading risk. Software may help traders to take more informed decisions, to diversify risk in a more rational manner, to replicate more closely past successful strategies – but risk will still be there. Disks may store, and CPUs may process, longer streams of historical data, but they will never alone guarantee risk-free return. Only networking speed and co-location seem to provide the possibility of falsifying the Efficient Market Hypothesis.

However, the obvious objection would be that such plusses could be replicated by anyone willing, and able, to invest enough time, resources and dedication to reach similar, or even better, results. In the end, the history of Mankind has always been one of today's losses and tomorrow's wins – the delta being the endeavour employed. Should not the main effort by regulators be the one to incentivise more HFT rather than less? To make competition easier for the many rather than harder for all? As suggested by Foresight (2012), "the more competitive the HFT industry, the more efficient will be the market in which they work" (p.54). The question is still a very open one.

7. Conclusion on HFT arbitraging

The considerations developed above lead to the conclusion that, although many academics support the beneficial role of HFT in improving market efficiency, price discovery, and reduced transaction costs, there are also contrasting views on this matter. Because of their speed advantage, a relatively small number of HF traders are able to beat, most of the times, the majority of slower competitors in adjusting their own limit orders, cancelling them out altogether, and aggressively picking off other investors' outstanding limit orders. Moreover, using a simulation of arbitrage opportunities, I found that a small number of HF traders could consistently achieve risk-free return by reaping most of the arbitrage-led gains rather than letting them being spread evenly among the larger number of investors, in contrast to the Efficient Market Hypothesis. Certainly, HF traders speed up price discovery - but at their own benefit and to the detriment of the many other slower investors, that is, to the detriment of the market at large. Yet, whether or not this violation of the market efficiency principles should be further regulated is a debated matter. HF traders also have the chance to exploit the anomaly permitted by many Exchanges' regulations of behaving like market makers when it suits them, profiting from the spread and taking advantage of liquidity providing rebates as incentives, without bearing any of the obligations market makers carry when market courses penalise this category of investors and drastically reducing the risk of being adversely selected. As a last point, the debate on whether or not HF traders are the major contributors to decreasing transaction costs seems still rather open, and if many academics support the view of HFT decreasing transaction costs, there are other, not less authoritative, voices carrying a different opinion also worth listening to.

References

Advent (2012): Advent Software Inc. (2012), *High-Frequency Trading: Useful Liquidity Tool or Weapon of Financial Mass Destruction?*, White Paper, San Francisco: Advent Software Inc.

Aldridge (2010): Aldridge, I. (2010) *High Frequency Trading*, Hoboken: John Wiley and Sons

Aitken et al. (2012): Aitken, M., deB. Harris, F., McInish, T., Aspris, A., Foley, S. (2012); *High frequency trading - assessing the impact on market efficiency and integrity*, DR28, London: Government Office for Science

Arnuk and Saluzzi (2009) - Arnuk, S., Saluzzi, J. (2009); *Latency Arbitrage: The Real Power Behind Predatory High Frequency Trading*; Themis Trading LLC, White paper

Brogaard et al. (2012): Brogaard, J., Hendershott, T., Hunt, S., Latza, T., Pedace, L., Ysusi, C. (2012); *High-frequency trading and the execution costs of institutional* investors, DR21, London: Government Office for Science

CFTC-SEC (2010): CFTC, SEC (2010); *Findings Regarding the Market Events of May 6, 2010*, Washington: CFTC-SEC

Durden (2014) - Durden, T. (2014); *High Frequency Trading: All You Need To Know*; www.ZeroHedge.com (accessed 08/04/2015)

Fama (1965): Fama, E. (1965) The Behavior Of Stock Market Prices, *The Journal of Business*, 38 (1), p. 34-105

Fama (1970): Fama, E. (1970) Efficient Capital Markets: A Review of Theory and Empirical Work, *The Journal of Finance*, 25 (2), p.383-417

Foresight (2012): Foresight (2012), *The Future of Computer Trading in Financial Markets*, Final Project Report, The Government Office for Science, London

Grossman and Stiglitz (1980): Grossman, S., Stiglitz, J. (1980) On the Impossibility of Informationally Efficient Markets, *The American Economic Review*, 70 (3), p. 393-408

Jain (2005): Jain, P. (2005), Financial Market Design and the Equity Premium: Electronic versus Floor Trading, *The Journal Of Finance*, LX (6), p. 2955.2985

Jarrow and Protter (2011): Jarrow, R., Protter, P. (2011) *A Dysfuntional Role of High Frequency Trading in Electronic Markets*, Johnson School Research Paper Series #08-2011, Johnson Cornell University (accessed 28/11/2013)

Kirkpatrick and Dahlquist (2007): Kirkpatrick, C., Dahlquist, J. (2007) *Technical Analysis - The Complete Resource For Financial Market Technicians*, Upper Saddle River: FT Press

Leshik and Cralle(2011): Leshik, E., Cralle, J. (2011) *An Introduction to Algorithmic Trading*, Padstow: John Wiley and Sons

Lewis (2014): Lewis, M. (2014) *An Adaptation from 'Flash Boys': A Wall Street Revolt*; NYTimes.com, March 2014 (accessed 12/05/2014)

Malkiel (2003): Malkiel, B. (2003) The Efficient Market Hypothesis and Its Critics, *Journal of Economic Perspectives*, 17 (1), p-59-82

McInish and Upson (2012): McInish, T., Upson, J. (2012) *Strategic Liquidity in a Market with Fast and Slow Traders*, Working Paper, University of Texas (accessed 24/12/2013)

Shleifer and Vishny (1997) - Shleifer, A., Vishny, R. (1997) The Limits of Arbitrage; *The Journal of Finance*, LII (1), p.35-55

Sornette and von der Becke (2011): Sornette, D., von der Becke, S. (2011) *Crashes and High Frequency* Trading, DR7, London: Government Office for Science

Van Horne (1995): Van Horne, J. (1995) *Financial Management and Policy*, Englewood Cliffs: Prentice Hall International Inc.

Wilson and Marashdeh (2007): Wilson, E., Marashdeh, H. (2007) Are Co-integrated Stock Prices Consistent with the Efficient Market Hypothesis?, *The Economic Record*, 83 (s1), s87-s93