ELSEVIER Journal of Financial Economics 42 (1996) 397–421

JOURNAL OF Financial ECONOMICS

Market timing ability and volatility implied in investment newsletters' asset allocation recommendations

John R. Graham[a], Campbell R. Harvey*[,b,c]

[a]*David Eccles School of Business, University of Utah, Salt Lake City, UT 84112, USA*
[b]*Fuqua School of Business, Duke University, Durham, NC 27708, USA*
[c]*National Bureau of Economic Research, Cambridge, MA 02138, USA*

(Received October 1994; final version received February 1996)

Abstract

We analyze the advice contained in a sample of 237 investment newsletter strategies over 1980–1992. Each newsletter strategy recommends a mix of equity and cash. We find no evidence that letters systematically increase equity weights before market rises or decrease weights before market declines. While there is no information in the newsletter strategies about future market returns, we document that disagreement among the newsletters is correlated with future realized and implied volatility.

Key words: Market timing; Investment newsletters; Performance evaluation
JEL classification: G12; G14

1. Introduction

This paper provides the first comprehensive analysis of investment newsletter recommendations. We evaluate the performance of 237 newsletter strategies

*Corresponding author.

The second author's research was supported by the Batterymarch fellowship. We have benefited from the comments of Stephen Brown, Wayne Ferson, Doug Foster, Will Goetzmann, Shmuel Kandel, Gautam Kaul, Chris Kirby, and Jaime Zender, as well as the seminar participants at the University of Utah and the 1995 Western Finance Meetings. We appreciate the detailed comments of the editor, Jerry Warner, and an anonymous referee. We also are grateful to Mark Hulbert for providing the data and answering many of our questions. The New York Stock Exchange provided volume data.

SSDI 0304-405X(96)00878-1

from June 1980 to December 1992. These newsletters recommend investment weights for equity. Rather than selecting specific stocks, the newsletters attempt to call the direction of the market as a whole.

Our paper investigates whether the newsletters offer any market timing ability. Timing implies that excess returns are positive after recommended increases in equity weights and negative after recommended decreases in equity weights. We find that the newsletters fail to offer advice consistent with market timing. When newsletters recommend an increase in equity weights, the subsequent one-month market return (in excess of the riskless rate) is positive 70.4% of the time. However, when a decrease in equity weights is recommended, the subsequent one-month market return is positive 69.4% of the time. This implies that, in the aggregate, changes in recommended investment weights do not provide information about future market returns.

Our analysis of individual newsletter performance does not reveal any systematic evidence that their long-term returns exceed those of a passive benchmark. We do, in contrast, find that newsletters with a run of correct recommendations (this is sometimes called 'hot hands') provide potentially valuable information about future market returns. However, while some newsletters appear to have short-term insights, an investor cannot use a 'hot streak' to identify a specific newsletter that will provide superior returns in the long term.

Finally, we investigate the aggregate information contained in the cross-section of newsletters' forecasts. Using the recommended asset weights, we infer each letter's forecasted market return by assuming an exponential utility function. We allow risk aversion to differ across newsletters, but assume that it is constant through time. We use these forecasts to explore the information in the dispersion of newsletter forecasts. Our evidence suggests that dispersion predicts future realized volatility, future implied volatility, and future trading volume.

Our paper is organized as follows. The second section details the nature of the data. Direct measures of market timing are investigated in the third section, which also addresses the relation between forecast dispersion, volume, and volatility. Some concluding remarks are offered in the final section.

2. Data

We have data from the *Hulbert Financial Digest* on 101 investment newsletters beginning in June 1980 and ending in December 1992. Since some newsletters offer more than one investment strategy, there are a total of 237 newsletter strategies. Hulbert compiles data on a broad set of newsletters that provide well-defined recommendations. A recommendation is a proposed portfolio composition in which recommended long equity plus short equity plus cash less margin equals 100%. In almost all cases, the nonequity category is cash. In some cases, it may be fixed income. However, to simplify the analysis, we assume that the nonequity investment is represented by the 30-day Treasury bill.

The date an observation is added to the raw file is the date Hulbert receives it in the mail or over the phone for letters with free hotlines, rather than the date published on the newsletter. If the letter has a free hotline, Hulbert calls this number each day to supplement the recommendations received by mail. Also, if the letter has previously expressed a 'stop-loss' position, such as selling if the Dow Jones Industrial Average reaches 6,000, Hulbert implements this as a recommendation if the condition occurs.

Our data has none of the survivorship problems related to letters dropping out of the sample. Newsletters are added on the day Hulbert first receives the letter; no data are deleted when a newsletter ceases to exist. This contrasts with the acute survivorship bias in most previous mutual fund studies (see discussion and references in Brown, Goetzmann, Ibbotson, and Ross, 1992). Indeed, if we required the newsletter to exist for the entire sample, we would be left with only 13 of the 237 newsletter strategies.

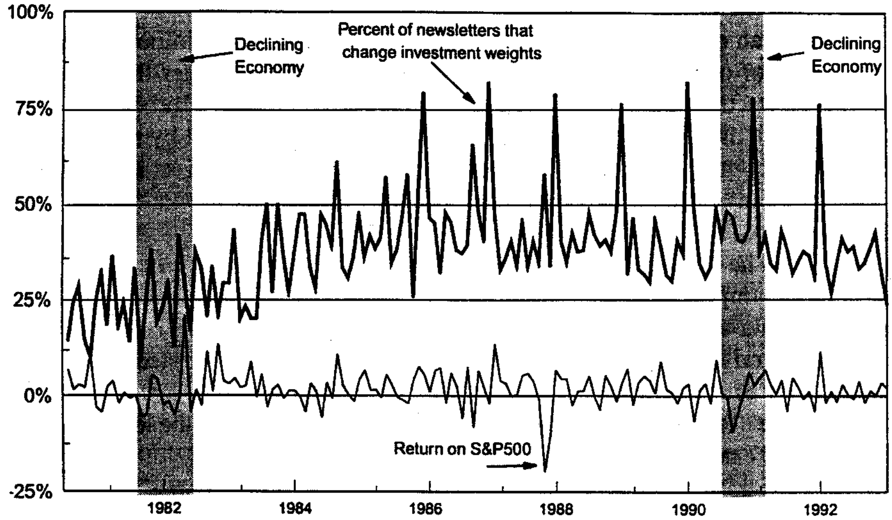
There are 15,133 total recommendations across all newsletters. An observation can occur on any day during a month, and multiple observations may occur in any month. However, for our tests, we concentrate on monthly recommendations. This allows us to link our work to the growing literature on conditional performance measurement, which utilizes monthly data. To this end, we use the last observation in a month as our 'monthly' asset weight recommendation. Later, we assess the sensitivity of our results to this assumption in separate estimation that uses daily S&P 500 returns and acts on the recommendation the day the investment letter is received.

3. The information in newsletter recommendations

3.1. Changing investment weights – Aggregate analysis

Investment letters frequently change their recommended positions. Panel A in Fig. 1 shows the time series of changes in newsletter equity weights. There is a 53% chance that the recommended investment weights will change in any month. There is a 75% chance that the investment weights will change in December. Panel B suggests that there are some distinct time-series patterns in the average market exposure across newsletters. In particular, equity weights are lowest in recessions. During the 1981–82 recession, the average market weight was only 20%. During the recovery and expansion that followed, the market exposure increased to 86% at the end of 1985 and then began a slow decline. The average market weight bottomed out during the most recent recession (July 1990–March 1991) at 28%. Over 1992, the average market weight was 54%. The letters' average equity weights are affected by the stage of the business cycle.

A: Times series of changes in investment newsletter equity weights



B: Times series of the average equity position across all newsletters

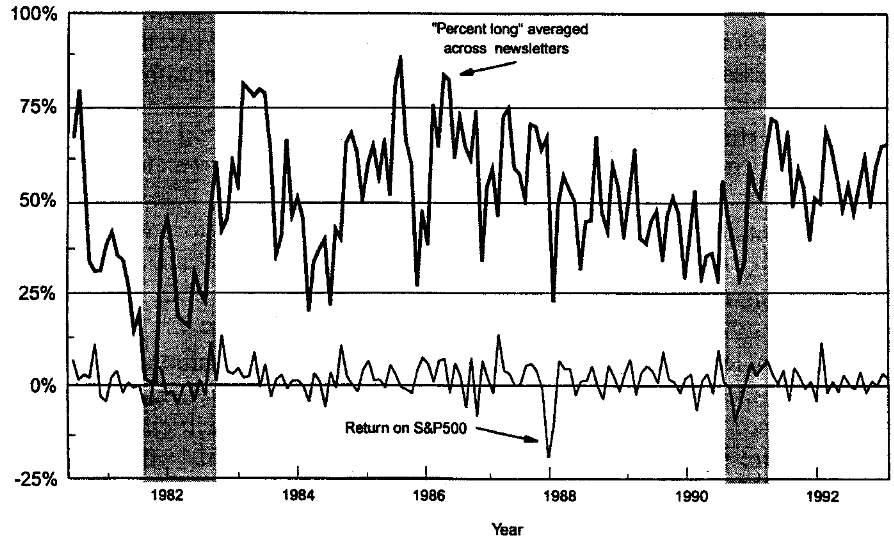


Fig. 1. Time series of investment newsletter equity weights.

Out of the sample of 237 newsletter strategies that give recommendations in any given month, panel A shows the percent that change their position from the previous month over the period June 1980 to December 1992. The monthly return on the S&P 500 (cash index prior to May 1982, futures index starting in May 1982) and the July 1981–November 1982 and July 1990–March 1991 recession periods, as defined by the NBER, are also shown. Panel B shows the mean equity weight recommendation.

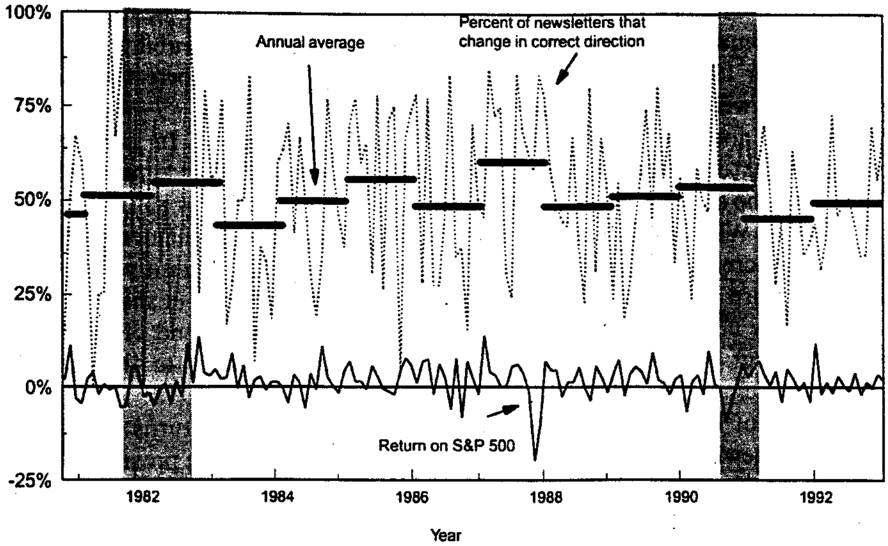


Fig. 2. Correct changes in recommended investment weights.

The figure shows the monthly time series of the percentage of newsletter that change their investment weights in the correct direction (i.e., in the same direction as the one-month-ahead market movement). The thick lines are the annual averages of the percentages. The shaded bars indicated periods of economic recession, as defined by the NBER (July 1981–November 1982 and July 1990–March 1991). The monthly return on the S&P 500 index (cash index prior to May 1982, futures index starting in May 1982) is also shown.

We now explore whether investment newsletters anticipate the market direction. The market return for our purposes is the S&P 500 futures index. We choose the futures index because it has relatively low transactions costs. This is especially important if the investment weights are frequently altered. Fig. 2 shows the percent of newsletters that change investment weights in the correct direction (i.e., in the same direction as the one-month-ahead market movement) over the June 1980–December 1992 period. The overall average is 50.1% and there appears to be random variation over time. A random investment strategy (50% increased weights, 50% decreased weights) would produce a 50% hit rate. The overall hit rate of 50.1% is statistically indistinguishable from a success rate generated by random investment strategy and indicates an inability to predict the market.[1]

The best year for newsletters was 1987 when on average 64% of the newsletters changed investment weights in the correct direction. Of the 94 newsletter

[1]If a letter increases weight 69.9% of the time (the average percent of positive excess returns for the months in our sample) but the selection is still random, the hit rate under the null hypothesis is 59.8% which is far above the observed rate of 50.1%.

strategies available for September 1987, 14 advocated increased equity weights and 21 recommended a lower equity exposure, revealing some ability to predict the October 1987 market crash. The percent of newsletters correctly changing investment weights dropped below 50% in 1991 and 1992.

We also look at higher-frequency investment strategies. In most of our analysis, we assume that the newsletters make monthly recommendations. However, the choice of monthly versus daily recommendations has no effect on our results. When the newsletter recommendations are implemented on a daily basis, an increased (decreased) equity weight followed by a positive (negative) market return in the 20-day period after the recommendation (a shorter period is used only if a new recommendation is released before the end of the 20-day period) occurs 49.3% of the time, which is slightly inferior to the monthly performance.

Panel A of Fig. 3 shows a scatter plot of the S&P 500 returns against the percent of newsletters that increased investment weights in the previous period. Each point on the graph represents a month in our sample. If newsletters correctly anticipate market upturns, there should be a positive relation. However, the correlation is only 0.027 and is not significantly different from zero. Panel B examined the S&P 500 return against the percent of newsletters that decrease investment weights. If the newsletters as a group correctly anticipate market declines, we would see a negative relation. While the results indicate that the correlation is negative, -0.038, it is not significantly different from zero.

When we focus on the market performance after recommended increases or decreases in weights, there is little evidence of market timing in Table 1. Here we examine S&P 500 returns in excess of the one-month Treasury bill rate. When using the S&P 500 futures, the excess return is defined as the percentage change in the S&P 500 futures index. For recommended equity weight increases, the subsequent one-month market return in excess is positive 70.4% of the time; the mean annualized excess return is 14.7%. The excess return is positive 69.9% of the time for all observations in our sample. For recommended equity weight decreases, the subsequent one-month market excess return is positive 69.4% of the time and the mean annualized excess return is 15.6%. A comparison of the percentages reveals no significant difference between 70.4% and 69.4%.

A one-month horizon may not be long enough to evaluate the ability to anticipate market direction. The second set of columns in Table 1 tracks the market return for six months following the changes in equity weights. However, we find that the future market return is more likely to be positive after decreases in recommended increased equity weights. The mean annualized six-month return is 12.7%, while the mean excess return following decreased equity weights is 16.2%. This is the opposite to what we expect if newsletters appropriately time the market.

We also examine how well the investment newsletters anticipate large market movements. Here, we examine the changes in weights before absolute market returns that are greater than one standard deviation. The results reveal weak

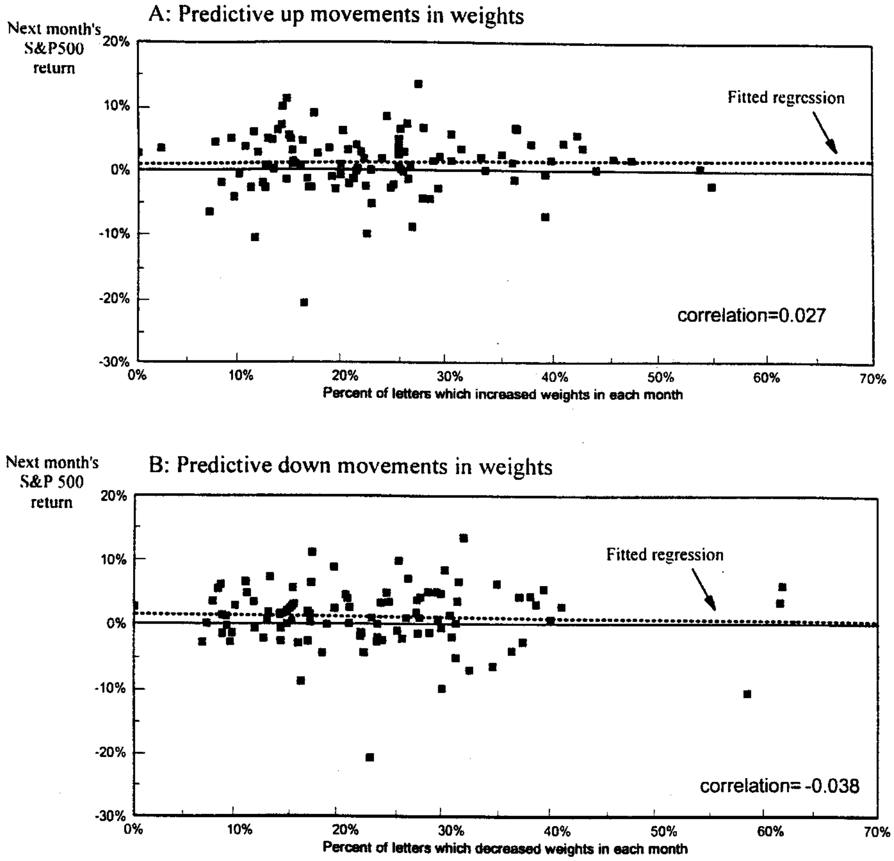


Fig 3. Predictive changes in recommended investment weights.

Panel A (B) shows the relation between positive (negative) changes in equity weights versus market returns; a positive (negative) relation indicates market timing ability. There is one observation for each month of the sample (May 1982 to December 1992). The percent of active letters which increased (decreased) equity weights in any given month appears on the horizontal axis in panel A (B). The corresponding one-month-ahead return on the S&P 500 futures index appears on the vertical axis of each graph.

evidence of market timing. Positive (negative) weight changes are followed by positive market excess returns 71.3% (68.1%) of the time, a difference that is significant at the 10% level.

3.1.1. Persistence and 'hot hands'

In the middle rows of Table 1, we condition on whether the newsletter's most recent recommendation was correct or incorrect. For those with correct past

Table 1
Measuring the ability of investment newsletters to anticipate market direction

Movements in the S&P 500 futures index are examined following changes in newsletter recommendations. The top portion of the table studies market movements after increases ($\Delta w_t > 0$; see row 2) and decreases ($\Delta w_t < 0$; see row 3) in recommended equity weights. The middle section explores market movements conditional on whether a newsletter correctly called the direction of the market in its last recommendation. The bottom section examines market movements after recommendations made by newsletters that have a 'hot hand' (i.e., they correctly anticipated the direction of the market in their last three recommendations) versus letters that have a 'cold hand' (i.e., they incorrectly anticipated the direction of the market in their last three recommendations). These scenarios are examined for the proportion of time the market increases and the excess return for a one-month horizon (columns one and two), a six-month horizon (columns three and four), and a one-month horizon for observations in which the absolute value of the market return is greater than one standard deviation. The excess return is defined as the percentage change in the S&P 500 futures index.

	Anticipating short-term return[a]		Anticipating longer-term return[b]		Predicting large movement[c]	
	Next month's S&P 500 excess return		Next six months' S&P 500 excess return		Next month's S&P 500 excess return	
	% > 0	mean	% > 0	mean	% > 0	mean
All observations:	69.9%	0.152	68.6%	0.145	69.6%	0.332
All observations in which recommended equity weights						
increased ($\Delta w_t > 0$)	70.4	0.147	68.2	0.127	71.3	0.380
decreased ($\Delta w_t < 0$)	69.4	0.156	69.0	0.162	68.1	0.293
(p-value)[d]	(0.220)	(0.295)	(0.265)	(0.495)	(0.095)	(0.091)
Given last recommendation was correct, recommended equity weights						
increased ($\Delta w_t > 0$)	75.5	0.186	72.2	0.164	73.0	0.417
decreased ($\Delta w_t < 0$)	65.9	0.133	65.9	0.135	67.3	0.251
(p-value)[d]	(0.001)	(0.023)	(0.004)	(0.096)	(0.057)	(0.052)
Given last recommendation was incorrect, recommended equity weights						
($\Delta w_t > 0$)	66.7	0.120	65.3	0.099	70.9	0.364
($\Delta w_t < 0$)	70.9	0.139	70.0	0.155	63.2	0.217
(p-value)[e]	(0.013)	(0.218)	(0.007)	(0.003)	(0.015)	(0.045)
Given last 3 recommendations were correct, recommended equity weights						
($\Delta w_t > 0$)	76.0	0.196	71.8	0.154	71.7	0.375
($\Delta w_t < 0$)	61.3	0.028	63.7	0.059	55.5	−0.092
(p-value)[d]	(0.001)	(0.001)	(0.019)	(0.021)	(0.024)	(0.009)
Given last 3 recommendations were incorrect recommended equity weights						
($\Delta w_t > 0$)	65.9	0.088	63.6	0.093	69.6	0.380
($\Delta w_t < 0$)	72.7	0.129	69.8	0.143	68.6	0.386
(p-value)[e]	(0.038)	(0.169)	(0.056)	(0.078)	(0.457)	(0.489)

[a]Annualized one-month return on the S&P 500 futures index. The sample begins in May 1982, which is the first month a return on the S&P 500 futures index data is available.

[b]Annualized six-month return on the S&P 500 futures index.

recommendations, a positive change in equity weight is followed by a positive excess return 75.5% of the time. In contrast, a negative change in equity weight is followed by a positive return 65.9% of the time. Similar results are found for the six-month horizon and the large movements sample. The statistical tests show significant differences between the proportions following the increased and decreased equity weights. While this analysis shows mild evidence of market timing, it should be interpreted cautiously given that the market excess return is positive 65.9% of the time following recommended reductions in equity weights.

The results for the newsletters whose last recommendation was incorrect are also reported in Table 1. Given a previous incorrect recommendation, an increased recommended weight is followed by a positive market excess return 66.7% of the time. Decreased recommended equity weights are followed by a positive market excess return 70.9% of the time. An investor is better off betting against these newsletters.

The final part of the table examines 'hot hands' and 'cold hands'. We define hot (cold) hands as those letters which correctly (incorrectly) anticipated the direction of the market in their last three recommendations. The results are consistent with the patterns that emerge from conditioning on whether the previous investment recommendation was correct. For the hot hands sample, the one-month market excess return is positive 76.0% of the time following recommended increased equity weights, with an annualized average excess return of 19.6%. The one-month market excess return is positive 61.3% of the time following recommended decreases in equity weights with an average annualized excess return of only 2.8%. These proportions are more impressive for the large movement sample. For this sample, the average excess return following recommended decreases in investment weights is -9.2%.

The cold hands sample is consistent with the patterns seen when we conditioned on the previous return being incorrect. The market excess return is positive 65.9% of the time following increased recommended weights, with an average return of 8.8%. The market excess return is positive 72.7% of the time following recommended decreases in equity weights with a mean return of 12.9%.

Footnotes to Table 1 (continued)

[c]Annualized one-month returns on the S&P 500 futures index which are greater in absolute value than the annualized standard deviation of the excess return on the S&P 500 futures index.

[d]p-value for a one-tailed ANOVA F-test testing the null hypothesis that the mean values in the two rows above are equal against the alternative hypothesis that the value associated with $\Delta w_t > 0$ is greater than the value for $\Delta w_t < 0$. A value of 0.05 or smaller indicates that the null is rejected in favor of the alternative at a 5% level of significance.

[e]p-value for a one-tailed ANOVA F-test testing the null hypothesis that the mean values in the two rows above are equal against the alternative hypothesis that the value associated with $\Delta w_t < 0$ is greater than the values for $\Delta w_t > 0$.

When viewing newsletters performance in the aggregate, there is little evidence of market timing. However, if an investor finds a letter with 'hot hands', there is some evidence that these recommendations contain information about future market returns. Of course, to achieve the 'hot hands' return described above, an investor would need to subscribe to a large portion of the newsletters in our sample. All the same, our analysis so far implies that, *if the hot hands phenomenon persists*, there may be some individual newsletters that are superior. In the next section, we investigate whether there are individual newsletters that are superior in the long term, and find no evidence that this is the case. That is, one cannot select a long-run superior newsletter by choosing a letter with a hot streak.

3.2. Market timing tests for individual newsletters

We test the market timing ability of each newsletter by estimating the model:

$$r_{m,t+1} = \delta_{i,1} + \delta_{i,2}\Delta w_{i,t} + \delta_i'\mathbf{Z}_t + \varepsilon_{i,t+1}\,, \tag{1}$$

where $\Delta w_{i,t}$ represents the change in net equity position at the end of month t (sampling only the months when weights changed) and $\mathbf{Z}_t$ is a set of common information variables available to all investors at time t. If the coefficient $\delta_{i,2} > 0$, on average the newsletter is increasing (decreasing) equity weights before the market excess return is positive (negative).

The information variables in (1) are designed to control for time variation in expected returns. The $\mathbf{Z}_t$ includes the lagged excess return on the Center for Research in Security Prices (CRSP) NYSE equally weighted index (persistence in returns), a January dummy variable, the lagged excess return on a three-month Treasury bill (expected inflation), the lagged Moody's Baa–Aaa yield spread (default risk), and the lagged excess dividend yield on the S&P 500 index (mean reversion). These variables have been shown in a number of papers to capture variation in expected returns (see Fama and French, 1989; Harvey, 1989).

We also investigate a model identical to (1) using $w_{i,t} - \bar{w}_{i,t-1}$ as an explanatory variable, where $\bar{w}_{i,t-1}$ is the average equity weight for newsletter i up to time $t-1$. This tells us whether the letter has a higher (lower) equity weight relative to its average when the market return increases (decreases). This may be a better test of market timing in some instances. For example, a newsletter with a recommended equity weight of zero might not choose to lower its weight any further. The $\Delta w_{i,t}$ variable will not pick up this observation. However, the $w_{i,t} - \bar{w}_{i,t-1}$ specification indicates whether the equity weight is correctly below its average return.

We also investigate differential abilities to time the market in up and down states. We estimate an indicator regression, inspired by Henriksson and Merton

(1981), that allows us to measure differential responses:

$$\Delta w_{i,t} = \delta_{i,3} I(r^{+}_{m,t+1}) + \delta_{i,4} I(r^{-}_{m,t+1}) + \varepsilon_{i,t+1}\,, \tag{2}$$

where $I(r^{+}_{m,t+1})$ is an indicator variable that takes on a value of one when market returns are positive. Essentially, (2) tells us the average increase (decrease) in equity weights when the market rises (falls).

The results of estimating (1) and (2) are presented in Table 2. The first panel shows that the coefficient on the weight variable is negative, albeit insignificantly so, for the regression with Δw_t defined over the pooled time-series cross-section of investment weights. A negative coefficient indicates that equity weights are increasing (decreasing) before the future market return is negative (positive). Furthermore, newsletter by newsletter, the coefficient is positive for only 43.5% of the regressions and significantly positive in only 8.3% of the estimates when testing at the 5% significance level.

Similar results are obtained when the investment weights minus their mean values are used as an explanatory variable. The slope coefficient for the pooled newsletter regression in Table 2 is negative, which is the wrong sign. Furthermore, for the individual investment letters only 45.0% of the sample had positive coefficients and only 7.6% are significantly positive. The lack of market timing inference does not change if the information variables, Z_t, are dropped from the specification in (1).

The results for the indicator variable specification in (2) are presented in panel B of Table 2. In the pooled newsletter regression, the point estimates of the coefficients suggest that market weights increase by 0.5% before positive market returns and decrease by 0.4% before negative market returns. While these are the correct signs, neither of these coefficients are statistically different from zero. Of the individual investment letters, 52.1% increased their recommended market weights (2.4% significant) before positive returns and 50.5% decreased weights (3.4% significant) before negative returns.

When the investment weight minus their average values are used as the dependent variable, one of the pooled newsletter coefficients has the wrong sign while the other is insignificantly different from zero. Of the individual letters, 49.2% are above the average weight when returns are positive and 53.1% are below the average weight when the returns are negative.

The regression analysis suggests that there are few, if any, individual newsletters that are statistically superior. We check this inference using a nonparametric Monte Carlo analysis that makes no assumptions about the distribution of market returns. The analysis consists of 500 simulations for each newsletter, where a single simulation calculates a hypothetical return for a letter based on a random ordering (without replacement) of its recommended investment weights. The newsletter's actual return is then compared to the distribution of 500 simulated returns. If a letter's actual return is larger than 90% of the simulated returns, the newsletter is assigned a p-value of 0.90, and is said to be

Table 2
Tests of market timing and extra market timing

Coefficients for the following regressions indicate whether there is market timing ability in the recommended investment weights of investment newsletters. The reported coefficient(s) is from a pooled regression using all 237 newsletters strategies. Also reported is the percent of the coefficients that are positive in 237 regressions, one for each newsletter. In each panel of this table, the 'changes in equity weights' and 'equity weights relative to mean' are separate regressions. In the latter regressions, the mean is calculated on a rolling historical average, where the historical period over which the weights are averaged includes all monthly recommendations for newsletter i which occurred up to year $t-1$. p-values are shown in brackets below the estimated coefficients.

A: $r_{m,t-1} = \delta_{1,i} + \delta_{2,i}W_{i,t} + \delta_{3,i}\boldsymbol{Z}_t + \varepsilon_{i,t+1}$, $i = 1, 2, \ldots, 237$, where $r_{m,t+1}$ is the one-month-ahead return on the S&P 500 futures index, $W_{i,t}$ is defined below in the table, and $\boldsymbol{Z}_t$ is a set of public information variables available at time t: the lagged excess return on the CRSP equally weighted NYSE index, the lagged excess return on a three-month Treasury bill, the lagged Moody's Baa–Aaa yield spread, the lagged excess dividend yield on the S&P 500 index, and a January dummy variable. A positive δ_2 coefficient indicates market time ability.

	Pooled regression	Individual regressions	
Investment weight specification	δ_2	% with $\delta_{2,i}$ positive	% $\delta_{2,i} > 0$ and sig.[a]
Changes in equity weights: $W_{i,t} = \Delta w_{i,t}$	− 0.00009 [0.312]	43.5%	8.3%
Equity weights relative to mean: $W_{i,t} = w_{i,t} - \bar{w}_{i,t-1}$	− 0.00002 [0.210]	45.0%	7.6%

B: $W_{i,t} = \delta_{3,i} I(r^+_{m,t+1}) + \delta_{4,i} I(r^-_{m,t+1}) + \varepsilon_{i,t+1}$, $i = 1, 2, \ldots, 237$, where $I(r^-_{m,t+1})$ and $I(r^-_{m,t+1})$ are dummy variables equal to one when the market return is positive and negative, respectively, and $W_{i,t}$ is defined below in the table. A positive (negative) coefficient on the δ_3 (δ_4) coefficient indicates market timing ability.

	Market up			Market down		
	Pooled	Individual $\delta_{3,i}$ pooled		Individual $\delta_{4,i}$		
Investment weight specification	δ_3 % > 0	% > 0, sig.[b]		δ_4	% < 0	% < 0, sig.[b]
Changes in equity weights: $W_{i,t} = \Delta w_{i,t}$	0.478 [0.720]	52.1%	2.4%	− 0.377 [0.773]	50.5%	3.4%
Equity weights relative to mean: $W_{i,t} = w_{i,t} - \bar{w}_{i,t-1}$	− 2.61 [0.033]	49.2%	15.6%	− 1.36 [0.293]	53.1%	14.4%

C: $W_{i,t} = \theta_{i,2} I(r^+_{m,t+1};\ \text{unexpected}) + \theta_{i,3} I(r^-_{m,t+1};\ \text{unexpected}) + \theta_{i,4} I(r^+_{m,t+1};\ \text{expected}) + \theta_{i,5} I(r^+_{m,t+1};\ \text{expected}) + \varepsilon_{i,t+1}$, $i = 1, 2, \ldots, 237$. The $I(.,.)$ are indicator variables which take on a value of one when their conditions are met and a value of zero otherwise. For example, $I(r^+_{m,t+1};\ \text{unexpected})$ has a value of one when the market return one period hence is positive but the common expectation of the market return is that is will be negative. The common expectation of the market return is determined from the regression of the market return on the investment variables detailed in panel A. The lower portion of this panel runs a separate set of regressions using an equation analogous to the equation in the top portion of this portion of this table, except that a third conditioning variable is added to the indicator functions, namely the direction of expected volatility. The 'N' or 'n' ('P' or 'p') superscript indicates 'unambiguously negative' ('positive') throughout panel C. See the footnotes for further details.

	Pooled	individ.	Pooled	individ.	Pooled	individ.	Pooled	individ.
	θ_2	$\%\theta_{i,2} > 0$	θ_3	$\%\theta_{i,3} < 0$	θ_4	$\%\theta_{i,4} > 0$	θ_5	$\%\theta_{i,5} < 0$
Investment weight specification	Mkt unexpected ↑		Mkt unexpected ↓		Mkt expected ↑		Mkt expected ↓	
Changes in equity weights: $W_{i,t} = \Delta w_{i,t}$	3.00^p [0.352]	50.9%	2.13^n [0.198]	42.4%	-0.75^p [0.534]	51.4%	-4.15^n [0.049]	56.2%
Equity weights relative to mean: $W_{i,t} = w_{i,t} - \bar{w}_{i,t-1}$	-1.42^p [0.636]	40.4%	2.04^n [0.206]	47.3%	-3.17^p [0.005]	49.4%	-6.50^n [0.002]	54.3%

Table 2 (continued)

Investment weight specification	Pooled θ_2 Mkt unexpected ↑	individ. $\%\theta_{i,2} > 0$	Pooled θ_3 Mkt unexpected ↓	individ. $\%\theta_{i,3} < 0$	Pooled θ_4 Mkt expected ↑	individ. $\%\theta_{i,4} > 0$	Pooled θ_5 Mkt expected ↓	individ. $\%\theta_{i,5} < 0$
Changes in equity weights: $W_{i,t} = \Delta w_{i,t}$								
Expected volatility increase:	2.14 [0.357]	51.6%	1.51^N [0.542]	44.6%	−3.91 [0.046]	50.8%	-1.91^N [0.576]	50.8%
Expected volatility decrease:	7.36^P [0.158]	64.7%	4.02 [0.353]	39.0%	3.47^P [0.125]	55.0%	−7.42 [0.073]	61.5%
Equity weights relative to mean: $W_{i,t} = w_{i,t} - \bar{w}_{i,t-1}$								
Expected volatility increase:	−4.25 [0.051]	39.7%	1.13^N [0.626]	45.5%	−4.32 [0.015]	46.6%	-4.84^N [0.128]	50.4%
Expected volatility decrease:	13.82^P [0.006]	50.7%	4.81 [0.231]	50.0%	-1.47^P [0.488]	49.3%	−9.01 [0.021]	58.6%

[a]The percentage of individual $\delta_{2,i}$ coefficients that are positive and significant at a 5% level.

[b]The percentage of individual coefficients that have the correct sign and are significant at a 5% level.

[n]Predicted to be negative because the excess market return in period $t+1$ is negative.

[p]Predicted to be negative because the excess market return in period $t+1$ is positive.

[N]Predicted to be unambiguously negative, in the sense that the expected change in volatility and the realized market return give reinforcing implications that optimal investment weights should be declining in the changes in weights specification and less than the mean in the weights relative to mean specification.

[P]Predicted to be unambiguously positive, in the sense that the expected change in volatility and the realized market return give reinforcing implications that optimal investment weights should be increasing in the changes in weights specification and greater than the mean in the weights relative to mean specification.

a superior performer at a 10% level of significance. Only 11 newsletter strategies are deemed superior by this experiment at the 10% level, compared to an expected number of 23. This suggests that individual newsletters are not superior performers. In contrast, 65 newsletters strategies are in the lower 10% tail of the distribution, implying that there are significantly inferior individual newsletters (see Fig. 4).

The results thus far suggest that the letters offer little market timing information. Following Grinblatt and Titman (1989), we also formed portfolios based on annual rather than monthly newsletter recommendations. If there is value to the monthly advice, the portfolio returns based on monthly updating should produce higher returns. Our results (not reported) indicate that in only 47.3% of the cases is there a loss associated with annual updating. Next, we more explicitly control for both conditioning information and time-varying volatility in equity returns. While (1) controls for time variation in expected returns, the indicator regression (2) does not. Neither specification controls for changes in market volatility. This may present a problem because an investment newsletter could reduce equity weight solely as a result of an anticipation of higher volatility.

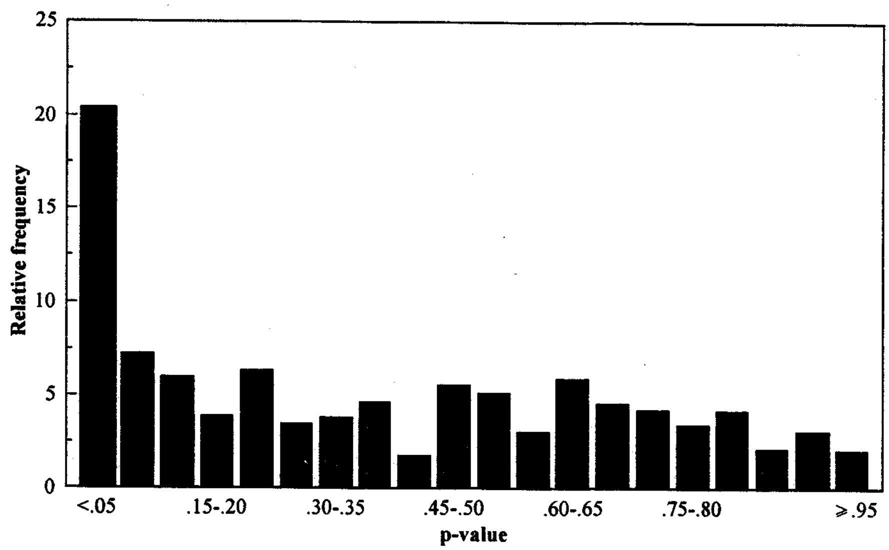


Fig. 4. Distribution of p-values based on Monte Carlo analysis.

This histogram summarizes the p-values from a Monte Carlo analysis of 500 simulations for each newsletter. A single simulation calculates a hypothetical return for a newsletter based on a random ordering (without replacement) of its recommended investment weights. The p-value is equal to the position of the newsletter's actual return in the distribution of 500 hypothetical returns. For example, if a letter's actual return is larger than 90% of the hypothetical returns, the newsletter is assigned a p-value of 0.90 (p-values below 0.05 indicate significantly poor market timing at the 5% level).

3.3. Market timing and conditioning information

Most performance evaluation studies have been executed within the paradigm of constant risk (for the underlying securities) and constant risk premiums. Viewed in this context, trading rules based on publicly available information could earn abnormal returns.[2] However, if we move away from the constant risk/constant risk premium framework, some predictability in returns may naturally arise.

We measure the timing skill over and above the common level of timing inherent in the base-line predictability. Extra market timing occurs when (i) a newsletter correctly anticipates the direction of the market and (ii) the common expected return does not correctly anticipate the direction. Extra market timing is the ability to outperform the common market forecast.

The following model provides a test for extra market timing:

$$\begin{aligned} \Delta w_{i,t} = {} & \theta_{i,2} I(r^{+}_{m,t+1}\text{: unexpected}) + \theta_{i,3} I(r^{-}_{m,t+1}\text{: unexpected}) \\ & \theta_{i,4} I(r^{+}_{m,t+1}\text{: expected}) + \theta_{i,5} I(r^{-}_{m,t+1}\text{: expected}) + \varepsilon_{i,t+1}\,. \end{aligned} \tag{3}$$

If the newsletters' forecasts are at least as good as the common forecast (holding conditional variances constant), $\theta_4 > 0$ and $\theta_5 < 0$. Positive values of θ_2 and negative values of θ_3 indicate extra timing ability. Likewise, coefficients $\theta_2 > 0$ and $\theta_3 < 0$ from a regression analogous to (3), but with $w_{i,t} - \bar{w}_{i,t-1}$ as the dependent variable, indicate that the weights are above or below average at times which correctly defy the common market expectation.

The results of estimating a pooled version of (3) are presented in panel C of Table 2. First, consider whether the newsletters correctly interpret the economy-wide information (θ_4 and θ_5). The signs on the coefficients are correct. However, only 51.4% of the individual letters increase market weights before market rises in which the common forecast was realized. Interestingly, 56.2% decrease weights before correctly anticipated market declines. Similar results are obtained when the investment weight relative to its average is used. Only 49.4% of the letters have weights above average before correctly anticipated market rises. However, 54.3% of the letters and weights below average when the common expected and realized market returns were negative.

There is no evidence that the investment letter portfolios exhibit any extra market timing (θ_2 and θ_3). Portfolio weights increase in 50.9% of the cases when the realized returns are positive and the expected market returns are negative.

---

[2]For example, a statistical model based on regressing the market return on the information variables, **Z**, in (1), could be used to design the following strategy: 100% equity if the predicted excess return is positive and 0% equity if the predicted excess return is negative. This strategy produces a 100bp extra annual return compared to a volatility-matched constant-weight benchmark. Studies that use conditioning information include Ferson and Schadt (1996), Chen and Knez (1996), and Bansal and Harvey (1996).

Portfolio weights decrease in only 42.4% of the cases when the realized market return is negative and the expected return is positive. Similar results are obtained when the weights are measured relative to their average.

Of course, if the market variance is forecasted to increase, a newsletter may decrease its weight in the market even if the excess market return is expected to be positive. The final part of Table 2 allows for both changing conditional means and variances. We use a GARCH (1,1) specification (see Engle, 1982; Bollerslev, 1986) in which the same information variables in (1) are allowed to influence the conditional mean. The number of indicator variables is increased to eight (four for expected volatility increases and four for expected volatility decreases). Four of the indicators have unambiguous signs. For example, if the market is expected to decline and volatility is forecasted to increase, then the equity weight should always decrease.

Allowing for time-varying volatility marginally improves the performance of the newsletter recommendations. When volatility is expected to decline and the market rises as expected, 55.0% of the newsletters increase equity weights (θ_4). When volatility is expected to rise and the market falls as expected, 50.8% of the newsletters decrease investment weights on average (θ_5). When the weights relative to their mean values are examined, the performance deteriorates. For example, only 49.3% of the letters have recommended weights above average prior to an increasing market when volatility is expected to decline.

Overall, controlling for time-varying volatility does not change our conclusion that newsletters lack extra market timing ability. When volatility is expected to decrease and the market return is unexpectedly positive (θ_2), 64.7% of the letters increase investment weights on average. While this appears impressive, when the weights are measured relative to their average level, only 50.7% of the letters are above their average weight. When volatility is expected to increase and the market return is unexpectedly negative, only 44.6% of the letters decrease investment weights on average. Similar results are obtained when weights are measured relative to their average levels. These results support our conclusions that there is little or no information in either the changes or the level of investment weights about the direction of future market returns.

3.4. Survivorship

Survivorship in mutual funds has recently been studied in detail by Brown, Goetzmann, Ibbotson, and Ross (1992) and Brown and Goetzmann (1995). For example, research in mutual fund performance evaluation often studies funds that have survived over the evaluation horizon. Given that newsletters are added on the day Hulbert first receives the letter and no data are deleted when a newsletter ceases to exist, there are no major survivorship biases in our sample. Interestingly, in the newsletter sample, tenure is not related to performance (see Table 3). Mean investment letter returns are presented by the

Table 3
Survivorship

A: Survivorship and performance. The top portion of the panel shows the average annual return for portfolios of newsletters, where the portfolios are based on the number of years that a newsletter exists in the sample. The bottom portion of the panel shows the returns for portfolios of letters, conditioned on whether the newsletter is still active in the last month of the sample (i.e., in December 1992).

Number of year in existence	Mean return
1	13.0%
2	13.2
3	11.8
4	15.3
5	12.9
6	12.8
7	10.9
8	14.1
9	16.9
10	13.7
11	11.9

Mean annual return for letters	Return
existing end-of-year 1992	12.5%
not existing end-of-year 1992	14.9%

B: Predicting the probability of survivorship. This panel contains the results from a logistic regression determining the characteristics of letters which cease to exist. The dependent variable is equal to one if the letter does not survive in period $t + 1$ and is equal to zero if the letter does survive. The explanatory variables are all measured in period t and are the number of months out of the last 12 that the newsletter correctly anticipated the direction of the market (*HIT*), the percent out of the last five years that the newsletter had a larger return than a volatility-matched passive portfolio of cash and equity (*PCTEFF*), the amount the letter's return was above that for the volatility-matched portfolio (*RETEFF*), and the letter's raw return over the previous year (*RETURN*). The regression standard errors are robust with respect to heteroskedasticity and serial correlation.

Variable	Coefficient	t-score
CONSTANT	− 2.00	5.40
HIT	− 0.14	4.62
PCTEFF	− 1.06	1.91
RETEFF	0.34	0.54
RETURN	0.67	0.25

number of years in existence in panel A of Table 3. There is little difference between the performance of letters that were in existence for one year and those that survived ten years. In addition, the average returns of the letters existing in December 1992 is 12.5%, which is less than the 14.9% for the letters that failed to survive through the last month in our sample.

The second panel of Table 3 attempts to predict the probability of survivorship. We run a logistic regression determining the characteristics of the letters that cease to exist. The dependent variable is equal to one if the letter ceases to exist in period $t + 1$ and zero otherwise. The explanatory variables are all measured in period t. The variables are the number of months out of the last 12 that the newsletter correctly anticipated the direction of the market (*HIT*), the percent out of the last five years that the newsletter had a larger return than a portfolio of cash and equity with the same standard deviation as the newsletter (*PCTEFF*), the amount the letter's return was above that for a volatility-matched portfolio (*RETEFF*), and the letter's raw return over the previous year (*RETURN*).

The results in Table 2 indicate that two of these variables significantly predict whether the newsletter will cease to exist. Lower hit rates greatly increase the probability of dropping out (heteroskedasticity-consistent t-ratio of 4.6). If the letter return is often less than a volatility-matched passive portfolio over the past five years, this also increases the probability of ceasing to exist (t-ratio of 1.91).

3.5. Newsletter disagreement, volume, and market volatility

Thus far we have examined whether investment newsletter recommendations contain information about future market returns. We now turn our attention to examining whether newsletters contain information about market volume and volatility. In particular, we test whether the degree of disagreement contains information about future market volatility and trading volume. Shalen (1993) presents a model in which dispersion in the agents' forecasts induces trading. Her model predicts that increased dispersion causes increased trading volume and increased volatility. Our data provide an ideal setting to test these predictions. Theoretical models have also been proposed that link trading volume, volatility of price changes, and agents' forecasts. For example, the Harris and Raviv (1993) model has implications about changes in mean forecasted returns and volume. Our data allows us to test these implications.[3]

Our sample consists of recommendations of asset allocation weights, not market forecasts. However, it is possible to infer the newsletters' forecasts of the market return. If we assume that newsletter subscribers have negative exponential utility functions and returns are normally distributed, the expected portfolio return will be proportional to the conditional variance of the market return multiplied by the equity weight. We solve for the proportionality coefficient, which is the relative risk aversion, λ_i. Assuming that each newsletter agrees on

[3]Tauchen and Pitts (1983) and Gallant, Tauchen, and Rossi (1992a,b) examine the volume and volatility relation for the market as a whole. Survey predictions of the market and examined in Cowles (1933), Lakonishok (1980), Brown and Maital (1981), Pearce (1984), and Dokko and Edelstein (1989). The only study that examines disagreement is Frankel and Froot (1990).

the variance of the market returns, the variance of the conditionally expected returns, and the unconditional mean return, we solve for

$$\hat{\lambda}_i = \frac{\mathrm{E}[r_i]}{(\mathrm{var}[r_i] - \mathrm{var}[\mathrm{E}\{r_i|\mathrm{Z}\}])\mathrm{E}[w_i]}, \tag{4}$$

where r_i is the newsletter return. The risk aversion coefficients range from 2.6 to 24.5. The mean (median) risk aversion coefficient of 9.83 (9.33) is consistent with the estimates presented in Campbell (1987) and Harvey (1989). To insure that the average returns and variances are meaningful, we require a letter to have at least four years of monthly data. While this induces some survivorship bias, it ensures that the returns span the average length of one business cycle.

To infer the time series of predicted returns for each newsletter, we calculate:

$$\mathrm{E}_t[r_{i,t+1}] = \hat{\lambda}_i \mathrm{E}_t[\sigma^2_{t+1}] w_{i,t+1}\,. \tag{5}$$

The constant risk aversion is multiplied by both the time-varying expected volatility proxy and the time-varying investment weights. We assume that all investors have the same forecast of volatility. The expected volatility, $\mathrm{E}_t[\sigma^2_{t+1}]$, is generated from a sequence of out-of-sample GARCH (1,1) forecasts beginning in May 1982. This date coincides with the first month of futures trading on the S&P 500.

Table 4 presents contemporaneous correlations between dispersion, volatility (both realized and expected), trading volume, and the change in the aggregated newsletter forecasted return. Dispersion is defined as the standard deviation of the newsletters' expected returns. This standard deviation is calculated each month. Two measures of market volatility are presented. The first is the ex post volatility of the daily S&P 500 returns within a month. The second volatility is the implied volatility on the S&P 100 index.[4] We use this volatility to measure 'expected' volatility. Below the diagonal, the correlations of the levels are displayed. Above the diagonal, the correlations of the first differences of the variables are presented. In level form, all of the variables are positively correlated. These are contemporaneous correlations, i.e., both are measured at time t. However, the forecasted returns used in the dispersion calculation are strictly based on information available at $t - 1$.

Fig. 5 summarizes the time-series patterns in these measures. Panel A shows the 32% correlation between dispersion and volume per share. Bessembinder, Chan, and Seguin (1996) also find a positive relation between divergence of

[4]See Harvey and Whaley (1992) for the methodology of constructing the implied volatilities. We use the Chicago Board of Options Exchange's Market Volatility Index. See Whaley (1993) for a description of how a basket of volatilities is combined into a single index. The time series properties of the index are examined by Fleming, Ostdiek, and Whaley (1995). We thank Barb Ostdiek for providing us with this data, which is available from January 1986.

Table 4

Correlation of forecast dispersion, market volatility, and trading activity

Correlations for variables in level form are below the diagonal and for the first difference of the variables are above the diagonal.

	Dispersion[a]	Realized volatility[b]	Implied volatility[c]	Volume per share[d]
Dispersion		0.497*	0.146	0.415*
Realized volatility	0.548*		− 0.459*	0.668*
Implied volatility	0.775*	0.325*		− 0.519*
Volume per share	0.325*	0.536	0.212**	
First absolute difference $E(r_{m,t+1})$				0.578*

[a]Dispersion measures the standard deviation over the cross-section of implied market return forecasts across nonmissing observations for a sample of 237 investment newsletters for each month in the period May 1982 through December 1992. The dispersion is from the period t newsletter forecasts, but each forecast is based on information available in period $t-1$.

[b]Monthly realized volatility is obtained by summing squared daily returns plus two times the autocovariance for daily returns for the near S&P 500 futures contract.

[c]Implied volatility is for the S&P 100 index and exists starting in January 1986.

[d]Volume per share is monthly NYSE volume for all shares divided by total number of shares outstanding.

[e]The absolute value of the first difference of the mean forecast is based on the average of the implied market forecasts made by the newsletters.

* indicates significant at $\alpha = 0.05$.
** indicates significant at $\alpha = 0.10$.

opinion and trading volume, although they use the open interest on the S&P 500 index futures as a proxy for divergence of opinion. The correlation between dispersion and both realized volatility (55%) and implied volatility (77%) is depicted in the next panels. These results are influenced by the GARCH volatility which enters each newsletter's forecast through (5). However, separate multiple regressions of realized volatility on fitted GARCH and the cross-sectional standard deviation of the investment weights (not reported) indicate that the dispersion in recommended weights has incremental explanatory power. We also link our analysis of market timing to dispersion. There is a weak positive correlation between dispersion and the percentage of newsletters moving weight in the correct direction. This suggests that in times of the greatest disagreement, the overall performance of the newsletters is marginally superior.

Table 4 also reports a test of one of the predictions of the Harris and Raviv (1993) model. They show that volume and forecast changes should be positively correlated because they are both driven by a third exogeneous factor, namely a signal. In support of their model, the contemporaneous correlation between

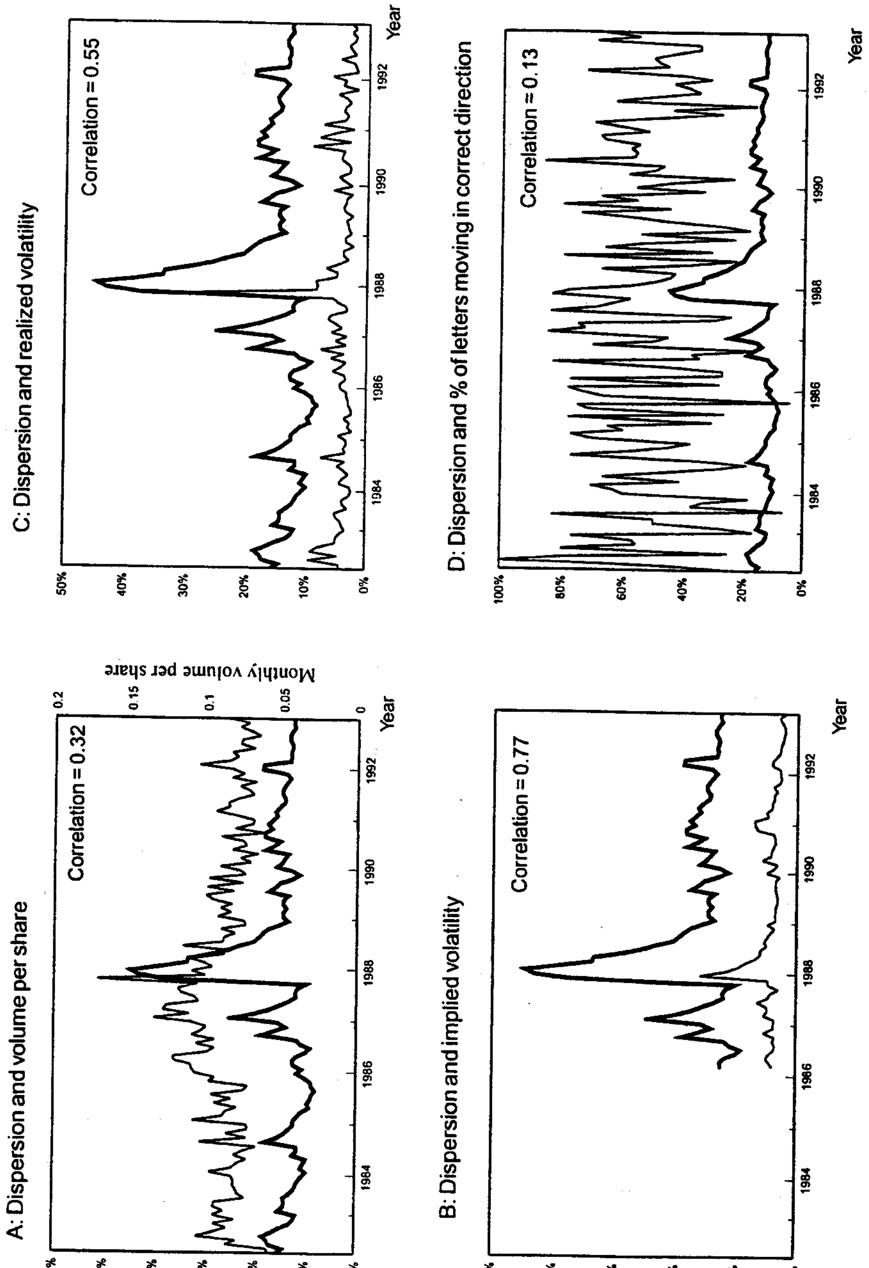


Fig. 5. Time-series relation between disagreement and volume, returns volatility, and correct changes in investment weights.

the change in the absolute value of the market forecast and volume is 58%, which is significant at the 1% level.

Consistent with the predictions of Shalen (1993), we also find (but do not report in detail) that changes in dispersion significantly predict changes in volatility and NYSE volume (adjusted by the total number of shares outstanding). These findings are robust to the choice of proxy for volatility. Hence, newsletter disagreement contains important economic information.

4. Conclusions

Our paper investigates the ability of newsletters to predict the direction of the market. In analyzing over 15,000 asset allocation recommendations for the 1980–1992 period, we find little evidence that recommended equity weights increase before future positive market returns or decrease before negative market returns. We argue that timing should be evaluated relative to the common-knowledge degree of predictability in the economy. Extra timing exists when a newsletter correctly anticipates the direction of the market the common expected return does not. We find no evidence that investment letters as a group have any knowledge over and above the common level of predictability.

While we find little evidence that investment newsletters as a group can time the market, we do identify a 'hot hands' phenomenon. We present some evidence that 'hot' newsletters' recommendations contain limited information about future market returns. However, we argue that the 'hot hands' phenomenon is fleeting. While some letters at certain times appear to have short-run insights, an investor cannot use a hot streak to identify a particular newsletter that will provide superior recommendations over the long term. Our Monte Carlo analysis indicates that the performance of investment newsletters is no better than, and potentially worse than, what would be expected from a set of letters that offer random recommendations.

While there is little information in the investment newsletters' opinions regarding stock market direction, we find that the degree of disagreement among letters predicts both realized and expected volatility as well as trading

Caption of Fig. 5 (continued)

Panel A shows the time series of the dispersion [thick line] of inferred market forecasts made by letters and volume per share on the NYSE [thin line] over the period May 1982 to December 1992. Panel B (C) shows dispersion [thick line] relative to implied (realized) market volatility [thin line]. The implied volatility series is only available from January 1986. Panel D show dispersion [thick line] relative to the percentage of letters changing their recommended equity weights in the same direction as the ensuing market movement.

volume. These results are consistent with the theoretical models proposed in Shalen (1993) and Harris and Raviv (1993).

References

Bansal, Ravi and Campbell R. Harvey, 1996, Performance in the presence of dynamic trading strategies, Unpublished working paper (Duke University, Durham, NC).

Bessembinder, Hendrick, Kalok Chan, and Paul J. Seguin, 1996, An empirical examination of information, differences of opinion, and trading activity, Journal of Financial Economics 40, 105–134.

Bollerslev, Tim, 1986, Generalized autoregressive conditional heteroskedasticity, Journal of Econometrics 31, 307–327.

Brown, Bryan W. and Shlomo Maital, 1981, What do economists know? An empirical test of experts' expectations, Econometrica 49, 491–504.

Brown, Stephen J. and William Goetzmann, 1995, Performance persistence, Journal of Finance 50, 679–698.

Brown, Stephen, J., William Goetzmann, Roger G. Ibbotson, and Stephen A. Ross, 1992, Survivorship bias in performance studies, Review of Financial Studies 5, 553–580.

Campbell, J.Y., 1987, Stock returns and the term structure, Journal of Financial Economics 18, 373–400.

Chen, Zhiwu and Peter J. Knez, 1996, Portfolio performance measurement: Theory and applications, Review of Financial Studies 51, 425–461.

Cowles, Alfred III, 1993, Can stock market forecasters forecast?, Econometrica 1, 309–324.

Dokko, Yoon and Robert H. Edelstein, 1989, How well do economists forecast stock market prices? A study of the Livingston surveys, American Economic Review 79, 865–871.

Engle, Robert F., 1982, Autoregressive conditional heteroskedasticity with estimates of U.K. inflation, Econometrica 50, 987–1008.

Fama, Eugene F. and Kenneth R. French, 1989, Business conditions and expected returns on stocks and bonds, Journal of Financial Economics 25, 23–50.

Ferson, Wayne E. and Rudi Schadt, 1996, Measuring fund strategy and performance in changing economic conditions, Journal of Finance, forthcoming.

Fleming, Jeffery, Barbara Ostdiek, and Robert E. Whaley, 1995, Predicting stock market volatility: A new measure, Journal of Futures Market, forthcoming.

Frankel, Jeffery and Kenneth Froot, 1990, Chartists, fundamentalists, and trading in the foreign exchange market, American Economic Review 80, 181–185.

Gallant, A. Ronald, George Tauchen, and Peter E. Rossi, 1992a, Nonlinear dynamic structures, Econometrica 61, 871–907.

Gallant, A. Ronald, George Tauchen, and Peter E. Rossi, 1992b, Stock prices and volume, Review of Financial Studies 5, 199–242.

Grinblatt, Mark and Sheridan Titman, 1989, Mutual fund performance: A analysis of quarterly portfolio holdings, Journal of Business 62, 393–416.

Harris, Milton and Arthur Raviv, 1993, Differences of opinion make a horse race, Review of Financial Studies 6, 473–506.

Harvey, Campbell R., 1989, Time-varying conditional covariances in tests of asset pricing model, Journal of Financial Economics 24, 289–317.

Harvey, Campbell R. and Robert E. Whaley, 1992, Market volatility prediction and the efficiency of the S&P 100 index option market, Journal of Financial Economics 31, 43–73.

Henriksson, Roy D. and Robert C. Merton, 1981, On market timing and investment performance: II. Statistical procedures for evaluation skills, Journals of Business 54, 513–553.

Lakonishok, Josef, 1980, Stock market return expectations: Some general properties, Journal of Finance 35, 921–930.

Pearce, Douglas K., 1984, An empirical analysis of expected stock returns, Journal of Money, Credit and Banking 16, 317–327.

Shalen, C., 1993, Volume, volatility and the dispersion of beliefs, Review of Financial Studies 6, 405–434.

Tauchen, George and Mark Pitts, 1983, The price variability–volume relationship on speculative markets, Econometrica 51, 485–505.

Whaley, Robert E., 1993, Derivatives on market volatility: Hedging tools long overdue, Journal of Derivatives, 71–84.