

Lecture IX: Market Efficiency Models, the Capital Asset Pricing Model and Joint Test of the CAPM and Market Efficiency

Charles B. Moss

November 16, 2013

I. Tests of Market Efficiency

A. Fama, E.F. and M. Blume. 1966. Filter Rules and Stock Market Trading. *Journal of Business* 39(1), 226-241.

1. The basic notion behind this article was to test for stock market efficiency by comparing the profitability of x -percent rules with the profitability of buy-and-hold strategies.
 - a. Under the random walk theory changes in price are uncorrelated with zero mean

$$x_t = x_{t-1} + \epsilon_t \tag{1}$$

Thus, past prices contain no information about future price movements that could be used to increase profits.

- b. The concept of a random walk was typically tested using statistical procedures discussed later. One exception was the work by Alexander who tested for market efficiency using a mechanical trading rule. The specific rule was an x -percent filter: "If the daily closing price of a particular security moves up at least x percent, buy and hold the security until its price moves down by at least x percent from a subsequent high, at which time simultaneously sell and go short. The short position is maintained until the

- daily closing price rises at least x percent above a subsequent low at which time one covers and buys.”
- c. Alexander reported the results for trading filters ranging between 5 and 50 percent. "In general, filters of all different sizes and for all different time periods yielded substantial profits – indeed profits significantly greater than those of simple buy-and-hold policy. This led Alexander to conclude that the independence assumption of the random walk model was not upheld by his data.”
2. Fama applies Alexander’s filter rules to individual securities on the Dow-Jones Industrial Average from January 1956/April 1958 to September 26, 1962. He generated twenty four different filters ranging from 0.5 percent to 50.0 percent.
 - a. "When commissions are taken into account the largest profits under the filter technique are those of the broker.”
 - b. Only four securities had positive average profits per filter after commissions.
 - c. When commissions were excluded, the filter did better, but the returns were still not as large as buying-and-holding. The difference between the results is postulated to be the difference in dividends:
 - Using closing prices alone, Alexander failed to adjust for the effect of dividends.
 - The critical point is that dividends declared and paid when the stock is short are the responsibility of the investor.
- B. Theil, H. and C.T. Leenders. 1965. Tomorrow on the Amsterdam Stock Exchange. *Journal of Business* 38(3), 277-284.
1. The point or concept tested in this paper is also slightly different than the traditional random walk hypothesis. Specifically, the goal is to use past information to predict the proportion of stocks that will advance on tomorrow’s trading.
 2. To accomplish this they use a variant of information theory.
 - a. Assume a game where a check is hidden somewhere on a checker board. Suppose that you could ask a series of true or false questions at a positive price. The problem

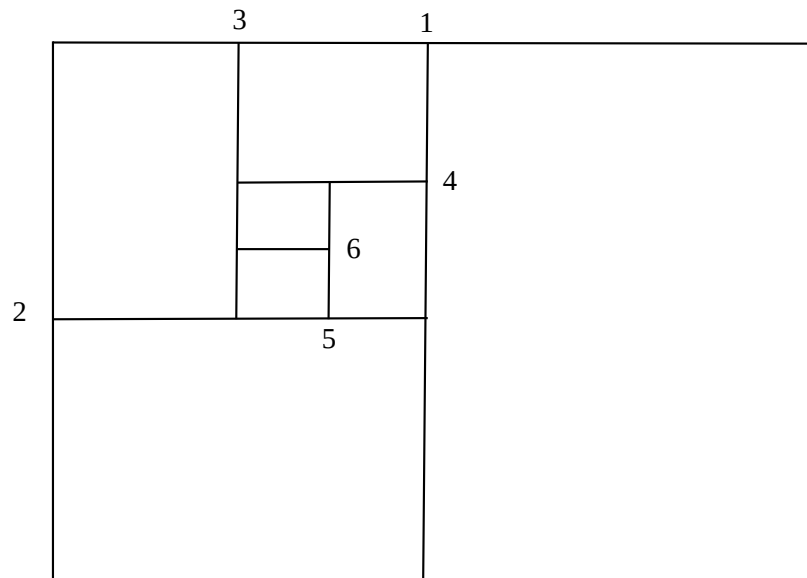


Figure 1: Minimum Cost Search Pattern

then becomes how many questions would you have to ask? Specifically, how do you get the most information from each question?

- b. The greatest amount of information is gained by halving the board each time as depicted in Figure 1. In the case of a chess board we require six questions $64 = 2^6$.
- c. In general, if there are n equi-probable alternatives then H bits of information are required

$$n = 2^H \text{ or } H = \log_2 n \quad (2)$$

- d. $\log_2 n$ implies

$$H = \log_2 \frac{1}{1/n} \quad (3)$$

which shows that there is a relationship between uncertainty and information. However, we have assumed that each outcome is equally likely. Generalizing the formula-

tion, if outcomes are not equally likely the information in outcome i is

$$\log_2 \frac{1}{0.80}, \log_2 \frac{1}{0.20} \rightarrow \text{warped coin} \quad (4)$$

- e. This leads us to the average value of information defined in this case as

$$0.80 \log_2 \frac{1}{0.80} + 0.20 \log_2 \frac{1}{0.20} = 0.722 \text{ bits} \quad (5)$$

- f. Generalizing this concept, assume that there are n mutually exclusive events $\{E_1, E_2, \dots, E_n\}$ with corresponding probabilities $\{p_1, p_2, \dots, p_n\}$. If E_i occurs, the informational content of that signal is $\log_2 (1/p_i)$. The expected informational content of the signal is

$$\sum_{i=1}^n p_i \log_2 \left(\frac{1}{p_i} \right) = - \sum_{i=1}^n p_i \log_2 p_i \quad (6)$$

- f. Further generalizing, suppose that a message does not tell us directly which of the n events was realized, merely that the odds have changed.
- The original probabilities (prior probabilities) p_i are changed to q_i (the posterior probabilities).
 - Question: How do we measure the informational content of such an indirect message?

$$I(q; p) = \sum_{i=1}^n q_i \log_2 \left(\frac{q_i}{p_i} \right) \quad (7)$$

3. "Suppose that it is our task to predict a number of probabilities, such as the fractions of stocks advancing, declining, and remaining constant on tomorrow's Stock Exchange. Let us write $p_1, \dots, p_n$ for these forecasts. At some point in time we shall know what the probabilities are that were actually realized. These are $q_1, \dots, q_n$. Question: 'How do we measure the accuracy of the forecasts, given the realizations as observed afterwards?' ... If the observed probabilities (q_i) supply us with much information, given the forecasts made

(p_i) , we regard the forecasts as highly inaccurate.... Going back to equation [7] we see that $I(q;p)$ is a natural measure, based on information theory, to measure the inaccuracy of the forecasts.”

4. A straightforward application of the information equation yields

$$\sum_{i=1}^3 q_{it} \log_2 \left(\frac{q_{it}}{p_{it}} \right) \quad (8)$$

where p_{it} is the forecast of the i^{th} fraction on the t^{th} day. This is the informational inaccuracy for a given day. The long-run informational inaccuracy is

$$\frac{1}{T} \sum_{t=1}^T \sum_{i=1}^3 q_{it} \log_2 \left(\frac{q_{it}}{p_{it}} \right) \quad (9)$$

- a. Different rules

- Long-run $p_{it} = \bar{q}_i$ $I = 0.0689$.
- Autorgressive $p_{it} = q_{i,t-1}$ $I = 0.0774$.
- Some mix as presented in Table 1. The optimal being $\alpha^* = 0.477$ implying an inaccuracy measure of 22% less than the long-run average.

$$p_{it} = \alpha q_{i,t-1} + (1 - \alpha) \bar{q}_i \quad (10)$$

Table 1: Information Inequality for Mixed Model

α	I
0.00	0.0689
0.30	0.0557
0.45	0.0536
0.47	0.0535
0.49	0.0535
0.51	0.0536
0.70	0.0572
1.00	0.0774

Table 2: Weighted Sum of Digits

N	α	I
1	0.477	0.0535
2	0.527	0.0544
3	0.551	0.0556
4	0.564	0.0567
5	0.578	0.0572

· Another generalization is

$$p_{it} = \alpha \frac{Nq_{i,t-1} + (N-1)q_{i,t-2} + \cdots q_{i,t-N}}{\frac{1}{2}N(N+1)} + (1-\alpha)\bar{q}_i \quad (11)$$

the results for this formulation are presented in Table 2.

5. "Our conclusion is, therefore, that the Amsterdam Stock Exchange has a memory of one day as regards to the fractions of stocks advancing in price or declining or remaining unchanged."

C. Granger, C.W.J. and O. Morgenstern 1963. Spectral Analysis of New York Stock Market Prices. *Kyklos* 16(1), 1-27.

1. This study approaches the problem from a statistical time series perspective

$$x_t - x_{t-1} = \epsilon_t$$

$$E[\epsilon_t] = 0 \quad (12)$$

$$E[\epsilon_t \epsilon_s] = \begin{cases} \sigma_\epsilon^2, & t = s \\ 0, & t \neq s \end{cases}$$

2. Quick explanation of the Fourier representation.

a. Frequency is proportional to the inverse of the "period."
If a component has a period of k weeks, the frequency will be $2\pi/k$.

- b. The concept is to break the time path of a sequence of data points into the sum of Fourier terms of different periods

$$f(t; \alpha, \beta) = \alpha_0 + \left[\sum_{k=1}^K \alpha_k \sin(\omega_k t) + \beta_k \cos(\omega_k t) \right] \quad (13)$$

for some series of frequencies ω_k , $k = 1, \dots, K$.

- c. In general, the objective is to determine the variance in the series explained by certain frequencies. If the frequency band $(\omega, \omega + d\omega)$ contributes $f(\omega) d\omega$ to the total variance, we call the function $f(\omega)$ the **power spectrum of the series**. A large value of the estimated $f(\omega)$ indicates that the frequencies of this band are of particular importance as they contribute a large amount of the total variance.
- In general, the economic time series have a spectrum of the shape depicted in Figure 2. The longer the period component, the larger the amplitude.
 - The spectrum of independent sequence terms (white noise) is a horizontal line parallel to the x -axis.
 - The section of the power spectrum of most interest to economists is the low-frequency range, within which all the "long-run" components are concentrated.
3. Results of the Analysis – "In every case the resulting spectrum was very flat over the whole frequency range.... The initial result was that a simple random walk model fits the data fairly well as very few points are outside the 95 percent confidence limits."

II. Market Equilibrium and the Capital Asset Pricing Model (CAPM)

- A. In the preceding section we addressed the question of market efficiency by looking at the value of information or the presence of profit opportunities through historical information. This effort takes another tack, looking at equilibrium from the price of risk perspective. Specifically, the CAPM is formulated based on the risk-return tradeoff.

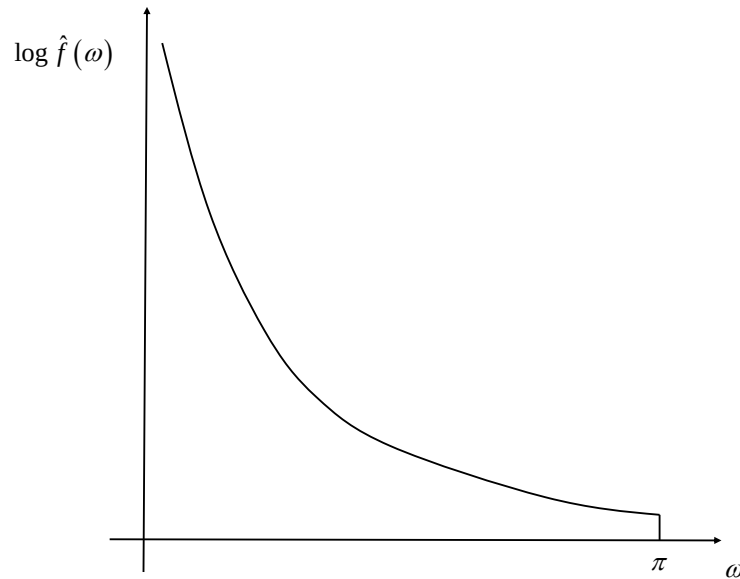


Figure 2: Time Series Spectrum

- B. "The CAPM is developed in a hypothetical world where the following assumptions are made about investors and the opportunity set:
1. Investors are risk-averse individuals who maximize the expected utility of their end of period wealth.
 2. Investors are price takers and have homogeneous expectations about asset returns that have a joint normal distribution.
 3. There exists a risk-free asset such that investors may borrow or lend unlimited amounts at the risk-free rate.
 4. The quantities of assets are fixed. Also, assets are marketable and perfectly divisible.
 5. Asset markets are frictionless and information is costless and simultaneously available to all investors.
 6. There are no market imperfections such as taxes, regulations, or restrictions on short-selling."
- C. Proof of the CAPM requires that in equilibrium, the market portfolio will be some efficient portfolio on the top of the minimum

variance set. This is established by homogeneous expectations which makes the minimum-variance set the same for all investors. Next, the assumption of a risk-free asset along with the assumption of expected utility maximizers fixes the same "market portfolio" of risk stocks for all individuals.

1. In the $E - S$ frontier we examined the case of a risk-free asset. Specifically, given a risk-free asset there was exactly one efficient combination of risk assets

$$\underline{w}_t = \frac{\Sigma^{-1}(\bar{z} - R\underline{1})}{B - AR} \quad (14)$$

All other efficient mean/standard deviation points could be derived by lending or leveraging the single risky portfolio at the risk-free rate along the Capital Market Line.

2. The covariance between any individual asset returns and the efficient portfolio was then shown to be proportional to the covariance of the individual asset returns and the efficient portfolio returns

$$\begin{aligned} \varrho_t &= \Sigma \underline{w}_t = \Sigma \left[\frac{\Sigma^{-1}(\bar{z} - R\underline{1})}{B - AR} \right] = \frac{(\bar{z} - R\underline{1})}{[B - AR]} \\ \sigma_t^2 &= \underline{w}_t' \varrho_t = \underline{w}_t' \left[\frac{\bar{z} - R\underline{1}}{B - AR} \right] \\ &= \frac{\underline{w}_t' \bar{z} - R \underline{w}_t' \underline{1}}{B - AR} \\ \Rightarrow \sigma_t^2 &= \frac{\bar{z}_t - R}{B - AR} \\ B - AR &= \frac{\bar{z}_t - R}{\sigma_t^2} \\ \varrho_t &= \frac{\bar{z} - R\underline{1}}{\left[\frac{\bar{z}_t - R}{\sigma_t^2} \right]} \end{aligned} \quad (15)$$

$$\Rightarrow [\bar{z}_t - R] \frac{\mathcal{Q}_t}{\sigma_t^2} = \bar{z} - R\underline{\beta} \quad (16)$$

Denoting $\underline{\beta} = \mathcal{Q}_t / \sigma_t^2$

$$\begin{aligned} [\bar{z}_t - R] \underline{\beta} &= \bar{z} - R\underline{\beta} \\ \bar{z}_i - R &= \beta_i [\bar{z}_t - R] \end{aligned} \quad (17)$$