A MARTINGALE SYSTEM THEOREM FOR STOCK INVESTMENTS

Robert J. VANDERBEI

AT&T Bell Laboratories, Murray Hill, NJ 07974, USA

In this paper a proof is given that the dollar-cost-averaging investment strategy yields no advantage over any other non-clairvoyant strategy by showing that the difference between any two strategies is a mean-zero martingale. An interesting corollary of this theorem is that if X_t is a continuous positive martingale then the process $Y_t = X_t \int_0^t X_s \, ds - t$ is also a martingale.

martingales * system theorems * dollar cost averaging

1. Introduction

In this paper we consider an investment strategy called dollar-cost-averaging (DCA). It is well known that the argument in favor of DCA is flawed: an investor cannot get something for nothing. Probabilists have made this type of flaw precise with theorems like the optional sampling theorem or, more generally, the martingale system theorem (see e.g. [4]). It is interesting that the usual martingale system theorem is not formulated in a manner that is convenient for debunking DCA. The purpose of this paper is to formulate a new martingale system theorem which directly shows that DCA yields no advantage.

The story behind DCA goes as follows. Financial planners argue (see e.g. [2] or [5]) that investing a fixed small amount periodically in a given stock is superior to a single large investment in that stock. These financial planners, perhaps motivated by a larger percentage commission on small investments, argue that, in buying fixed small amounts frequently, the investor will benefit through buying relatively more when the price is low than when it is high. In this way, they claim that the investor can make money even when the share price does not exhibit an upward trend. This investment strategy is called *dollar-cost-averaging*.

We compare a general investment strategy to the mean performance of the stock. By introducing a time dependent scale factor to correct for any drift in the mean performance, we may without loss of generality assume that the mean performance would pay one dollar for every dollar invested no matter when the investment is liquidated. For this reason, we will refer to the mean performance as *pocketing the money* but the reader should realize that what we have in mind is a riskless investment that matches the mean performance of the stock investment (at least over some reasonable time frame). Of course in reality, there is no riskless investment that matches the performance of a risky investment. Our comparison to the mean performance merely serves as a benchmark to judge the relative merits of possible investment strategies.

In Section 3, we propose a better model and show that, on the average, no strategy has an edge. In Section 4, we show how to correct for trends so that the martingale theorem in Section 3 applies in general. Finally, in Section 5 we study the continuous version of the model introduced in Section 3.

2. Dollar-cost-averaging

Suppose that the share price of a certain stock is modeled by a stochastic process X_k, $k = 0, 1,$

2, ... If a dollars are invested every month in this stock, then the number of shares purchased in the k-th month is a/X_k and the total number of shares accumulated up through month n is $\sum_{k=0}^{n} a/X_k$. The value of this investment after the n-th month is obtained by multiplying this sum by the final share price X_n. Hence, the random variable

$$Y_n = X_n \sum_{k=0}^{n} \frac{a}{X_k} - a(n+1) \qquad (2.1)$$

denotes the difference between how we would do investing in this stock compared to simply pocketing a dollars for each of the $n+1$ months.

In order to say anything concrete, we need to give more details about the stochastic process X_k. One possibility is to assume that X_k is a simple random walk conditioned to return to its starting value x_0 at some fixed time, say n. This is the process that assigns equal probability to each of the feasible trajectories $\{x_k\}_{k=0,1,...,n}$: $x_n = x_0$ and $x_{k+1} = x_k \pm 1$ (for $k = 0, 1, ..., n-1$). If we assume that x_0 is sufficiently large (i.e., that $n < 2x_0$), then the feasible trajectories take only positive values. Let Ω denote the set of feasible trajectories. (Note that for Ω to be nonempty, it is necessary that n be an even integer.) The usual argument in favor of investing on a regular basis relies on this simple random walk model and is summarized by the following theorem.

Theorem 1. *If X_k is simple random walk conditioned to return to its starting point x_0 at time n, and n is an even integer satisfying $n < 2x_0$, then* $E[Y_n] > 0$.

Proof. For every sequence $\{x_k\}_{k=0,1,...,n}$ in Ω, let $\{\tilde{x}_k\}_{k=0,1,...,n}$ denote the 'reflected' sequence: $\tilde{x}_k = x_0 + (x_0 - x_k) = 2x_0 - x_k$. Let y_n be defined as in (2.1) using the sequence x_k and let $\tilde{y}_n$ be the analogous quantity obtained using the reflected sequence $\tilde{x}_k$. Then

$$\tfrac{1}{2}(y_n + \tilde{y}_n) = x_0 a \sum_{k=0}^{n} \frac{1}{2}\left(\frac{1}{x_k} + \frac{1}{\tilde{x}_k}\right) - a(n+1). \qquad (2.2)$$

Since the harmonic mean of a pair of numbers is smaller than the arithmetic mean, we see that

$$\left(\tfrac{1}{2}\left(x_k^{-1} + \tilde{x}_k^{-1}\right)\right)^{-1} \leqslant \tfrac{1}{2}(x_k + \tilde{x}_k) = x_0. \qquad (2.3)$$

Combining (2.2) and (2.3), we get

$$\tfrac{1}{2}(y_n + \tilde{y}_n) \geqslant 0$$

with equality holding if and only if $x_k = \tilde{x}_k$ for $k = 0, 1, ..., n$ (which is impossible for trajectories in Ω). The conclusion of the theorem now follows from the observation that Ω is invariant under reflection and that a trajectory and its reflection have the same probability. $\square$

The principle difficulty with this model is that we conditioned the process to return to its starting point at some future time. There is no way in real life that this can be guaranteed. In fact, if we knew that we were going to return to x_0 at some specified future time, then it would be possible to do even much better than DCA. Namely, invest only when X_k is less than x_0.

3. A more realistic model

In this section we introduce an improved model. Instead of a simple random walk conditioned to return to its starting point, we let X_k be a positive martingale. Roughly speaking, a martingale is a stochastic process which is 'fair'. That is, the conditional expectation of X_{k+1} given the values of $X_0, X_1, ..., X_k$ is simply equal to X_k. It is often convenient to give more information than simply past values of X_k. In general, one might want to condition on the entire past history of the universe. Let $\mathscr{F}_k$ denote a collection of random variables constituting all (relevant) history up to and including time k. The collection $\{\mathscr{F}_k\}_{k \geqslant 0}$ is called a *filtration*, and a stochastic process X_k is called a *martingale* adapted to the filtration $\{\mathscr{F}_k\}_{k \geqslant 0}$ if, for each k, X_k is a function of the random variables in $\mathscr{F}_k$ and

$$E[X_{k+1} \mid \mathscr{F}_k] = X_k$$

(see [3] for the definition and properties of conditional expectation). Associated with any martingale X_k is another stochastic process $\langle X \rangle_k$ that measures the variability of X_k. This process is called the *quadratic variation process* and is defined as

$$\langle X \rangle_k = \sum_{j=0}^{k-1} \left(E\left[X_{j+1}^2 \mid \mathscr{F}_j \right] - X_j^2 \right).$$

Instead of assuming that we invest a dollars each month, we let the size of the investment vary from month to month as a random variable A_k (not necessarily positive). The only assumption that we need to make is that A_k is adapted to the filtration $\{\mathscr{F}_k\}_{k \geqslant 0}$. This means that, for each k, A_k is a function of some or all of the random variables in $\mathscr{F}_k$. In this more general setting, the difference between the value of our stock investment at time n and the total number of dollars invested up through time n is

$$Y_n = \sum_{k=0}^{n} A_k \left(\frac{X_n}{X_k} - 1 \right) \tag{3.1}$$

and the total number of shares owned at time n is

$$S_n = \sum_{k=0}^{n} \frac{A_k}{X_k}. \tag{3.2}$$

The following theorem shows that no choice of investment streams A_k can give an advantage over simply pocketing the money.

Theorem 2. *The process Y_n is a mean-zero martingale whose quadratic variation is given by*

$$\langle Y \rangle_n = \sum_{j=0}^{n-1} \left(\langle X \rangle_{j+1} - \langle X \rangle_j \right) S_j^2. \tag{3.3}$$

Proof. First note that Y_n can be written as follows:

$$Y_n = \sum_{k=0}^{n-1} A_k \left(\frac{X_n}{X_k} - 1 \right). \tag{3.4}$$

Using (3.1) and (3.4) and the definition of S_n, we see that

$$Y_{n+1} - Y_n = (X_{n+1} - X_n) S_n.$$

Taking conditional expectations of both sides shows that Y_n is a martingale. Hence,

$$\begin{aligned}
E\left[Y_{n+1}^2 \mid \mathscr{F}_n \right] - Y_n^2 &= E\left[\left(Y_{n+1} - Y_n \right)^2 \mid \mathscr{F}_n \right] \\
&= E\left[(X_{n+1} - X_n)^2 \mid \mathscr{F}_n \right] S_n^2 \\
&= \left(E\left[X_{n+1}^2 \mid \mathscr{F}_n \right] - X_n^2 \right) S_n^2.
\end{aligned} \tag{3.5}$$

From the definition of quadratic variation, (3.5) implies that

$$\langle Y \rangle_{n+1} - \langle Y \rangle_n = \left(\langle X \rangle_{n+1} - \langle X \rangle_n \right) S_n^2.$$

Summing this formula gives us (3.3), thereby completing the proof. □

Remarks. (1) It is not necessary to assume that X_k is strictly positive as long as we make the extra stipulation that $X_k = 0$ implies $A_k = 0$ and we use the convention that zero times infinity is zero. This relaxation of the positivity assumption allows us to consider stocks of firms that can go bankrupt. The above stipulation says that we can't buy shares of a bankrupt stock. Of course, in real life once a stock price hits zero it stays there but the mathematics does not require this for Theorem 2 to hold. The mathematics also does not require nonnegativity of X_k, but it is hard to imagine a real-world interpretation for this level of generality.

(2) In the investment strategy model we should think of the quadratic variation as measuring the level of risk.

(3) The usual martingale system theorem says that if X_k is a martingale and B_k is adapted to the filtration $\{\mathscr{F}_k\}_{k \geqslant 0}$ then the process

$$Z_n = \sum_{k=0}^{n-1} B_k (X_{k+1} - X_k)$$

is a martingale. The fact that Y_n is a martingale can be derived from this result by taking

$$B_k = \sum_{j=0}^{k} \frac{A_j}{X_j},$$

however this seems like a round-about way of proving it.

4. Adjusting for trends

In this section we show that the assumption that the share price process is a martingale actually entails no loss of generality. To start, let X_k be an arbitrary positive stochastic process representing the share price. Put

$$\tilde{X}_k = R_k X_k,$$

where R_k is the *discount factor* given by

$$R_k = \prod_{j=0}^{k-1} \frac{X_j}{E\left[X_{j+1} \mid \mathscr{F}_j \right]}.$$

Note that R_k is a function of some or all of the random variables in $\mathscr{F}_{k-1}$ and that

$$R_{k+1} = R_k \frac{X_k}{E\left[X_{k+1} \mid \mathscr{F}_k \right]}.$$

The factor R_k is called the discount factor because its application turns the share price X_k into a martingale $\tilde{X}_k$:

$$E[\tilde{X}_{k+1} \mid \mathcal{F}_k] = R_{k+1}E[X_{k+1} \mid \mathcal{F}_k]$$
$$= R_k X_k$$
$$= \tilde{X}_k.$$

At time n, the value of our investment is $X_n S_n$. Hence, the present value of the investment is $R_n X_n S_n$. The present value of the actual investment stream is $\sum_{k=0}^n R_k A_k$. We are interested in the difference between these two present values:

$$Y_n = R_n X_n S_n - \sum_{k=0}^n R_k A_k$$
$$= \sum_{k=0}^n \left(R_n X_n \frac{A_k}{X_k} - R_k A_k \right)$$
$$= \sum_{k=0}^n \tilde{A}_k \left(\frac{\tilde{X}_n}{\tilde{X}_k} - 1 \right),$$

where $\tilde{A}_k = R_k A_k$ represents the present value of the dollar investment at time k. This last formula is in the form that we studied in the previous section.

5. A continuous version

In this section we assume that the reader is familiar with Ito calculus for continuous parameter semi-martingales (see e.g. [1]).

Let X_t be a positive continuous local martingale representing the share price at time t, let B_t be a semi-martingale representing the number of dollars invested in the stock at time t and put

$$Y_t = X_t S_t - B_t,$$

where

$$S_t = \int_0^t \frac{d B_u}{X_u}$$

represents the total number of shares owned at time t. Note that B_t represents the total dollar investment at time t and so it is the continuous analog of $\sum_{k=0}^n A_k$. Since B_t is a semi-martingale, we see that we are allowing the investor to buy and sell shares at a very fast pace.

Theorem 3. *The value Y_t is given by*

$$Y_t = \int_0^t S_u \, d X_u + \langle X, S \rangle_t.$$

If B_t is locally of bounded variation, then Y_t is a continuous local martingale with quadratic variation

$$\langle Y \rangle_t = \int_0^t S_u^2 \, d\langle X \rangle_u. \tag{5.1}$$

Proof. Applying the integration by parts formula from Ito calculus (see e.g. [1], p.68), we see that

$$X_t S_t = \int_0^t S_u \, d X_u + \int_0^t X_u \, d S_u + \langle X, S \rangle_t.$$

From the definition of S_t and the associative law for stochastic integrals, we get

$$\int_0^t X_u \, d S_u = B_t.$$

Therefore,

$$Y_t = \int_0^t S_u \, d X_u + \langle X, S \rangle_t.$$

If B_t locally has bounded variation, then S_t also has this property and so the joint variation $\langle X, S \rangle_t$ vanishes. Hence,

$$Y_t = \int_0^t S_u \, d X_u. \tag{5.2}$$

Since S_t is continuous, it is predictable and hence the stochastic integral is a martingale. Formula (5.1) follows from (5.2) (see e.g. [1], p.56). This completes the proof. □

Remarks. (1) Choosing $B_t = t$, the definition of Y_t reduces to the formula given in the abstract.

(2) If we restrict our attention to the case where shares are purchased but not sold, then B_t is increasing and hence is locally of bounded variation.

(3) If B_t has unbounded variation, then in general $\langle X, S \rangle_t$ will not vanish. This means that if we are able to vary our investment stream fast enough, then we can potentially make the value of our investment into a submartingale.

Example. Suppose that $B_t = X_t$. Then, using Ito's formula, we see that

$$S_t = \ln(X_t) - \ln(X_0) + \frac{1}{2} \int_0^t \frac{d\langle X \rangle_u}{X_u^2}.$$

Hence,

$$\langle X, S \rangle_t = \langle X, \ln(X) \rangle_t.$$

Since $\langle X, \ln(X) \rangle_t$ is obviously increasing, the process Y_t is a submartingale. The economic interpretation is simple. If we could actually react so fast to the market that we could make our dollar investment match the share price of the stock, then we would be in the enviable position of having a large amount invested when the price is high and a small amount invested when the price is low. This anomaly is an artifact of the continuous time model and as we showed in the previous section this behavior is impossible in discrete time.

References

[1] R. Durrett, *Brownian Motion and Martingale in Analysis*, Wadsworth, Belmont, CA, 1984.

[2] L. Engel, *How to buy Stocks*, 6th edition, Little-Brown, Boston, MA, 1976.

[3] S. Karlin and H. Taylor, *First Course in Stochastic Processes*, 2nd edition, Academic Press, New York, 1975.

[4] J. Lamperti, *Stochastic Processes: A Survey of the Mathematical Theory*, Springer-Verlag, Berlin, 1977.

[5] J. McGee, *Lasser's Personal Financial Planner 1988*, Simon and Shuster, New York, 1987.