A computational view of market efficiency

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.
A Computational View of Market Efficiency*

Jasmina Hasanhodzic†, Andrew W. Lo‡, and Emanuele Viola††

August 31, 2009

Abstract

We propose to study market efficiency from a computational viewpoint. Borrowing from theoretical computer science, we define a market to be efficient with respect to resources $S$ (e.g., time, memory) if no strategy using resources $S$ can make a profit. As a first step, we consider memory-$m$ strategies whose action at time $t$ depends only on the $m$ previous observations at times $t - m, \ldots, t - 1$. We introduce and study a simple model of market evolution, where strategies impact the market by their decision to buy or sell. We show that the effect of optimal strategies using memory $m$ can lead to “market conditions” that were not present initially, such as (1) market bubbles and (2) the possibility for a strategy using memory $m' > m$ to make a bigger profit than was initially possible. We suggest ours as a framework to rationalize the technological arms race of quantitative trading firms.

Keywords: Market Efficiency; Computational Complexity.

1 Introduction

Market efficiency—the idea that prices fully reflect all available information—is one of the most important concepts in economics. A large number of articles have been devoted to its formulation, statistical implementation, and refutation since Samuelson (1965) and Fama (1965a,b; 1970) first argued that price changes must be unforecastable if they fully incorporate the information and expectations of all market participants. The more efficient the

---

*The views and opinions expressed in this article are those of the authors only, and do not necessarily represent the views and opinions of AlphaSimplex Group, MIT, Northeastern University, any of their affiliates and employees, or any of the individuals acknowledged below. The authors make no representations or warranty, either expressed or implied, as to the accuracy or completeness of the information contained in this article, nor are they recommending that this article serve as the basis for any investment decision—this article is for information purposes only. This research was supported by the MIT Laboratory for Financial Engineering and AlphaSimplex Group, LLC.

†Senior Research Scientist, AlphaSimplex Group.

‡Harris & Harris Group Professor, MIT Sloan School of Management.

††Assistant Professor, College of Computer and Information Science, Northeastern University, 440 Huntington Ave, #246WVH, Boston, MA 02115, USA (corresponding author).
market, the more random the sequence of price changes generated by it, and the most efficient market of all is one in which price changes are completely random and unpredictable.

According to the proponents of market efficiency, this randomness is a direct result of many active market participants attempting to profit from their information. Driven by profit opportunities, legions of investors pounce on even the smallest informational advantages at their disposal, and in doing so, they incorporate their information into market prices and quickly eliminate the profit opportunities that first motivated their trades. If this occurs instantaneously, as in an idealized world of frictionless markets and costless trading, then prices fully reflect all available information. Therefore, no profits can be garnered from information-based trading because such profits must have already been captured. In mathematical terms, prices follow martingales.

This stands in sharp contrast to finance practitioners who attempt to forecast future prices based on past ones. And perhaps surprisingly, some of them do appear to make consistent profits that cannot be attributed to chance alone.

1.1 Our Contribution

In this paper we suggest that a reinterpretation of market efficiency in computational terms might be the key to reconciling this theory with the possibility of making profits based on past prices alone. We believe that it does not make sense to talk about market efficiency without taking into account that market participants have bounded resources. In other words, instead of saying that a market is “efficient” we should say, borrowing from theoretical computer science, that a market is efficient with respect to resources $S$, e.g., time, memory, etc., if no strategy using resources $S$ can generate a substantial profit. Similarly, we cannot say that investors act optimally given all the available information, but rather they act optimally within their resources. This allows for markets to be efficient for some investors, but not for others; for example, a computationally powerful hedge fund may extract profits from a market which looks very efficient from the point of view of a day-trader who has less resources at his disposal—arguably the status quo.

As is well-known, suggestions in this same spirit have already been made in the literature. For example, Simon (1955) argued that agents are not rational but boundedly rational, which can be interpreted in modern terms as bounded in computational resources. Many other works in this direction are discussed in §1.2. The main difference between this line of research and ours is: while it appears that most previous works use sophisticated continuous-time models, in particular explicitly addressing the market-making mechanism (which sets the price given agents’ actions), our model is simple, discrete, and abstracts from the market-making mechanism.

Our model and results: We consider an economy where there exists only one good. The daily returns (price-differences) of this good follow a pattern, e.g.:

\[ (-1, 1, -1, 1, -1, 1, 2, -1, 1, -1, 1, 2, -1, 1, -1, 1, -1, 1, 2, \ldots) \]

which we write as \([-1, 1, -1, 1, -1, 1, 2]\), see Figure 1 on Page 4. A good whose returns follow a pattern may arise from a variety of causes ranging from biological to political; think of
seasonal cycles for commodities, or the 4-year presidential cycle. Though one can consider more complicated dependencies, working with these finite patterns keeps the mathematics simple and is sufficient for the points made below.

Now we consider a new agent $A$ who is allowed to trade the good daily, with no transaction costs. Intuitively, agent $A$ can profit whenever she can predict the sign of the next return. However, the key point is that $A$ is computationally bounded: her prediction can only depend on the previous $m$ returns. Continuing the above example and setting memory $m = 2$, we see that upon seeing returns $(-1, 1)$, $A$'s best strategy is to guess that the next return will be negative: this will be correct for 2 out of the 3 occurrences of $(-1, 1)$ (in each occurrence of the pattern).

We now consider a model of market evolution where $A$’s strategy impacts the market by pushing the return closer to 0 whenever it correctly forecasts the sign of the return — thereby exploiting the existing possibility of profit — or pushing it away from 0 when the forecast is incorrect — thereby creating a new possibility for profit. In other words, the evolved market is the difference between the original market and the strategy guess (we think of the guess in $\{-1, 0, 1\}$). For example, Figure 2 shows how the most profitable strategy using memory $m = 2$ evolves the market given by the pattern $[-2, 2, -2, 2, -2, 2, 3]$ (similar to the previous example).

The main question addressed in this work is: how do markets evolved by memory-bounded strategies look like?

We show that the effect of optimal strategies using memory $m$ can lead to market conditions that were not present initially, such as (1) market bubbles and (2) the possibility that a strategy using memory $m' > m$ can make larger profits than previously possible. By market bubbles (1), we mean that some returns will grow much higher, an effect already anticipated in Figure 2. For the point (2), we consider a new agent $B$ that has memory $m'$ that is larger than that of $A$. We let $B$ trade on the market evolved by $A$. We show that, for some initial market, $B$ may make more profit than what would have been possible if $A$ had not evolved the market, or even if another agent with large memory $m'$ had evolved the market instead of $A$. Thus, it is precisely the presence of low-memory agents ($A$) that allows high-memory agents ($B$) to make large profits.

Regarding the framework for (2), we stress that we only consider agents trading sequentially, not simultaneously: after $A$ evolves the market, it gets incorporated into a new market, on which $B$ may trade. While a natural direction is extending our model to simultaneous agents, we argue that this sequentiality is not unrealistic. Nowadays, some popular strategies $A$ are updated only once a month. Within any such month, a higher-frequency strategy $B$ can indeed trade in the market evolved by $A$ without $A$ making any adjustment.

### 1.2 More related work

Since Samuelson’s (1965) and Fama’s (1965a,b; 1970) landmark papers, many others extended their original framework, yielding a “neoclassical” version of the efficient market hypothesis where price changes, properly weighted by aggregate marginal utilities, are unforecastable (see, for example, LeRoy, 1973; Rubinstein, 1976; and Lucas, 1978). In markets where, according to Lucas (1978), all investors have “rational expectations,” prices do fully
Figure 1: The market corresponding to a market pattern.

Figure 2: Optimal memory-2 strategy evolves the market and creates a bubble.

Figure 3: Optimal memory-3 strategy evolves the market without creating a bubble.
reflect all available information and marginal-utility-weighted prices follow martingales. Market efficiency has been extended in many other directions, but the general thrust is the same: individual investors form expectations rationally, markets aggregate information efficiently, and equilibrium prices incorporate all available information instantaneously. See Lo (1997, 2007) for a more detailed summary of the market efficiency literature in economics and finance.

There are two branches of the market efficiency literature that are particularly relevant for our paper: the asymmetric information literature, and the literature on asset bubbles and crashes. In Fischer Black's (1986) presidential address to the American Finance Association, he argued that financial market prices were subject to “noise”, which could temporarily create inefficiencies that would ultimately be eliminated through intelligent investors competing against each other to generate profitable trades. Since then, many authors have modeled financial markets by hypothesizing two types of traders—informed and uninformed—where informed traders have private information regarding the true economic value of a security, and uninformed traders have no information at all, but merely trade for liquidity needs (see, for example, Grossman and Stiglitz, 1980; Diamond and Verrecchia, 1981; Admati, 1985; Kyle, 1985; and Campbell and Kyle, 1993). In this context, Grossman and Stiglitz (1980) argue that market efficiency is impossible because if markets were truly efficient, there would be no incentive for investors to gather private information and trade. DeLong et al. (1990, 1991) provide a more detailed analysis in which certain types of uninformed traders can destabilize market prices for periods of time even in the presence of informed traders. More recently, studies by Luo (1995, 1998, 2001, 2003), Hirshleifer and Luo (2001), and Kogan et al. (2006) have focused on the long-term viability of noise traders when competing for survival against informed traders; while noise traders are exploited by informed traders as expected, certain conditions do allow them to persist, at least in limited numbers.

The second relevant offshoot of the market efficiency literature involves “rational bubbles”, in which financial asset prices can become arbitrarily large even when all agents are acting rationally (see, for example, Blanchard and Watson, 1982; Allen, Morris, and Postlewaite, 1993; Santos and Woodford, 1997; and Abreu and Brunnermeier, 2003). LeRoy (2004) provides a comprehensive review of this literature.

Finally, in the computer science literature, market efficiency has yet to be addressed explicitly. However, a number of studies have touched upon this concept tangentially. For example, computational learning methods have been applied to the pricing and trading of financial securities by Hutchinson, Lo, and Poggio (1994). Evolutionary approaches in the dynamical systems and complexity literature have also been applied to financial markets by Farmer and Lo (1999), Farmer (2002), Farmer and Joshi (2002), Lo (2004), Farmer, Patelli, and Zovko (2005), Lo (2005), and Bouchaud, Farmer, and Lillo (2008). And simulations of market dynamics using autonomous agents have coalesced into a distinct literature on “agent-based models” of financial markets by Arthur et al. (1994, 1997), Arthur, LeBaron, and Palmer (1999), Farmer (2001), and LeBaron (2001a-c, 2002, 2006). While none of the papers in these distinct strands of the computational markets literature focus on a new definition of market efficiency, nevertheless, they all touch upon different characteristics of this aspect of financial markets.
Organization. We provide our computational definition of market efficiency in Section 2. We describe the dynamics of market evolution in Section 3 and conclude in Section 4.

2 A computational definition of market efficiency

We model a market by an infinite sequence of random variables $X_1, X_2, \ldots$ where $X_t \in \mathbb{R}$ is the return at time $t$. For the points made in this work, it is enough to consider markets that are obtained by the repetition of a pattern. This has the benefit of keeping the mathematics to a minimum level of sophistication.

Definition 1 (Market). A market pattern of length $p$ is a sequence of $p$ random variables $[X_1, \ldots, X_p]$, where each $X_i \in \mathbb{R}$.

A market pattern $[X_1, \ldots, X_p]$ gives rise to a market obtained by the independent repetition of the patterns: $X_1^1, X_2^1, \ldots, X_p^1, X_1^2, X_2^2, \ldots, X_p^2, \ldots$, where each block of random variables $X_i^1, \ldots, X_i^j$ are distributed like $X_1, \ldots, X_p$ and independent from $X_1^j, \ldots, X_p^j$ for $j \neq 1$. Figure 1 shows an example for a market pattern in which the random variables are constant.

We now define strategies. A memory-$m$ strategy takes as input the previous $m$ observations of market returns, and outputs 1 if it thinks that the next return will be positive. We can think of this “1” as corresponding to a buy-and-sell order (which is profitable if the next return is indeed positive). Similarly, a strategy output of $-1$ corresponds to a sell-and-buy order, while 0 corresponds to no order.

Definition 2. A memory-$m$ strategy is a map $s : \mathbb{R}^m \to \{-1, 0, 1\}$.

We now define the gain of a strategy over a market. At every market observation, the gain of the strategy is the sign of the product of the strategy output and the market return: the strategy gains a profit when it correctly predicts whether the next return will be positive or negative, and loses a profit otherwise. Since we think of markets as defined by the repetition of a pattern, it is enough to define the gain of the strategy over this pattern.

Definition 3 (Gain of a strategy). The gain of a memory-$m$ strategy $s : \mathbb{R}^m \to \{-1, 0, 1\}$ over market pattern $[X_1, \ldots, X_p]$ is:

$$\sum_{i=1}^{p} E_{X_1, \ldots, X_p} \left[ \text{sign} \left( s(X_{i-m}, X_{i-m+1}, \ldots, X_{i-1}) \cdot X_i \right) \right],$$

where for any $k$, the random variable $(X_{1-kp}, \ldots, X_{kp})$ is independent from all the others and distributed like $(X_1, \ldots, X_p)$.

A memory-$m$ strategy $s$ is optimal over a market pattern if no memory-$m$ strategy has a bigger gain than $s$ over that market pattern.

The gain of an optimal memory-1 strategy over a certain market is a concept which is intuitively related to the (first-order) autocorrelation of the market, i.e. the correlation between the return at time $t$ and that at time $t+1$ (for random $t$), a well-studied measure.
of market efficiency (cf. Lo 2005, figure 2). This analogy can be made exact up to a normalization for markets that are given by balanced sequences of +1, −1. In this case higher memory can be thought of an extension of autocorrelation. It may be interesting to apply this measure to actual data.

Given a deterministic market pattern (which is just a sequence of numbers such as [2, −1, 5, ...]), optimal strategies can be easily computed: it is easy to see that an optimal strategy outputs +1 if and only if more than half the occurrences of x in the market pattern are followed by a positive value.

We are now ready to give our definition of market efficiency.

**Definition 4 (Market efficiency).** A market pattern \([X_1, \ldots, X_p]\) is efficient with respect to memory-\(m\) strategies if no such strategy has strictly positive gain over \([X_1, \ldots, X_p]\).

For example, let \(X_1\) be a \([-1,1]\) random variable that is −1 with probability 1/2. Then the market pattern \([X_1]\) is efficient with respect to memory-\(m\) strategies for every \(m\).

A standard “parity” argument gives the following hierarchy, which we prove for the sake of completeness.

**Claim 5.** For every \(m\), there is a market pattern that is efficient for memory-\(m\) strategies, but is not efficient for memory-(\(m + 1\)) strategies.

**Proof.** Let \(X_1, X_2, \ldots, X_{m+1}\) be i.i.d. random variables with range \([-1,1]\) whose probability of being 1 is 1/2. Consider the market pattern

\[
\left[ X_1, X_2, \ldots, X_{m+1}, X_{m+2} := \prod_{i=1}^{m+1} X_i \right]
\]

of length \(p = m + 2\). By the definition of the pattern, the distribution of each variable is independent from the previous \(m\), and therefore any memory-\(m\) strategy has gain 0.

Consider the memory-(\(m + 1\)) strategy \(s(x_1, x_2, \ldots, x_{m+1}) := \prod_{i=1}^{m+1} x_i\). Its gain over the market pattern is

\[
\sum_{i=1}^{p} E_{X_1, \ldots, X_p} \left[ \text{sign} \left( X_{i-m-1} \cdot X_{i-m} \cdots X_{i-1} \cdot X_i \right) \right] = E_{X_1, \ldots, X_p} \left[ \text{sign} \left( X_1 \cdot X_2 \cdots X_{m+1} \cdot X_{m+2} \right) \right] = 1.
\]

\[\square\]

### 3 Market evolution

In this section we describe the dynamics of our market model. We consider a simple model of evolution where the strategies enter the market sequentially. After a strategy enters, its “impact on the market” is recorded in the market and produces a new evolved market. The way in which a strategy impacts the market is by subtracting the strategy output from the market data.
Definition 6 (Market evolution). We say that a memory-$m$ strategy $s: \mathbb{R}^m \rightarrow \{-1, 0, 1\}$ evolves a market pattern $[X_1, X_2, \ldots, X_p]$ into the market pattern

$$[X_1 - s(X_{1-m}, \ldots, X_{1-1}), X_2 - s(X_{2-m}, \ldots, X_{2-1}), \ldots, X_p - s(X_{p-m}, \ldots, X_{p-1})].$$

Figure 2 shows the evolution of a market pattern.

Definition 6 can be readily extended to multiple strategies acting simultaneously, just by subtracting all the impacts of the strategies on the market, but for the points made in this paper the above one is enough.

In the next two subsections we point out two consequences of the above definition.

### 3.1 Market bubbles

In this section we point out how low-memory strategies can give rise to market bubbles, i.e. we show that an optimal low-memory strategy can evolve a market pattern into another one which has values that are much bigger than those of the original market pattern (in absolute value). Consider for example the market pattern $[-2, 2, -2, 2, -2, 2, 3]$ in Figure 2. Note how an optimal memory-2 strategy will profit on most time instances but not all. In particular, on input $(-2, 2)$, the best output is $-1$, which agrees with the sign of the market in 2 out of 3 occurrences of $(-2, 2)$. This shrinks the returns towards $0$ on most time instances, but will make the last return in the pattern rise. The optimal memory-2 strategy evolves the original market pattern into $[-1, 1, -1, 1, -1, 1, 3]$. The situation then repeats, and an optimal memory-2 strategy evolves the latter pattern into $[0, 0, 0, 0, 0, 0, 5]$. This is an example of how an optimal strategy creates market conditions which were not initially present.

We point out that the market bubble does not form with memory 3; see Figure 3.
We also point out that the formation of such bubbles is not an isolated phenomenon, nor specific to memory 2: we have generated several random markets and plotted their evolutions, and bubbles often arise with various memories. We report one such example in Figure 4 where a random pattern of length 20 is evolved by optimal memory-3 strategies into a bubble after 16 iterations. In fact, the framework proposed in this paper gives rise to a “game” in the spirit of Conway’s Game of Life, showing how simple trading rules can lead to apparently chaotic market dynamics. Perhaps the main difference between our game and Conway’s is that in ours strategies perform optimally within their resources.

3.2 High-memory strategies feed off low-memory ones

The next claim shows a pattern where a high-memory strategy can make a bigger profit after a low-memory strategy has acted and modified the market pattern. This profit is bigger than the profit that is obtainable by a high-memory strategy without the low-memory strategy acting beforehand, and even bigger than the profit obtainable after another high-memory strategy acts beforehand. Thus it is precisely the presence of low-memory strategies that creates opportunities for high-memory strategies which were not present initially. This example provides explanation for the real-life status quo which sees a growing quantitative sophistication among asset managers.

Informally, the proof of the claim exhibits a market with a certain “symmetry.” For high-memory strategies, the best choice is to maintain the symmetry by profiting in multiple points. But a low-memory strategy will be unable to do. Its optimal choice will be to “break the symmetry,” creating new profit opportunities for high-memory strategies.

Claim 7. For every \( m < m' \) there are a market pattern \( P = [X_1, \ldots, X_p] \), an optimal memory-\( m \) strategy \( s_m \) that evolves \( P \) into \( P_m \), and an optimal memory-\( m' \) strategy \( s_{m'} \) that evolves \( P \) into \( P_{m'} \) such that the gain of an optimal memory-\( m' \) strategy over \( P_m \) is bigger than either of the following:

- the gain of \( s_{m'} \) over \( P \),
- the gain of any memory-\( m \) strategy over \( P_m \),
- the gain of any memory-\( m' \) strategy over \( P_{m'} \).

Proof. Consider the market pattern

\[
\left[ X_1, \ldots, X_{m-1}, a \cdot X_m, b \prod_{i=1}^{m} X_i, Y_2, \ldots, Y_{m'-1}, c \cdot Y_{m'}, \prod_{i=2}^{m'} Y_i, X_1, \ldots, X_{m'-1}, a' \cdot X_{m'}, b \prod_{i=1}^{m'} X_i, Y_2, \ldots, Y_{m'-1}, c \cdot Y_{m'}, -\prod_{i=2}^{m} X_i \prod_{i=2}^{m'} Y_i \right],
\]

where \( X_1, \ldots, X_{m'} \) and \( Y_2, \ldots, Y_{m'} \) are independent \( \{-1, 1\} \) random variables (note some of these variables appear multiple times in the pattern), and \( a = 10, a' = 20, b = 30, c = 40. \)
We now analyze the gains of various strategies. For this it is convenient to define the latest input of a strategy \( s(x_1, \ldots, x_t) \) as \( x_t \); this corresponds to the most recent observation the strategy is taking into consideration.

The strategy \( s_m \): We note that there is an optimal strategy \( s_m \) on \( P \) that outputs 0 unless its latest input has absolute value \( a \). This is because in all other instances the random variable the strategy is trying to predict is independent from the previous \( m \). Thus, \( s_m(x_1, \ldots, x_m) = 0 \) unless \( |x_m| = a \) in which case it outputs the sign of \( \prod_{i=1}^{m} x_i \). The gain of this strategy is 1. The strategy evolves the pattern into the same pattern except that the first occurrence of \( b \) is replaced with \( b - 1 \):

\[
P_m := \left[ X_1, \ldots, X_{m-1}, a \cdot X_m, (b - 1) \prod_{i=1}^{m} X_i, Y_2, \ldots, Y_{m'-1}, c \cdot Y_{m'}; \prod_{i=1}^{m} X_i \prod_{i=2}^{m'} Y_i \right].
\]

The strategy \( s_{m'} \): We note that there is an optimal memory-\( m' \) strategy over \( P \) that outputs 0 unless its latest input has absolute value \( a, a' \), or \( c \). Again, this is because in all other instances the strategy is trying to predict a variable that is independent from the previous \( m' \). Moreover, when its latest input has absolute value \( c \), we can also assume that the strategy outputs 0. This is because the corresponding contribution is

\[
E \left[ \text{sign} \left( s_{m'}(b \cdot \prod_{i=1}^{m} X_i, Y_2, \ldots, Y_{m'-1}, c \cdot Y_{m'}; \prod_{i=1}^{m} X_i \prod_{i=2}^{m'} Y_i) \right) \right] + E \left[ \text{sign} \left( s_{m'}(b \cdot \prod_{i=1}^{m'} X_i, Y_2, \ldots, Y_{m'-1}, c \cdot Y_{m'}; \prod_{i=1}^{m} X_i \prod_{i=2}^{m'} Y_i) \right) \right] = 0.
\]

Thus there is an optimal memory-\( m' \) strategy with gain 2 that evolves the market pattern into the pattern \( P_{m'} \) that is like \( P \) with \( b \) replaced by \( b - 1 \) in both occurrences (as opposed to \( P_m \) which has \( b \) replaced by \( b - 1 \) only in the first occurrence).

The optimal memory-\( m' \) strategy on \( P_{m'} \). Since \( P_{m'} \) is like \( P \) with \( b \) replaced by \( b - 1 \), the gain of the optimal memory-\( m' \) strategy on \( P_{m'} \) is 2.

The optimal memory-\( m \) strategy on \( P_m \). Essentially the same argument for \( s_m \) can be applied again to argue that any memory-\( m \) strategy on \( P_m \) has gain 1 at most.

The optimal memory-\( m' \) strategy on \( P_m \). Finally, note that there is a memory-\( m' \) strategy whose gain is 4 on \( P_m \). This is because the replacement of the first occurrence of \( b \) in \( P \) with \( b - 1 \) allows a memory-\( m' \) strategy to predict the sign of the market correctly when its latest input has absolute value \( c \).

We make few final comments regarding Claim \[?] First, the same example can be obtained for deterministic market patterns by considering the exponentially longer market pattern where the random variables take all possible combinations. However the randomized example is easier to describe. Also, although in Claim \[?] we talk about a optimal strategy, in fact the
particular strategies considered there are natural in that they are those that minimize the number of \{-1, 1\} outputs (i.e., the amount of trading). We could have forced this to be the case by adding in our definition of gain a negligible transaction cost, in which case the claim would talk about the optimal strategy, but we preferred a simpler definition of gain.

4 Conclusion

In this work we have suggested to study market efficiency from a computational point of view. We have put forth a specific memory-based framework which is simple and tractable, yet capable of modeling market dynamics such as the formation of market bubbles and the possibility for a high-memory strategy to “feed off” low-memory ones. Our results may provide an analytical framework for studying the technological arms race that portfolio managers have been engaged in since the advent of organized financial markets.

Our framework also gives rise to a few technical questions, such as how many evolutions does it take a pattern to reach a certain other pattern, to what extent does this number of evolutions depends on different levels of memory, and to what extent does it depend on the simultaneous interaction among strategies using different memories.

References


