THE JOURNAL OF FINANCE • VOL. XLI, NO. 3 • JULY 1986

Discrete Expectational Data and Portfolio Performance

EDWIN J. ELTON, MARTIN J. GRUBER, and SETH GROSSMAN*

ABSTRACT

In this article we examine the information content in analysts' recommendations which are made on a five-point buy, hold, or sell scale. Our data set includes data on 10,000 forecasts per month. Unlike most prior studies, our data set does not suffer from selection or survivorship bias. We find information in analysts' changes in recommendations. Approximately 4.5% extra return can be earned by purchasing new buys rather than new sells.

IN THIS ARTICLE we examine the information contained in analysts' buy, hold and sell recommendations. There are a number of reasons why such a study is important. As we discuss in more detail below, this study provides an evaluation of forecast ability and a test of strong form efficiency. In addition, this study provides some insight into whether information is lost in the investment process. What makes this study unusual is that it is one of the few cases in evaluating informational content where the forecaster is recommending a clear and unequivocal course of action rather than producing an estimate of a number, the interpretation of which is up to the user. In addition, the large number of forecasts collected and the manner in which they were collected (with no ex post bias) differentiates this study from most others.

While expectations play an important role in most economic theories, studies evaluating expectational data are not common. In addition, studies which do evaluate forecasts generally assume an arbitrary payoff (penalty) function for forecast errors. The squared forecast error is the most commonly used criteria for judging the accuracy of forecasts. The value of analysts' recommenations is conceptually easy to measure. A more natural metric exists (risk-adjusted excess returns) for measuring the value of a recommendation and whether the recommendations are useful or not.

Our study also has implications for the efficient market literature. Strong form tests of efficient markets are perhaps the most important and least frequently performed tests. All other efficiency tests require a theory of what type of information (e.g., earnings increases) should affect stock prices and then a test of whether the information hypothesized as affecting stock prices does in fact affect stock prices. For example, in previous papers [2, 3] it was found that although analysts' forecasts of earnings per share are quite accurate relative to naïve models, the market is efficient with respect to these forecasts. In this paper, we examine whether analysts' recommendations can lead to superior returns. No

* New York University. We would like to thank Banker's Trust Company for supplying the data.

knowledge is needed or hypothesized about how the analysts would or do use any type of information. The question examined is whether analysts, using any method and information they deem effective, can form forecasts which can be used to produce an excess return.

Studies evaluating mutual fund performance have generally found that mutual fund performance does not dominate random selection. This study will help us to understand whether this performance has been due to poor forecasts of returns on securities or a poor ability to utilize these forecasts.

This clearly is not the first study on the value of analysts' recommendations. Dimson and Marsh [1] provide an excellent review of previous studies and we will not repeat the review here. Our study is different from past studies in that we employ a larger and more carefully constructed data base than those used by most past researchers and we use a more robust methodology to adjust for risk. Most previous studies of recommendations use Value Line data. There is nothing wrong with this per se. Most psychological studies analyze sophomore college students. One wonders whether these studies generalize to the "normal population". The same is true with Value Line studies. There is another problem with Value Line studies. Value Line has the reputation of producing good forecasts. Researchers who have studied this have found it to be true. Thus, studies of Value Line involve an ex post selection bias (Gregory, [4]). This same selection bias is present in many of the other studies reviewed by Dimson and Marsh. Namely, for many of the studies the data was requested and provided long after the forecasts were made. One would expect that data would more likely be provided if analysts' recommendations were good. The current study analyzes by far the most extensive data base of analysts' recommendation ever assembled and does not have the ex post selection bias of many previous studies. For this study a large number of brokerage firms were asked to submit rankings. Once a brokerage firm started submitting rankings, it continued for the entire time period of the study with only one exception. In the next section we will present the details of the data base, but to give an idea of its size, it contains data from 720 analysts at 33 brokerage firms, producing over 10,000 forecasts per month. The data base is three times larger than the total of the data used in all prior studies.

In addition, this study has the advantage that the timing of availability of the data is fairly precise. The collection procedure and the dissemination of data followed a consistent procedure over time.

Our methodology differs from that used in previous studies in that we do not assume that a particular capital asset pricing model holds. We make the more robust assumption that beta is either in general a sufficient metric for risk or the only metric that systematically differs across recommendations. Thus, we develop comparison portfolios based on differing recommendations but having the same beta.

The paper is divided into three sections. In the first section we discuss the data in more detail. In the second section we analyze the effect of a change in an analyst's recommendation. In the final section we look at the value of a firm's recommendation list.

We place major emphasis on analyzing changes in analysts' recommenations because a change in recommendation should be the direct signal that analysts expect a discontinuity in a stock's performance. Analysts' recommendations at any moment in time represent an accumulation of past changes. If information is quickly incorporated into share price and analysts revise recommendations more slowly (recommendations are sticky) then a change in recommendation on a stock should contain more information than a continuation of a past recommendation.

I. Data

The data we utilized was compiled by the investment group at Banker's Trust Company and was marketed under the name IBOS. Starting in March, 1981, forecast data was collected from 33 brokerage firms. One brokerage firm was changed over the period so that we have data from 34 firms. Although data is still being collected, we end our study in November, 1983 because of the need for data on returns after the forecast is made. Thus we have 33 months of data.

The data collected consisted of the rank of each stock on a five-point scale. A rating of 3 was a neutral rating. A rating of 1 or 2 was a buy recommendation with a 1 the more desirable. A rating of 4 or 5 was a sell recommendation with a 5 the least desirable. The average number of analysts utilized by the 33 brokerage firms was 659 in 1981, 734 in 1982 and 766 in 1983. The data was collected over the last week of every month. However, the recommendation could have been made at any time over the month. This data base was constructed by an institution the business of which was valued by the brokerage firms. Thus, the brokerage firms had an incentive to supply good forecasts and the forecasts are likely to be more accurate than those employed in normal academic surveys.

As should come as no surprise to anyone who has observed brokerage firms, there were many more buy recommendations than there were sell recommendations. Table I shows the breakdown by rating for our three sample years and the overall period. Forty-eight percent of the recommendations were buys (1's or 2's) and only two percent were rated 5's (a grading distribution that would warm the heart of most students). Table II shows the distribution of changes in classifications. On average, about 11% of the classifications are changed in a month.

We deleted some stocks from our data base. As we will discuss later, we were concerned about holding constant other influences. One influence we would like to control is any small firm effect. To control for this, we effectively eliminated small firms. We eliminated all firms that were not followed by three or more analysts.[1] We found that this rule led to a sample of large capitalization stocks. This eliminated 12% of the forecasts on average over the three years.

We also utilized the CRSP data base to calculate monthly returns and betas.

[1] A brokerage firm may follow a small stock because of a special underwriting relationship. However, firms followed by multiple brokerage firms are large capitalization stocks. We could have used size directly. However, examining several samples showed this procedure served the same function and was much easier to implement.

Table I

Number of Firms in Each Rating Category

Rating	1981*	1982	1983**	Averaged Over 33 Months
1	1475	1592	1654	1577
2	3185	3135	3420	3245
3	2785	3920	4575	3794
4	895	1144	1307	1122
5	127	264	313	239
Total	8467	10655	11269	9977

* 10 months
** 11 months

Table II

Number of Changes to Each Rating Category

New Rating	1981*	1982	1983**	Averaged Over 33 Months
1	168	150	178	165
2	278	385	386	353
3	338	485	401	412
4	116	228	189	181
5	127	58	46	45
Total	927	1306	1200	1156

* 10 months
** 11 months

II. The Value of a Change in Recommendation

In this section we analyze whether there is information in changes in analysts' recommendations. We analyze this in two ways, first by pooling the data from all brokerage firms, and second by using the data for each brokerage firm separately. In all cases we analyze the information content by calculating the return on a portfolio of stocks that were upgraded or downgraded in ranking and comparing it to a neutral portfolio with the same beta. Each of these analyses will now be discussed.

A. Aggregate Comparison

The analysis is easiest to understand in the context of a specific example. Consider securities which were upgraded to 1's and 2's. For all 33 brokerage firms for each month we collect newly upgraded 1's and 2's. We also collect downgrades to 4's and 5's as a benchmark portfolio. For each month, we then construct a portfolio of upgraded 1's and 2's that has the same beta as a portfolio of downgraded 4's and 5's. This was accomplished as follows. First we calculated a beta on each security using the market model for the five years prior to the month in question. The market index was a value-weighted index of all stocks on the NYSE. Then the two groups of upgraded 1's and 2's and downgraded 4's and 5's were each ranked from highest to lowest beta. The beta on an equally-weighted portfolio was calculated for each group. For the group with the largest number of firms, firms were deleted at the extreme (e.g., if its beta was higher than the other groups, the biggest beta stocks were deleted) until the betas on the two groups were within .01 of each other.[2]

Having formed two portfolios with the same beta, one with a change to a more positive and one with a change to a more negative recommendation, we then computed the differential return on the two groups to see if changes in analysts' recommendations contain information.

This procedure differs from the procedure used in past studies of analyst performance. We believe that our procedure is superior for two reasons. First, our procedure only assumes that beta is a reasonable measure of risk and thus it is robust across many alternative definitions of the "correct" equilibrium process. The last decade has produced a large number of alternative general equilibrium models. Not only are there a dozen alternative CAPM models, but arbitrage pricing theory has been put forward as a competing paradigm. Since arbitrage pricing theory does not require any particular set of factors influencing expected returns, each test of APT can be viewed as producing a unique model. Additional research in evaluating expectational data requires either a delay until equilibrium theory has been sorted out or the use of a research design that is likely to be robust across a number of alternative theories. We believe our design is robust. The beta adjustments we make would be appropriate if either the zero-beta or the standard CAPM is a reasonable model and avoids specifying a specific intercept. If an alternative CAPM is the correct equilibrium model, then the adjustment is still appropriate if rcommendations are not based on sensitivity to other influences. For example, if the post-tax CAPM were appropriate, then our adjustment would hold constant the market influence but not the effect of dividend yield. Not adjusting for dividend yield just adds random noise unless rankings correspond to dividend yield. Furthermore, the first index in most APT models is very closely related to a market portfolio. Once again, as long as brokerage selection is not based on sensitivity to other factors, this adjustment is appropriate.

Second, our results are not affected if brokerage firms as a whole follow only a homogenous subsection of the market. Brokerage firms tend to concentrate on certain types of stocks. At the very least, they concentrate on large capitalization companies. In addition they may well concentrate on certain sectors. Insofar as these sectors outperform or under-perform the market as a whole, then comparison of new 1's to the average securities in the market is in part an examination of how the sample of stocks followed by the brokerage firms performed as a whole. By comparing a broker's 1's and 2's to the broker's 4's and 5's, we alleviate this problem.

Before examining the results, one other issue needs to be examined. A stock which is upgraded to a 1 by one broker may be rated a new 5 by a second. Thus,

[2] Whether it was high or low betas that were deleted varied randomly over time.

the same firms may be in both groups. We did do a second analysis where we eliminated all firms that were in both portfolios. This did not change the results so we only report results where the overlaps were not eliminated. In addition, a single firm may be upgraded to a 1 by more than one brokerage firm. Our procedure weights the firm by the number of times it has been upgraded to a 1. For example, a firm that is upgraded to a 1 by four different brokerage firms receives four times the weight of a firm that has been upgraded once. There are an enormous number of alternative weighting schemes that could be tried. We are making no claims that the two weighing schemes we tried are the only possible ones or even the best. However, insofar as they are imperfect, it biases the results against finding information. Thus if we find no information a justifiable comment is that a superior weighing scheme might have disclosed real information. To search for the optimal weighing scheme, while interesting, is well beyond the purpose of the paper.

In Table III we present results of the performance of a portfolio of newly recommended 1's compared to a portfolio of new 3's. Similarly results on newly downgraded 4's are compared to new 3's. This pattern is similar for all comparisons we make and therefore this is the only table in which we present an extensive time series of results. Let's examine the results. Before time zero the differential return is insignificant.[3] In the month of the recommendation, time zero, the differential return is large and positive. The large positive return persists in the month after the change in recommendation, and persists but is somewhat smaller in the second month after the change in recommendation. For downgraded stocks a negative return occurs in the month of the change in recommendation. A negative return of a larger size occurs in the month after the recommendation. While there are additional negative returns in later months none of these are statistically significant. In subsequent tables we will present results only for the month of the recommendation and the subsequent two months, since these are the only months with significant returns.

Table IV presents results for all possible changes in recommendations as well as a comparison of new 1's and 2's to new 4's and 5's. Except for securities that were upgraded to 4's, the results for upgrades are uniformly of the correct sign and highly significant.[4] For upgrades there is a large differential return in the month of the announcement. An additional increment in return occurs in the month after the announcement and a smaller, generally insignificant return two months after. For downgrades the negative returns generally are insignificant in the month of the announcement. However, there is a substantial negative return in the month after the announcement. Once again, two months after the announcement there are insignificant returns. Also, note the monotonicity across changes. Upgraded 1's result in greater returns than upgraded 2's and upgraded 2's give higher returns than upgrades to 4's. Rank monotonicity is observed with downgrades as well. The magnitude of these numbers is substantial. For example,

[3] In this particular case the return is primarily positive for upgrades and negative for downgrades. This pattern varies with different upgrades and downgrades and is never significant.

[4] For each of our 33 months we have a differential return. The mean is the average return over the 33 months and the standard deviation used in the t-test is the standard deviation of the 33 differential returns divided by the square root of 33.

Table III

Return in Percent

Month	up to 1 compared to 3	down to 4 compared to 3
−6	0.07	−0.33
−5	0.55	−0.24
−4	0.19	−0.08
−3	−0.07	−0.10
−2	0.17	0.06
−1	−0.27	−0.02
0	1.91	−0.56
1	1.24	−0.74
2	0.28	−0.08
3	−0.11	0.04
4	0.37	0.17
5	0.06	0.00
6	0.28	−0.14
7	0.04	−0.07
8	−0.04	0.22
9	−0.05	0.13
10	−0.11	0.13
11	−0.34	−0.33
12	−0.34	0.53
13	0.18	−0.04

for new 1's, the total excess return was almost 3.43% over the three months. This is clearly in excess of transaction costs.[5]

Another way to examine the effect of upgrades is to look at the sign of the differential return on upgraded stocks for each of the 33 months (cases) in which forecasts were made. This is a different test since the return numbers shown in Table IV are the average over all 33 months. For upgrades, negative returns occurred in only 5 of 33 months (cases) for the announcement month, 3 of 33 for the month after and 11 of 33 for two months after. In only 2 of 33 cases was the total return negative over the three-month period.

Finally, comparing a portfolio of upgrades with a portfolio of downgrades gives the greatest excess return of all. It is large and significant in each of the periods although only at a 5% level for two months after the forecast was made.

It is hard to attribute these results to anything but information in analysts' rankings. Brokerage firms concentrate on large capitalization stocks. As discussed earlier our sampling rule eliminated all low capitalization stocks. Thus, excess returns on small firms is not an explanation.[6] As an example of other possible

[5] We can not be certain that the return in month zero is after the announcement so the 3.43% may be an overestimate. However, we do not see any excess return in the month before the announcement. If there were returns before the announcement we should see an effect in month −1 for firms with announcements at the beginning of the month. We do not see such an effect. Thus the data suggests the effect is after the announcement.

[6] If small firms was an explanation the effect would be concentrated in January and we do not observe such a pattern. Furthermore, small firm effect can not explain the persistent month-to-month pattern.

Table IV
Excess Portfolio Return Comparisons–Combined

Change in Class	Comparison	Month of Recommendation			1 Month After			2 Months After		
		Mean	t Value	Number Negative	Mean	t Value	Number Negative	Mean	t Value	Number Negative
up to 1 or 2	down to 4 & 5	2.43	6.04	3	1.86	6.83	4	0.37	1.68	9
up to 1	3	1.91	7.25	5	1.24	5.17	6	0.28	1.51	13
up to 2	3	1.65	8.64	4	0.84	5.53	6	0.20	1.03	13
up to 4	3	0.68	0.93	11	0.09	0.14	18	−0.80	−1.12	19
down to 2	3	0.29	1.30	13	−0.37	−1.77	23	−0.09	−0.38	20
down to 4	3	−0.56	−1.78	22	−0.74	−3.19	26	−0.08	−0.39	17
down to 5	3	−0.38	−0.60	17	−1.48	−4.05	26	−0.40	−1.17	19

omitted influences, consider dividend policy. Some researchers find that high dividend stocks give higher risk-adjusted return. One possible explanation of our results is that upgrades are high dividend stocks and downgrades are low dividend stocks. Not only is this implausible but it cannot explain the consistency of our results from month to month. Even the most ardent advocates of excess returns on high dividend stocks have not argued that this occurs in every month. In fact, no one has found a non-beta influence that could account for the consistency of these results.

B. Broker by Broker Comparison

Table V disaggregates the results to the brokerage firm level. The major difference in disaggregating at the brokerage firm level is that each brokerage firm is examined separately and matching portfolios are formed for each broker as opposed to across all brokers. There are several changes in the research design when the results are disaggregated to the brokerage firm level. First, at the brokerage firm level we do not have as many firms with changes in recommendations. Thus, for upgrades, we combine upgrades to 1's and 2's. Likewise for downgrades, we combine downgrades to 4's and 5's. The comparison group is also changed. There are not enough new 3's. Thus we use a comparison group, 3's whose rating had not changed from the prior month. The second change is in the way we construct our portfolios. Consider upgrades. For each brokerage firm we match the beta on the upgrades with the beta on the unchanged 3's. Matching at the brokerage firm level has the disadvantage that we had to combine 1's and 2's and also 4's and 5's. It has an advantage in controlling for other factors. Insofar as a brokerage firm specializes in a sector, this design eliminates the effect of abnormal performance in that sector on our results.

Once again the same pattern emerges for the upgrades. A large return in the month of an announcement is followed by a substantial return in the month after and a smaller return two months after the announcement. Similarly, for downgrades there is a substantial negative return in the month of the announcement (more so than the prior table) and a still substantial return the month after. Note that these results are the average results across the 33 brokerage firms. In Table IV all data was combined. Thus the portfolios used to produce the results in Tables IV and V were not constructed in the same manner. Table V also shows two results that do not have a parallel in Table IV. First, for upgrades the total return was negative over the 33 periods for only 5 of the 34 brokerage firms in the month of the announcement, 5 in the month after and 5 two months after. These are not the same five firms in all cases. Only three of the brokerage firms had negative returns over the full three-month period. Correspondingly for downgrades, 29 of 34 brokerage firms had negative differential returns in the month of the announcement and 28 of 34 in the month after the announcement while only 7 of 34 firms had the wrong sign over the 3 months.

All of these results are statistically significant so that we can state that analysts' recommenations have value at a brokerage firm level. We made several comparisons to see if the ranking across brokerage firms was consistent over time. We ranked brokerage firms by the size of the differential excess return earned on

Table V

Excess Portfolio Return Comparisons–Broker

Change in Class	Number of Brokers	Month of Recommendation			1 Month After			2 Months After		
		Mean	*t* Value	Number Negative	Mean	*t* Value	Number Negative	Mean	*t* Value	Number Negative
up to 1 or 2	all	1.05	5.50	5	0.87	4.73	5	0.54	4.44	5
up to 1 or 2	Best 5	1.50			2.08			0.67		
down to 4 or 5	all	−1.74	−3.88	29	−1.30	−4.19	28	−0.06	−0.22	15
down to 4 or 5	Best 5	−2.08			−1.64			1.38		

their upgraded 1's and 2's versus unchanged 3's in each month. We then computed the Spearman rank correlation coefficient between this ranking across all pairs of adjacent months. The average correlation coefficient was .10. Likewise, the rank correlation of the excess return on firms downgraded to 4 or 5 in adjacent months was on average .06. These are not very high and are insignificant. We tried one other way to see if superior brokerage firms could be selected. In each month we examined the differential return for the five firms ranked highest in the prior month. This did lead to some improvement in return (see Table V) compared to the sample as a whole. However, we can not claim statistical differences in the return. Furthermore, the results in the month of the announcement, while indicative of information, are not amenable to a trading rule since the best five firms cannot be determined until the end of the month.

In this section we have shown that there is substantial information in the change in analysts' forecasts, both in the aggregate and at the individual brokerage firm level. In the next section we will examine performance of the recommended list as a whole.

III. Brokers' Buy and Sell Lists

In the previous section of the article we examined the size of differential returns which could be earned if an investor took appropriate action (bought or sold stocks) when analysts changed their recommendations. However, there is a second question worth examining. How much information is conveyed by looking at the list of analysts' recommendations put out by brokers at a point in time?

This question is of interest because this is the way in which the brokerage firms are suggesting that customers use their data. In the previous section we showed that when an analyst changes his or her recommendation about a stock (reacts to new information), positive returns can be earned and that the information is absorbed into stock prices within three months. On the other hand, only 11.6% of recommendations are changed by analysts every month. This suggests that, on average, a stock stays in a particular recommended grouping for more than eight months. Thus, at any point in time, a recommendation category is going to contain a mixture of stocks that were placed there recently (and thus about which one has real information) and stocks which have been in the group a number of months. According to results of Section II, the latter should have no differentially positive return. Given the nature of the recommended groups, does it pay to act on the recommendations of brokerage firms at a point in time? This is the question to which we now turn. More specifically, in this section we shall examine whether excess returns can be made by following brokerage firm recommendations. In particular we shall examine whether holding the buy list versus holding the sell list, with risk adjusted appropriately, leads to excess return. We shall do this in two different ways. First we construct portfolios of buys and sells, holding beta and industry composition constant. This will enable us to comment on the impact of industry selection. Finally, we shall examine whether brokers who do well in recommendations one period have a tendency to do well in subsequent periods.

A. Performance of Recommended List

In this part we examine whether an investor could earn a differential return by selecting a brokerage firm at random and purchasing the stocks that are ranked 1 while selling those ranked 4's and 5's. To compare 1's with 4's and 5's, it was necessary to hold beta constant between the two groups, or else what appeared to be excess returns could be due simply to risk differentials. Following this we will examine what happens when we constrain the industry composition as well as the beta to be the same for both groups.

For each broker and for each quarter starting March, 1981, all stocks identified as 1's of the date under study were placed in one equally-weighted portfolio while all stocks labelled 4's and 5's were placed in a second portfolio. All 1's were compared with 4's and 5's rather than just 5's because of the small number of 5's identified by most brokers.

If the beta on these two portfolios for any broker differed by more than .01, then stocks were deleted from the appropriate extreme for the larger of the two portfolios until the betas differed by less than .01.

Table VI presents information about the differential return on 1's versus the return on 4's and 5's. The number can be considered the return from buying 1's rather than 4's and 5's with the beta on the overall portfolio remaining the same. All numbers in this table are expressed in terms of monthly differential returns. We see that the differential returns from buying 1's and selling 4's and 5's was .8% per month in the month the recommended list is completed and .57% per month in the following month. Both of these numbers are significantly different from zero at the 1% significance level. Standard deviations were computed across brokerage firms. The question we are examining is this: if an investor picked a brokerage firm at random and followed its recommendation, what are the odds of positive differential returns? Perhaps this is easier to see by noting that, in the month of the recommendation, 28 out of 34 brokerage firms had positive differential returns on average while in the following month it was 27 out of 34. One month after the recommendation 27 firms would produce positive differential returns for their customers while only 7 would produce returns below random selection. This is reasonably impressive performance.

Returning to the table, we see that stocks did not perform very differently in the month prior to the publication of a recommendation list than they did in the publication month. Remember that 78% of the stocks on the recommended list had the same classification two months earlier. From the earlier section we know that the extra return occurs in the month of change and in the subsequent months. At any point in time a recommendation list is a combination of changes that occurred in the prior month, two months ago, three months ago, etc. Thus, given the results discussed in an earlier section, we would expect positive differential returns in the months before we examine the recommended list. Any benefit that can be gained from buying a recommended list disappears within two months of compilation of the list.

Before examining the case where industry membership was held constant, it is worthwhile to compare the results of using the recommended list with results presented in Tables IV and V of analyzing the excess returns from a change in

Table VI

Overall Performance of a Randomly Selected Broker*

	Average in the 3 months before Listing	Month of Listing	One Month After Listing	Two Months After Listing
A. Non Industry Matched Samples				
Return	0.87	0.80	0.57	−0.32
t-Value	3.07	2.90	2.40	−1.07
Number plus	25/34	28/34	27/34	17/34
B. Industry Matched Sample				
Return	0.53	0.52	0.10	0.11
t-Value	3.80	3.87	0.87	0.77
Number plus	25/34	25/34	18/34	21/34

* Differential performance from selecting one broker and following its advice over the sample period—1's minus 4's and 5's.

recommendation. While Table V did not compare upgrades to 1's and downgrades to 4's and 5's, it did show upgrades to 1's and 2's compared to downgrades to 4's and 5's. The differential returns were 2.43% in the recommendation month and 1.86% in the next month. Looking at changes gives us returns that are 279% and 230% of the returns to be attained by acting on the recommended list. Clearly it pays to buy and sell a portfolio of securities based on changes in recommendation. However, enough of this information is captured by the contents of a firm's recommended list at any point in time that it conveys useful information.

B. Recommended List Performance With Industry Constraints

In this section we repeat the procedure of the last section except that we match high-rated stocks with low-rated stocks by industry classification as well as by beta.[7] Because we are selecting stocks within an industry and thus from a smaller sample, we often had to take a 2 rather than a 1 (i.e., no stock in the industry was rated a 1) or a 4 rather than a 5. It is logical to ask whether performance increases or decreases as we hold industry composition constant. This, in essence, eliminates industry betas from stock selection. There are two reasons for this further refinement. First, insofar as there are industry factors in an equilibrium model, holding industry classification constant accounts for this influence. Second, analysts generally follow an industry. Insofar as analysts have different ranking methods, this corrects the problem. Comparing the two parts of Table VI clearly shows that holding industry composition constant eliminates most of

[7] The algorithm for forming these groups was more complex than the previous algorithm. Industries were defined by 2 digit SEC code. An industry was selected at random. The stock with a beta closest to one was selected from the smallest group (high or low classified stock). Then the stock with the closest beta was selected from the larger group. This was repeated. When a stock was selected from the larger group an attempt was made to eliminate prior differences. After all industries were examined replacements were made if the differences in betas on an equally weighted portfolio was greater than .05.

the payoff from following the recommended list. While there is a statistically differential return earned in the month of the recommendation, it is smaller than in the non-industry-matched sample. In the month following the compilation of the recommended list, any differential return has all but disappeared. Most of the payoff from recommendation seems to come from differentiating between industries rather than from selecting individual issues.

C. Performance of Indivdiual Firms

The purpose of this section is to see whether an investor would be better off by following the recommended list of those brokers which did best in the past rather than selecting brokerage firms at random. Two types of tests were conducted. In the first test, the correlation between differential returns for brokerage firms in successive periods was calculated. In the second part the differential returns earned by following the recommendations of the top five brokers were compared with the return from the average brokerage firm. We review each test in turn.

In the first set of tests all brokerage firms were ranked by the size of the three month differential return (after the recommended list is determined). They were then ranked by the three month differential return after their next list was determined (one quarter later). Spearman rank correlations were calculated between all adjacent quarters. The average rank correlation when an industry constraint was not imposed was .10 while when an industry constraint was imposed it was .05. There does not seem to be very much correlation between performance in successive quarters.

It is still possible that some increase in return could be earned by selecting the very best brokerage firm even though the performance across all brokers show very little correlation from period to period. To examine this we computed the incremental return that would have been earned by following the advice of the five brokerage firms that performed best over the previous quarter rather than by selecting a brokerage firm at random. The added return for the non-industry constrained sample is .32% one month after the forecast and −.37% two months after the forecast. The differences are even smaller in the case of industry constraints. Not only are these differences statistically insignificant, but they were also of a very small order of magnitude. It does not appear that looking for the best brokerage firm list has any payoff.

IV. Conclusions

In this paper we have shown that there is information in brokerage firm recommendations, in that excess returns are found in the month during which there is a change in the brokerage firm recommendations, as well as during the next two months. There is a consistency across changes and across time that suggests that the results cannot be due to an extraneous factor. Buying the brokerage firm list leads to excess returns, but excess returns not as large as those to be earned by acting on changes. We can find no evidence that a superior brokerage firm can be identified or that one brokerage firm is consistently better than a second.

REFERENCES

1. Elroy Dimson and Paul Marsh. "An Analysis of Broker's and Analysts' Unpublished Forecasts of U.K. Stock Returns." *Journal of Finance* 39 (December 1984), 1257–1293.
2. Edwin Elton, Martin Gruber, and Mustafa Gultekin. "Professional Expectations: Accuracy and Diagnosis of Errors." *Journal of Financial and Quantitative Analysis* 19 (December 1984), 351–63.
3. Edwin Elton, Martin Gruber and Mustafa Gultekin. "Expectations and Share Prices." *Management Science* 27 (September, 1981), 975–87.
4. N. A. Gregory. "Testing an Aggressive Investment Strategy Using Value Line Ranks: A Comment." *Journal of Finance* 38 (March 1983), 257.

DISCUSSION

DENNIS E. LOGUE*: This paper explores the value of brokerage firm stock recommendations. It first asks whether significant changes of opinion regarding a stock have some information content. In other words, when analysts boost their evaluation of a stock, does the price rise? It then asks whether there is persistence to the recommendations. For instance, can an investor who buys the stocks presently rated "best" do better than by buying some other portfolios of stocks? Finally, it addresses the persistence of the forecasting ability of any brokerage firm. Can a specific broker continually pick those stocks that subsequently rise relative to other stocks?

The paper shows first that changes of opinion by analysts are correlated with subsequent stock price movements. Stocks elevated to the highest categories have superior risk-adjusted excess returns relative to those either not elevated or downgraded. What was not shown is whether the analysts had true forecasting ability or whether the strength of the marketing effort associated with upgrades led to the price rise. That is, are analysts really identifying positively altered company situations or are the salesmen simply convincing enough investors that company prospects have improved because the experts said so? I do not think we will ever get to untangle this question with the kind of data available, but it is worth pondering.

The authors also evaluate "buy lists." These contain securities that were recently upgraded, but preponderantly have securities that have been labelled "buy" for a while. The findings here are surprising. Buy lists relative to sell lists do very well on a risk adjusted basis. However, we do not know how well they would do if upgrades and downgrades were excluded. The high returns might result from the impact of the upgrades.

Finally, the authors ask whether there is a "best" broker or group of brokers. Specifically, they choose the five best from a prior period, then see if their ability to pick stocks continues into the future. The answer is no: last month's "hot" brokers are not likely to be this month's.

* Amos Tuck School, Dartmouth College.