Time-variation in market efficiency: a mixture-of-distributions approach

John P. Hussman, Ph.D.

April 1998

Abstract

Market efficiency implies a positive relationship between market risk and expected return. Using ex-ante conditioning variables implied by the definition of total return, the time series of U.S. stock market data is partitioned into a set of conditional distributions. These distributions suggest that market risk is not universally efficient, but displays varying degrees of efficiency in generating expected return. The conditional distributions also exhibit differences in the expected mix of positive and negative outlier returns, Sharpe ratios, maximal investment loss, and business cycle characteristics. A broader definition of efficiency compatible with these results is discussed.

Introduction

The positive relationship between expected return and risk is among the most fundamental and widely accepted premises of finance. An important qualification is that this premise applies not to all financial risks, but strictly to "efficient" risks. Many financial risks are "inefficient" and can be diversified away or systematically avoided without compromising expected return. For example, the Capital Asset Pricing Model gives the restrictions on individual security returns that are necessary for the market portfolio to be the single efficient portfolio. Under these restrictions, market risk is compensated linearly by expected return, while no amount of diversifiable risk is compensated (Sharpe [1964], Lintner [1965]). Increased risk-bearing as a strategy to increase expected portfolio return is unreasonable unless the portfolio is already efficient in terms of expected return per unit of risk.

In finance literature, the term "market efficiency" applies to two wholly different concepts. In cross-sectional research, the market is efficient if no alternate portfolio of equal risk offers a higher expected return in the presence of risk-free borrowing and lending. In time series research, the market is efficient if price is a sufficient statistic for all current information about future asset returns (Grossman [1989]), so that no set of conditioning information can be used to earn abnormal risk-adjusted returns over time. Clearly, efficiency of the market according to the cross-sectional definition does not imply efficiency according to the time-series definition, nor vice versa. The time-series definition implies that the ex-ante Sharpe ratio is constant across periods. That is, market returns (in excess of the risk-free rate) must be drawn from a single probability distribution that cannot be partitioned ex-ante using conditioning information. Otherwise, abnormal risk-adjusted returns may be achieved by using leverage during periods with high Sharpe ratios, and avoiding market risk during periods with low Sharpe ratios.

The main assertion of this paper is that market risk is not universally efficient. The time-series of market returns is not drawn from a single probability distribution, but rather from a mixture of "conditional distributions" with varying degrees of efficiency in generating expected return. Moreover, by exploiting the definition of total return, a small amount of conditioning information may be found to partition these distributions ex-ante. At least one of these distributions is associated with expected returns that are predictably below the risk-free rate of return. The avoidance of risk in such "inefficient" distributions may increase expected return over time without an increase in volatility or maximal loss. This is particularly true if leverage is selectively used in those distributions with unusually high expected return per unit of risk. This evidence suggests that the market is not "efficient", in the sense that conditional risk-adjusted returns exhibit wide and predictable variation over time. However, the practical impact of this inefficiency depends on assumptions regarding transaction costs, tax treatment and differences in borrowing and lending rates.

It is important to note at the outset that these results do not necessarily imply inefficiency in a broader economic sense. If risk aversion is time varying, or investors face variation in the "insurance" properties of equities (i.e., time variation in the covariance of investment returns with consumption risks), then standard deviation or maximal loss are incomplete measures of the risk that investors subjectively perceive. For example, standard deviation may be low in some periods, but the subjective disutility of equity risk to investors may be high, leading investors to demand a high expected return. An investor having constant risk aversion and a lack of insurance motives would then earn high expected returns by bearing risk that other market participants find highly aversive. In this sense, these returns could be viewed as efficient in the broader sense of being compensation for the provision of scarce resources.

1. Return and risk of the market portfolio over time

If the market portfolio is efficient in a time-series sense and the data set is sufficiently large, one might expect to observe higher average returns during periods of higher volatility. Of course, there is no assurance of this, as efficiency is an ex-ante concept. It is notable, however, that monthly data from 1946 through 1997 fails to accept the hypothesis of a positive relationship between return and volatility risk. Indeed, the estimated return-risk relationship is negative.

Throughout this paper, the S&P 500 Index will be used as a proxy for the market portfolio. Using data for the S&P 500 Index, the average monthly total return in excess of the three-month Treasury bill yield was calculated for non-overlapping annual periods, with the corresponding standard deviation of monthly returns. Ordinary least squares regression produced the following relationship:

$$\text{Average monthly excess return} = 1.634 - 0.303 \text{ Standard deviation} \qquad (t = -2.307) \qquad (1)$$

On average, periods of high market volatility have been associated with below-average returns. While these periods have clearly offered unattractive return-risk characteristics ex-post, efficiency is properly defined ex-ante. The relevant question is whether any set of conditioning information might have been capable of partitioning favorable and unfavorable return-risk climates in advance. This is the focus of the next section.

2. Partitioning conditional return-risk distributions using total return factors

The total return on the market portfolio is determined by three factors: the level of the dividend yield, the change in the yield, and the growth rate of dividends. The dividend yield determines the income component of total return. The capital gain over any holding period can be calculated from the beginning and terminal dividend yields, and the rate of intervening dividend growth. In periods when the dividend yield falls, capital gains exceed the rate of dividend growth. In periods when the dividend yield rises, capital gains fall short of the rate of dividend growth.

Historically, variation in the rate of dividend growth on the S&P 500 Index has not been a significant source of variation in capital gains. From 1946 to 1997, the standard deviation of annual capital gains on the S&P 500 Index was 14.91% annually. The standard deviation of dividend growth was 5.73%. Subtracting dividend growth from capital gains to obtain the gain or loss attributable to changes in the dividend yield, the standard deviation was 15.91%. The smaller variability of capital gain itself arises because dividend growth and capital gains attributable to yield variation demonstrate a slight negative correlation in the data.

Due to the minor influence of dividend growth variations, it is evident that variation in total returns may be characterized primarily by the two remaining factors: the level of dividend yield, and its trend. Unfortunately, the future trend of the dividend yield is not known ex-ante, so a proxy must be obtained for this factor. While a wide range of proxies is possible, the prevailing trends in the dividend yield and closely competing yields such as interest rates are natural candidates.

The analysis of total return therefore suggests two dimensions by which conditional expected return distributions may be partitioned. These two dimensions yield four "distributions" which may be determined ex-ante, based on whether these dimensions are individually graded as "favorable" or "unfavorable".

The first dimension is the level of the dividend yield on the S&P 500 Index, which is directly observable. For simplicity, yield levels above the 1946-1997 median of 3.66% are graded as "favorable" on this dimension, while yield levels of 3.66% or below are graded as "unfavorable". While it is true that this exact median was not available ex-ante, the results below are robust to alternative criteria such as arbitrary moving averages of past yields.

The second dimension is the yield trend, which is proxied by the uniformity of trends in the following three variables: the dividend yield on the S&P 500 Index, the yield on three-month Treasury bills, and the yield on the ten-year Treasury bond. An odd number of yields is chosen to obtain a simple measure of trend uniformity. If at least two of these yields are below their levels of 26 weeks earlier, the trend dimension is graded as "favorable". Otherwise, trends are graded as "unfavorable".

3. Characteristics of conditional return-risk distributions

Using weekly closing prices and yields from 1946 to 1997, the data was separated into four conditional distributions based on the yield level and trend criteria described in section 3, and returns were computed using data for the following week. The conditional distributions are defined as follows:

Distribution I:	yield level favorable, yield trend favorable
Distribution II:	yield level unfavorable, yield trend favorable
Distribution III:	yield level favorable, yield trend unfavorable
Distribution IV:	yield level unfavorable, yield trend unfavorable.

Table 1 presents the average performance of the market by partition, for the entire span from 1946-1997 and for the split samples 1946-1971 and 1972-1997.

TABLE 1

CHARACTERISTICS OF CONDITIONAL RETURN-RISK DISTRIBUTIONS

1946-1997	Annualized Return	Frequency	Excess Return	Standard Deviation	Sharpe
Distribution I	33.59%	.174	0.4656%	1.90%	0.246
Distribution II	14.56%	.261	0.1752%	1.57%	0.112
Distribution III	8.99%	.326	0.0604%	2.17%	0.028
Distribution IV	1.41%	.239	-0.0676%	1.81%	-0.037

1946-1971	Annualized Return	Frequency	Excess Return	Standard Deviation	Sharpe
Distribution I	27.13%	.162	0.4360%	1.68%	0.260
Distribution II	15.27%	.192	0.2054%	1.31%	0.157
Distribution III	13.07%	.346	0.1969%	1.95%	0.101
Distribution IV	0.71%	.300	-0.0668%	1.54%	-0.043

1972-1997	Annualized Return	Frequency	Excess Return	Standard Deviation	Sharpe
Distribution I	39.47%	.187	0.4913%	2.07%	0.238
Distribution II	14.15%	.330	0.1579%	1.70%	0.086
Distribution III	4.55%	.306	-0.0939%	2.39%	-0.039
Distribution IV	2.62%	.187	-0.0689%	2.20%	-0.031

Annualized return is the compound annualized total return for each distribution. Frequency is the proportion of data points assigned to each distribution. Standard deviation is the weekly standard deviation of returns. Excess return is the compound weekly average return in excess of the three-month Treasury bill yield. Sharpe ratio is the ratio of the weekly excess return to the weekly standard deviation of returns.

a) Rank Sums

The conditional distributions vary significantly in their return-risk profiles, yet are highly similar across sub-periods. The hypothesis that the returns in these four partitions are drawn from a single probability distribution is strongly rejected. The standard *F* test for equality of sample means is not valid in this case, as variances in each distribution are unequal and standard normality assumptions are not ensured for financial data. While the test statistic F(3,2710) = 7.77 is extremely high, the statistic may not be *F* distributed. The nonparametric Kruskal-Wallis H test requires no assumptions about population probability distributions, and is based on rank-sums for returns in each distribution. This test yields H = 18.92 and is distributed $\chi^2(3)$. This statistic significantly exceeds the critical value of χ^2 even at the 0.005 level. Values for the first and second sub-periods also indicate significant difference across distributions, with H values of 12.79 and 9.16, respectively. Test statistics using returns in excess of the risk-free rate were even more extreme, with H values of 19.46, 16.02 and 9.93 for the full sample, the first sub-period and the second sub-period, respectively.

b) Sharpe Ratios

The possibility of achieving abnormal risk-adjusted returns largely rests on variation in Sharpe ratios across conditional distributions. Such variation is evident in Table 1. Distribution I is consistently associated with the highest annualized return and Sharpe ratio. Distribution II consistently exhibits the lowest standard deviation, and is also associated with relatively high returns despite an unfavorably low dividend yield. Distribution III consistently exhibits the highest standard deviation, but the return profile differs across sub-periods. During the 1946-1971 period, the Sharpe ratio was favorable at 0.101, but the 1972-1997 period exhibits a negative Sharpe ratio. This latter period captures much of the persistent rise in interest rates during the 1970's. The level of rates was sufficient to eliminate the excess return from holding stocks in this particular climate. Distribution IV generates a consistently negative Sharpe ratio, suggesting that risk taken in this conditional distribution has historically been uncompensated by return in excess of the risk-free rate.

c) Maximal Loss

Maximal loss or "drawdown" reflects the deepest peak-to-trough decline in portfolio value over a given holding period. Maximal loss in each distribution can be computed under the assumption that equities are held for an additional week as long as the market remains in a given conditional distribution, and that the return is zero during intervening periods. For the S&P 500, the maximal loss in weekly data was -44.64% during the period 1946-1997. This loss was registered during the bear market decline of 1973-1974. For Distribution I, maximal loss was modest at -16.67%. Distribution II exhibited a maximal loss of -16.92%. Distribution III exhibited a significantly deeper maximal loss of -32.08%. The deepest loss was recorded for Distribution IV, with a maximal loss of -51.35%. The reason this figure is deeper than the maximal loss for the S&P 500 is that while the index recovered frequently to new highs, the tendency to do so while in Distribution IV was relatively weak. Thus, entry into Distribution IV after an intervening period was capable of extending the maximal loss for this climate, even though the S&P 500 itself might have recovered during that intervening period. The performance of the market in Distribution IV is therefore not measured only by a slightly negative Sharpe ratio, but by a tendency to sustain severe interim losses.

d) Outliers

A frequent argument against active asset allocation is based on the notion that an investor missing even a small number of the best market periods would have underperformed Treasury bills over time. The frequent counter-argument is that an investor missing even a small number of the worst periods would have significantly outperformed a buy-and-hold strategy over time. The conditional distributions presented here are useful in evaluating this debate, as the occurrence of positive or negative outliers is not uniform across distributions.

Weekly market returns from 1946 through 1997 were sorted, noting the conditional distribution in effect at the end of the prior week. The 30 most positive and negative outlier returns were examined (30 observations were chosen to ensure a minimum expected frequency of five in each distribution). Of the top 30 weeks, 10 of these instances were allocated to Distribution I, 4 to Distribution II, 13 to Distribution III, and 3 to Distribution IV. This represents a significant difference from the relative frequencies of these distributions. The standard chi-square test yields a value of 9.70, which is distributed $\chi^2(3)$ and is significant at the 0.025 level. Clearly, the majority of the top market weeks occurred during periods of above-median dividend yield.

In contrast, of the 30 most negative market weeks, only 2 were allocated to Distribution I, 4 to Distribution II, 18 to Distribution III, and 6 to Distribution 4. Again, this profile differs from the relative frequencies of these distributions in a highly significant manner ($\chi^2(3) = 10.96$). Notably, while the top 30 weeks occurred primarily in environments of above-median dividend yield (Distributions I and III), the most negative 30 weeks were not distinguished primarily by low yield, but rather by an unfavorable trend condition (Distributions III and IV). The most extreme losing weeks were dominant in periods when the trend condition of the market had already deteriorated measurably. Distribution III held more extreme weeks, both positive and negative, than any other distribution, which is consistent with the high standard deviation of returns observed in the data.

e) Business Cycle Variation

It is notable that a significant proportion of Distribution I observations occur in the latter half of recessionary periods (when these recessions are likely to be widely recognized). Based on official NBER U.S. recession dates, 23.3% of Distribution I observations occurred in the latter half of U.S. recessions, compared with 3.9% of Distribution II, 8.0% of Distribution III, and none of Distribution IV. In contrast, only 17.3% of Distribution I observations occurred in the latter half of U.S. expansions, compared with 44.1% of Distribution II, 43.3% of Distribution III, and 60.8% of Distribution IV. Stated differently, Distribution I was active during 52.6% of late-recessions, but only 7.0% of late-expansions. This pattern suggests that the high average returns observed in Distribution I may represent compensation for perceived risks which are not completely characterized by standard deviation alone.

4. Asset Allocation

The significant differences in return and risk characteristics across conditional distributions suggests a role for active asset allocation. However, the returns from an active approach are dependent on assumptions regarding trading costs and taxation. The most obvious strategy is simply to hold the risk-free security during periods characterized by Distribution IV, where the Sharpe ratio is negative. In the absence of trading costs, such a strategy is associated with an annualized return of 13.28% from 1946 through 1997 (0.24% weekly), and a weekly standard deviation of 1.66%. This compares with a passive return of 12.35% for the S&P 500 Index (0.22% weekly), and a weekly standard deviation of 1.88%. Transaction costs in S&P 500 Index futures generally amount to a small fraction of one percent. Clearly, however, the foregoing return advantage could be eroded by sufficiently high transaction costs. Assuming a high cost of fully 1% per transaction, the annualized return declines to 10.98%, with a weekly standard deviation of 1.67%.

It is well known that under constant relative risk aversion, the single-period investment demand function has the form:

$$w = [\ ER_m - R_f\ ] / A\sigma^2_m \qquad (2)$$

where w is the proportion of funds allocated to the market portfolio, ER_m is the expected return of the market portfolio, R_f is the risk free rate, σ^2_m is the variance of the market portfolio, and A is the coefficient of relative risk aversion. Merton [1969] demonstrates that when expected returns are stochastic and investors maximize the expected utility of terminal wealth, the demand function generally lacks an analytical solution, and includes an additional term which represents "hedging demand". This optimal control problem requires evaluation of the investor's value function using a guess-and-verify method, iteration, or numerical solution (see e.g. Brennan, Schwartz and Lagnado, [1995]), and is beyond the scope of this paper. For analytical simplicity, the demand function in equation (2) will be used to obtain single-period or "myopic" asset allocation weights.

Using the data in Table 1, it is straightforward to calculate the values for $[\ ER_m - R_f\ ] / \sigma^2_m$ for each distribution. The value of A is then chosen to equate the standard deviation of returns for the active strategy with the standard deviation of returns for a passive investment in the market portfolio.

In the absence of transaction costs, the risk-equating value of coefficient A is 6.28. This implies percentage allocations w for Distributions I through IV of 206.3%, 20.5%, 113.8% and -32.7%, respectively. These allocations imply variation in the investment allocation from a highly leveraged position to a modest short position. In data from 1946 through 1997, the these exposures are associated with an annualized return of 17.72% compounded annually, with a weekly standard deviation (1.88%) identical to a passive investment in the S&P 500 Index. Moreover, the maximal loss for this strategy is -35.30%, compared with a maximal loss of -44.64% for the passive strategy.

Based on the conditioning information used to partition the return distributions, 106 shifts are identified during the 52-year sample period, or approximately two shifts per year. Assuming a high 1% transaction cost applied to changes in allocation, the risk-equating value of A is 6.40. After transaction costs, the compound annual return is 12.30%, nearly equal to the 12.35% return from the passive approach. The maximal loss from the asset allocation approach (-38.06%) remains smaller than the maximal loss for the passive strategy. Amortizing this level of transaction costs over the holding period, purchases in Distribution III and short positions in Distribution IV are associated with risk, but negative excess return. Setting allocations to zero for these distributions, the risk-equating value of A is 6.26, the compound annual return rises to 12.64%, and the maximal loss is -37.70%. In addition to transaction costs, the rate of taxation and differences in borrowing and lending rates are important in evaluating the desirability of an active approach.

5. Concluding Remarks

The results presented here suggest that even a small amount of conditioning information is capable of partitioning market returns into conditional distributions with widely dissimilar characteristics. No stable relation between expected return and risk is observed, indicating that market risk exhibits predictable, time-varying efficiency in generating expected return. Evidently, stock prices are not sufficient statistics for all current information about future asset returns.

Again, it is important to emphasize that predictability of excess returns is not inconsistent with broadly defined concepts of economic efficiency and equilibrium. In a competitive market, profits represent compensation for the provision of scarce resources to other market participants. If a portion of trading is non-speculative "noise" driven by liquidity or insurance (consumption covariance) motives, excess returns will be available in equilibrium to compensate other traders for providing liquidity and bearing consumption risk. As demonstrated by Hussman [1992], such "noisy" environments may induce trading and divergence of opinion even between fully rational traders, and may create profit opportunities on the basis of both private and public information in equilibrium.

The substantial differences in expected returns across distributions suggest a role for active asset allocation, though the effect on long-term returns is dependent on the impact of trading costs and taxation. Trading costs create a region in which the market may be "inefficient" while still excluding the possibility of abnormal risk-adjusted returns. Yet even if sufficiently high trading costs reduce long term returns below those of a passive approach, an active approach may still be optimal from the standpoint of utility maximization. This is particularly true for "myopic" investors whose subjective utility is defined over the sequence of returns during individual holding periods and not only over terminal wealth.

Finally, it is notable that the profile of returns in Table 1 permits the emergence of stock market "bubbles". Despite the below-median dividend yield in Distribution II, the expected return is significantly above-average, and the standard deviation of returns is lower than in any other distribution. In effect, a favorable trend in yields may be expected to generate a further short-term market advance, regardless of the specific level of yields. During a sustained "bubble", the observation of above-average price advances with low volatility might reasonably lead to a reduction in the risk aversion of market participants, resulting in a further decline in yields (i.e., long-term capitalization rates) and correspondingly positive short-term returns.

Based on the persistent advance in the S&P 500 Index from 1995 through 1998, coupled with record low dividend yields, this process appears to be a reasonable characterization of the U.S. stock market in recent years.

References

Brennan, M. J., E. S. Schwartz, and R. Lagnado, 1995, "Strategic Asset Allocation," Working paper 11-93, Anderson Graduate School of Management, University of California, Los Angeles, CA, February.

Grossman, S. J., 1989, *The informational role of prices*, MIT Press, Cambridge, MA, chap. 3.

Hussman, J. P., 1992, "Market efficiency and inefficiency in rational expectations equilibria: dynamic effects of heterogeneous information and noise," *Journal of Economic Dynamics and Control*, 16, 655-680.

Lintner, J., 1965, "The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets," *Review of Economics and Statistics*, 47, 13-37.

Merton, R. C., 1969, "Lifetime portfolio selection under uncertainty: the continuous-time case," *Review of Economics and Statistics*, 51, 247-57.

Sharpe, W. F., 1964, "Capital asset prices: A theory of market equilibrium under conditions of risk," *Journal of Finance*, 19, 425-442.