Relative strength does persist!

A meticulous set of tests proves the point beyond chance.

Charles A. Akemann and Werner E. Keller

The assumption that the historical sequence of securities prices does not contain any useful information for predicting future prices is one of the cornerstones of the efficient market hypothesis. The literature is filled with reports by investigators who have tested a wide range of possible correlations and trading strategies to the point that there remains little doubt within the academic community that successive price changes are essentially random. It is therefore not surprising that the validity of technical analysis, which presupposes the existence of trends in price behavior, is, in the words of a recent article, "no longer considered an interesting area for academic inquiry."

The study summarized in this article has nevertheless identified some long-term dependencies in stock prices that are not only statistically significant: they can be exploited to produce investment results that, after transactions charges, are superior to buying and holding a market index. The approach used, in addition to setting its filters to rather long time spans, focuses on the prices of industry stock *groups* rather than those of individual stocks as its basic data. The effect of both of these choices is to greatly reduce, through averaging, the amount of background "noise" created by shorter-term fluctuations that may very well have random characteristics, as the work of many investigators has suggested.

An approach using industry groups has a further advantage in that it leads directly to a central problem of investment management, that of industry emphasis within the portfolio. Since portfolio managers tend to be more interested in obtaining guidance in the choice of industries than in the choice of stocks, it is not surprising to see them give considerable attention to relative price movement among groups as, for example, the behavior of bank stocks versus the general market, international oils versus domestic oils, and so on. Their interest in such comparisons is based on the belief that such trends in relative movement may persist for some time into the future. Even though the literature on the question of trends has generally identified only very minor and non-exploitable dependencies, it is remarkable how widely the general notion of strength persistence is held within the financial community. Catering to this belief is a diverse group of mostly high-priced services that rank, smooth, chart, and otherwise interpret the movements of industry groups.

TECHNIQUE OF ANALYSIS

In the most straightforward test of strength persistence (the one we have examined statistically), the price change of an industry group index is observed over a specified time period, often several months, and its return (price change only) over this observation period is compared to the returns of other groups. All groups are then ranked by rates of return, and investment decisions are made on the presumption that the superior performance exhibited by highly ranked groups will tend to persist. The absolute size of the price change over the observation period is unimportant. As a matter of fact, in a very weak market, the highest ranked group may be distinguished only by a decline that is less than any of the others. Strength, then, is always defined in relative terms (hence the widely used term, "relative strength").

With the advent of on-line stock data bases, computer users are now able to construct industry groupings as they see fit and generate index values as frequently as they wish. Still, the indices used most widely by the investment community, and the only ones for which consistent historical data are available, are those published weekly by Standard & Poor's Corporation. As the historical series of these indices is not

available in computer-readable form, we assembled our own basic data base of 461 weeks, beginning with January, 1967 and extending through October, 1975. In the construction of this data base, certain clearly non-homogenous groups such as Miscellaneous Metals and Specialty Machinery were eliminated, as were the composite groups such as Food Composite or Retail Composite, which simply represent the aggregate of more specifically defined subgroups (e.g., Canned Foods, Dairy Products, etc.). Also eliminated for procedural reasons were all groups that were not continuously in existence throughout the period of the data base. The remaining industry groups nevertheless accounted for over 95% of the weighted capitalization of the S & P 500 Composite Stock Index.

Our investigation used concepts and parameters that seemed logical and non-controversial within the context of the relative strength approach. Furthermore, although certain refinements were tested, the basic selection and investment procedures remained the same throughout. The observation period was standardized at thirteen weeks (one quarter), to minimize any recurring price biases from earnings announcements, ex-dividend dates, and so on. The weekly index values were slightly smoothed by averaging two weeks of data at the beginning and the end of the observation period.

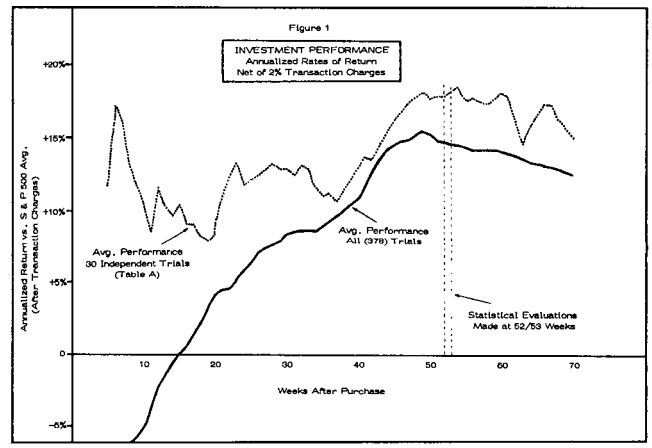
Thus, for example, the program would average the index value of weeks numbered twelve and thirteen and those for weeks one and two, and then calculate the percentage price change over the period. Next, the price changes of all groups would be ranked and the group with the greatest gain (or smallest loss) over the observation period would be selected for investment. Purchase was assumed to occur over the two weeks following the end of the observation period, or at the average prices prevailing during weeks fourteen and fifteen in our previous example. Then, the performance of the selected group was tracked week-by-week against the performance of the S & P 500 Index for the ensuing seventy weeks. Each week after purchase the program reduced the selected group's performance by a 2% transaction charge[1] and then compared the annualized net performance with the annualized change in the market index. For example, if the group had risen 8% and the market index 5% at the thirteenth week after purchase, the program would reduce the 8% to 6% (the 2% transaction charge), annualize the 6% to 24% and the 5% to 20% (multiplying by 52/13) and report the difference of 4% (24% - 20%).

The length of the data base permitted 378 individual trials of the concept, each trial consisting of a thirteen-week observation period and a subsequent seventy-week holding period.

1. Footnotes appear at the end of the article.

STRENGTH PERSISTS

The composite results of all these trials revealed a rather pronounced and long-lived strength persistence effect. The routine of purchasing the strongest stock group in a given thirteen-week period showed itself capable of selecting investments that, on average, exceeded the performance of the market index by more than 10% annually, after transactions costs. Although losses predominated during short holding periods, favorable investment results were obtained for holding periods ranging from approximately fifteen weeks to more than eighteen months, with the maximum performance relative to the index, 13%-15%, reached between 45 and 65 weeks after purchase. The shape of this performance curve is depicted by the solid line in Figure 1.



WHY?

In view of the well-documented failure of other investigators to uncover such material relationships among historical stock prices, we considered it imperative to develop both conceptual and statistical validations of these findings.

On the conceptual side, we feel there are two broad reasons why such a system could be successful. First, as already noted, substantial averaging occurs from both the use of stock *groups* and from the rather long thirteen-week observation periods. Second, our selection technique represents, in concept, an attempt to isolate stock groups exhibiting *unusual* price behavior — those undergoing substantial fundamental revaluation. It does not attempt to say that all price movements have predictive value. To the contrary, our evidence indicates that only certain types of price action have predictive value, but that these can be

meaningful in terms of portfolio performance. In selecting only the single best performing groups from among the 73 being monitored, there is at least a theoretical likelihood that our procedure isolates a thirteen-week price change that is not a "normal" fluctuation, but rather one that reflects an important shift in investor attitudes toward that group.

The tendency of such a group to continue to outperform the market index for a considerable time suggests the existence of a "ripple effect" in investor enthusiasm that is both long-lived and *price-insensitive.* In other words, investors seem to show a tendency to buy a group if they like the prospects (or the "market action"), even though the price may have already risen substantially. Such behavior is, of course, fundamentally contrary to the efficient market hypothesis.

In order to pursue this conceptualization a bit further, we compared the stock groups chosen in our independent trial series (Table A) to a compilation of the five groups that did, in fact, have the largest and most unusual price movements over the period of the trials. These ideal buy-and-holds were:

LARGEST INDUSTRY GROUP PRICE GAINS

April 5, 1966–June 25, 1975

Oil Well Machinery	+555%
Truckers	322%
Gold Mining	281%
Coal	210%
Variety Stores	110%
S & P "500" Average	+ 8%

The comparison shows that, over this eight-year period, our rather simple decision mechanism identified one of the top four groups for investment in thirteen out of thirty trials, a sufficient frequency to reinforce the notion that high relative performance tends to be preceded by unusual price activity (defined in our context as a rapid positive price change).

THE RESULTS IN MORE DETAIL

In our tests of statistical significance, we chose to consider non-overlapping observation periods; that is, for any two tests of the theory (trials), the investment decisions were based on disjoint sequences of prices. Thus, if the observation period for the first trial spanned the thirteen weeks numbered five to seventeen, for example, the second trial would begin at week number eighteen and run to week number thirty, and so on. This procedure provided thirty independent trials (consisting of an observation period plus an investment period) under a variety of market conditions, including two full bull-bear market cycles. We standardized the life of all investments at one year and, accordingly, caused the program to "sell" the group at the average price prevailing 52 and 53 weeks

TABLE A

PERFORMANCE SUMMARY

30 Independent Trials
One-Year Holding Period

Trial Number	Group Selected	Date of First Purchase*	Return vs S & P-500 At 52 wks.*
1	Machine Tools	4/5/67	3.1%
2	Oil Well Machinery	7/5/67	22.0
3	Savings and Loan	10/4/67	64.7
4	Gold Mining	1/3/68	24.9
5	Gold Mining	4/3/68	28.4
6	Property/Liability Insurance	7/3/68	-2.8
7	Cement Products	10/2/68	-12.4
8	Truckers	1/2/69	-17.4
9	Mail Order Stores	4/2/69	11.0
10	Cosmetics	7/2/69	15.5
11	Coal	10/1/69	68.7
12	Coal	12/31/69	56.0
13	Gold Mining	4/1/70	20.3
14	Tobacco	7/1/70	7.6
15	Movies	9/30/70	-9.6
16	Tobacco	12/30/70	3.7%
17	Aerospace	3/31/71	9.0
18	Steam Gen. Machinery	6/30/71	-17.6
19	Truckers	9/29/71	1.4
20	Airlines	12/29/71	-30.2
21	Gold Mining	3/29/72	37.0
22	Brewers	6/28/72	5.5
23	Farm Machinery	9/27/72	27.8
24	Oil Well Machinery	12/27/72	56.8
25	Coal	3/28/73	25.3
26	Gold Mining	6/27/73	59.7
27	Farm Machinery	9/26/73	-7.5
28	Gold Mining	12/26/73	28.1
29	Movies	3/27/74	122.9
30	Soaps	6/26/74	-6.0

* As discussed in the text, the groups were both purchased and sold on 2 consecutive weeks.

Average Performance	+19.9%
- Transaction Costs	- 2.0
Net Performance vs S & P-500	+17.9%
Standard Deviation	32.4

after purchase. Beyond facilitating computations, this interval is appropriate in light of the new holding period for long-term capital gains. Only the week-to-week performance history in Figure 1 gives results for holding periods other than one year.

When we began the simulation with the first week of our data base, we obtained the sequence of group selection and investment performance summarized in Table A. The results of the individual trials can be seen to be volatile but clearly biased toward a price performance superior to that of the index. Even after subtracting transactions charges, for example, 21 (or 70%) of the trials resulted in an investment that subsequently outperformed the market index.

The trials detailed in Table A were begun, as noted, with the first week of data but, by selecting different starting points, other sets of independent trials, with differing performance characteristics, could have been chosen. The investment performance of each of these sets relative to the S & P 500, together with their statistical properties, is presented in Table B. The last column in Table B is the left-hand endpoint for the 90% one-sided confidence interval, which is constructed for the average return versus the S & P 500. To determine the interval, we assumed (from the Central Limit Theorem) that the average of the results of thirty independent trials would be normally distributed with the same mean as the original (unknown) distribution of possible trial outcomes and with variance reduced by a factor of thirty.[2] While this assumption is necessarily approximate, all statisticians would agree that it should be substantially accurate, given the number of independent trials available.

With this assumption in force, the construction of the confidence interval used the standard method. For example, in the first row of Table B, the value 12.1 in the last column is given by

$$19.9 - \frac{(1.3114)(32.4)}{\sqrt{30}},$$

where 32.4 is the sample standard deviation, 30 is the number of trials, and 1.3114 is from the table for Student's t-Distribution for 29 degrees of freedom at the 90% level. We can interpret the first row of Table B to mean that, with 90% certainty, the average annual return for the investment procedure we describe will be at least 12.1 percentage points above that of merely buying and holding the market index (before transactions costs).

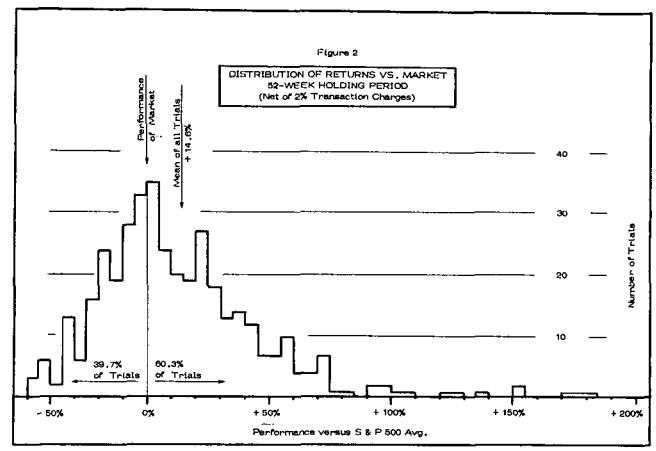
Although our relatively uncomplicated relative strength filter showed itself capable of selecting industry stock groups that, on average, tend to outperform the general market by a significant margin, obviously not every stock group selected by these procedures will perform according to the composite curve of Figure 1. A histogram of the 52-week investment results of all 378 possible trials provides a visual perspective on the variability of the phenomenon (Figure 2). After the 2% transaction charge, the peak of the distribution is only slightly positive (i.e., better than the market's performance), but the rather substantial 14.6% *mean* performance differential of all these trials arises from the positive skew of the distribution. Thus we might generalize that over the eight and a half years covered by our investigation, the approach succeeds by identifying a number of the "unique" winners and by avoiding catastrophic losses.

TABLE B

INVESTMENT PERFORMANCE
AND STATISTICAL CHARACTERISTICS
All Possible Sets of Independent Trials

Independent Trial Beginning at Week Number	Average Percentage Points Above S & P 500	Standard Deviation	Lower Confidence Limit One-sided t-test 90% Level
1	+19.9	32.4	+12.1
2	23.2	31.1	15.5
3	13.8	35.5	5.3
4	10.9	36.5	2.2
5	12.0	29.0	5.1
6	9.5	26.6	3.1
7	12.7	40.7	3.0
8	15.2	41.3	5.3
9	14.1	39.2	4.7
10	9.3	37.3	0.4
11	21.4	38.7	12.1
12	26.4	43.3	16.0
13	19.5	30.9	12.1

Average of Column 2	= +16.01	percentage points
Less transaction charges	= -2.00	percentage points
Net annual performance differential	= +14.01	percentage points



FORMULATING A STRATEGY

These observations lead directly to the problem of applying such a selection technique to a practical portfolio management situation. Clearly, the variability of returns would argue against investing the entire portfolio in one stock group, but the optimum investment strategy, given the statistical properties of the phenomenon, would be difficult to determine.[3]

As an initial step in this direction, we chose a set of portfolio management rules that (1) preserved strict statistical independence, and (2) followed naturally from our standardized thirteen-week observation/52-week investment procedure. Thus, we structured an *ex post* simulation by dividing a hypothetical $100 portfolio into four equal parts and investing one of the $25 slices as soon as our data base provided the first group selection (i.e., after week number thirteen). After passage of another thirteen weeks, the second $25 was invested in the group then selected and so on, until all four parts of the portfolio were invested. Then, as each investment attained its one-year holding period, it was sold and the proceeds were reinvested in the highest strength group selected at that time. This procedure was continued to the end of the data base, when, on a quarterly basis, each of the four sub-portfolios was sold and the proceeds held in cash until the last holding was sold, at which point the simulation ended. The cash holdings of the phase-in and phase-out periods were assumed to earn a 5% annual return, and the customary 2% transaction charge was applied throughout.

The results on an absolute basis are summarized in Table C and should be examined with reference to Table A, which details the groups selected in this simulation. The substantial difference between the 14.22% compounded annual growth rate for our "high-strength group" investment procedure and the 0.68% annual rate for the market index is worthy of note, particularly since our rather active strategy (100% annual turnover) was more successful than buying and holding 69 of the 73 industry groups individually over this period.

Since, as a practical matter, it was not possible to construct a dividend history for the S & P industry groups, yield was not taken into consideration in computing the performance of either the groups or that of the market index. On the basis of the sample of trials detailed in Table A, however, it does appear that the system tends to select groups whose component stocks have below average dividend yields, and thus the performance differential on a total return basis should be adjusted downward by perhaps 1% to 2%.

Interestingly, however, the average beta coefficient (calculated on a trailing 52-week basis) for these thirty groups was 0.98 at the time of purchase, suggesting an average portfolio risk level essentially equal to that of the market index. Since adjustment for yield variations would affect the size of the performance margin, but not the conclusions of the study, we deferred the rather substantial research required to quantify the historical yield differentials precisely.

Considerations of statistical procedure have necessarily focused our attention on independent trials with non-overlapping observation periods for the critical tests, but it is not clear that varying degrees of non-independence (by having observation periods overlap to some degree) would either increase risk or reduce performance. Thus, since the selection filter produces a fresh candidate each week, a real-world application might well divide the portfolio into a larger number of sub-portfolios and transact more frequently.

TABLE C

SUMMARY OF PORTFOLIO SIMULATION

Date First Investment Purchased	April 5, 1967
Date Last Investment Sold	June 25, 1975
Time Elapsed 429 weeks	(8.223 years)
Final Value of $100 (after transaction costs)	$298.39
Compounded annual rate of return	+14.22%
Corresponding Compound Return of S & P 500	+ 0.68%

FURTHER PROOF

Additional support for the concept of strength persistence comes from the price behavior of groups ranked immediately behind the best performing group during the selection process. Ideally, we would expect the stock group with the *second highest* performance rank over the observation period also to outperform the market index, but by a smaller differential than the first ranked group. We tested this hypothesis and found perfect rank/performance correlation among the top five groups in thirty independent trials corresponding to the sequence in Table A. These five groups, taken individually, achieved the following positive performance differentials against the S & P 500 (after 2% transactions charges):

First-Ranked Group	+17.9%
Second-Ranked Group	9.7%
Third-Ranked Group	8.1%
Fourth-Ranked Group	7.9%
Fifth-Ranked Group	0.8%

The sharp decline between groups four and five does not, of course, suggest that all groups of rank six or lower underperformed the market, but only that the statistical advantage of this selection procedure seems to be lost beyond the few top-ranked groups. In fact, the rather large differential that is already present between groups one and two further supports the notion of the highest ranked groups' "uniqueness."

We feel that the 461-week data base constructed for these studies is sufficiently diverse in terms of market environment to have provided a fair test for the strength persistence approach to investment selection, as we have defined it, and that the phenomenon can be demonstrated with a satisfactory degree of statistical certainty. Although we tested some variations in the selection formula to rule out the possibility that our findings were the result of an elaborate coincidence, we did not explore rigorously any other selection procedure, let alone any of the sophisticated weighting and other manipulations of which computers are capable. There is, however, some preliminary evidence that the present selection filter can be significantly optimized.

When it comes to exploiting the phenomenon in the market, we must recognize that, as in all systematic approaches, technical or otherwise, not everyone can play the same game. Relative strength investing is, after all, a somewhat parasitic exercise wherein someone else detects underevaluation, attractiveness, or some other reason to buy, begins to act on it, and the relative strength player comes along for the ride. Obviously, the larger the number of investors who wish to ride the same wave, the shorter the ride, or discounting interval. However, we believe that the mechanical nature of the approach we have described is likely to be sufficiently unpalatable to the average investment manager as to make widespread application of these techniques unlikely.

[1] Whether 2% is an appropriate estimate of commission and other charges is a separate debate, but estimates of "real costs" of 5% and more suggested by some authors seem difficult to defend. If an institution takes a discount in order to sell a block of stock, there is, after all, someone on the other side (most often, another institution) who receives the benefit. Under the pre-May 1, 1975 schedules, commissions on a typical $50,000 value trade (small by institutional standards) were approximately 1% each way, and today they are at least 30% less. Furthermore, since the S & P Index pays no commissions or management fees, its performance tends to overstate the actual performance attainable by an index fund, for example. We thus considered the 2% transaction cost level reasonable, but this assumption is not critical to our findings.

[2] Some investigators, notably Eugene F. Fama (e.g., "The Behavior of Stock-Market Prices," *Journal of Business*, 38 (1965), pp. 34-105), have made a strong case for the hypothesis that the differences of the natural logarithms of successive security prices have stable Paretian distributions with $\alpha < 2$ to which the usual Central Limit Theorem would not apply. However, since we do not use logarithms, our maximum negative result is -100%. Thus the only "long tail" of our distribution is in the positive direction. Our assumption of normality for the distribution of the thirty trial average may understate our case, but can never overstate it.

[3] A most readable discussion of the difficulties inherent in the selection of optimal investment strategies can be found in E. O. Thorp, "Optimal Gambling Systems for Favorable Games," *Review of the International Statistical Institute*, Vol. 37:3, 1969. Our strategy selection was guided in part by criteria discussed in this paper.